Preface

The area of stochastic programming was created in the middle of the last century, following fundamental achievements in linear and nonlinear programming. While it has been quickly realized that the presence of uncertainty in optimization models creates a need for new problem formulations, many years have passed until the basic stochastic programming models have been formulated and analyzed. Today, stochastic programming theory offers a variety of models to address the presence of random data in optimization problems: chance-constrained models, two- and multi-stage models, models involving risk measures. New problem formulations appear almost every year and this variety is one of the strengths of the field.

Stochastic programming can be quite involved, starting with sophisticated modeling and is based on advanced mathematical tools such as nonsmooth calculus, abstract optimization, probability theory and statistical techniques. One of the objectives of this Handbook is to bring these techniques together and to show how they can be used to analyze and solve stochastic programming models.

Because of the inherent difficulty of stochastic optimization problems, it took a long time until efficient solution methods have been developed. In the last two decades a dramatic change in our abilities to solve stochastic programming problems took place. It is partially due to the progress in large scale linear and nonlinear programming, in nonsmooth optimization and integer programming, but mainly it follows the development of techniques exploiting specific properties of stochastic programming problems. Computational advances are also due to modern parallel processing technology. Nowadays we can solve stochastic optimization problems involving tens of millions of variables and constraints.

Our intention was to bring together leading experts in the most important sub-fields of stochastic programming to present a rigorous overview of basic models, methods, and applications of stochastic programming. We hope that this Handbook will prove useful to researchers, students, engineers and economists, who encounter in their work optimization problems involving uncertainty. We also hope that our work will encourage many to undertake research in this exciting and practically important field.

We want to thank all the Authors involved in this project for their contributions. We also want to thank Darinka Dentcheva, Shabbir Ahmed, Tito Homem-de-Mello and Anton Kleywegt, who have helped us to review and improve several chapters of this Handbook.

Andrzej Ruszczyński and Alexander Shapiro December 2002.

Contents

A. Ruszczyński and A. Shapiro, Eds., *Handbooks in OR & MS, Vol. 10*

Chapter 1

Stochastic Programming Models

Andrzej Ruszczyński

Department of Management Science and Information Systems, Rutgers University, 94 Rockefeller Rd, Piscataway, NJ 08854, USA

Alexander Shapiro

School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA

Abstract

In this introductory chapter we discuss some basic approaches to modeling of stochastic optimization problems. We start with motivating examples and then proceed to formulation of linear, and later nonlinear, two stage stochastic programming problems. We give a functional description of two stage programs. After that we proceed to a discussion of multistage stochastic programming and its connections with dynamic programming. We end this chapter by introducing robust and min–max approaches to stochastic programming. Finally, in the appendix, we introduce and briefly discuss some relevant concepts from probability and optimization theories.

Key words: Two stage stochastic programming, expected value solution, stochastic programming with recourse, nonanticipativity constraints, multistage stochastic programming, dynamic programming, chance constraints, value at risk, scenario tree, robust stochastic programming, mean–risk models.

1 Introduction

1.1 Motivation

Uncertainty is the key ingredient in many decision problems. Financial planning, airline scheduling, unit commitment in power systems are just few examples of areas in which ignoring uncertainty may lead to inferior or simply wrong decisions. Often there is a variety of ways in which the uncertainty can be

formalized and over the years various approaches to optimization under uncertainty were developed. We discuss a particular approach based on probabilistic models of uncertainty. By averaging possible outcomes or considering probabilities of events of interest we can define the objectives and the constraints of the corresponding mathematical programming model.

To formulate a problem in a consistent way, a number of fundamental assumptions need to be made about the nature of uncertainty, our knowledge of it, and the relations of decisions to the observations made. In order to motivate the main concepts let us start by discussing the following classical example.

Example 1 (Newsvendor Problem). A newsvendor has to decide about the quantity x of newspapers which he purchases from a distributor at the beginning of a day at the cost of c per unit. He can sell a newspaper at the price s per unit and unsold newspapers can be returned to the vendor at the price of r per unit. It is assumed that $0 \leq r < c < s$. If the demand D, i.e., the quantity of newspapers which he is able to sell at a particular day, turns out to be greater than or equal to the order quantity x, then he makes the profit $sx - cx = (s-c)x$, while if D is less than x, his profit is $sD + r(x-D) - cx = (r-c)x + (s-r)D$. Thus the profit is a function of x and D and is given by

$$F(x, D) = \begin{cases} (s-c)x, & \text{if } x \leq D, \\ (r-c)x + (s-r)D, & \text{if } x > D. \end{cases} \tag{1.1}$$

The objective of the newsvendor is to maximize his profit. We assume that the newsvendor is very intelligent (he has Ph.D. degree in mathematics from a prestigious university and sells newspapers now), so he knows what he is doing. The function $F(\cdot, D)$ is a continuous piecewise linear function with positive slope $s-c$ for $x < D$ and negative slope $r-c$ for $x > D$. Therefore, if the demand D is known, then the best decision is to choose the order quantity $x^* = D$. However, in reality D is not known at the time the order decision has to be made, and consequently the problem becomes more involved.

Since the newsvendor has this job for a while he collected data and has quite a good idea about the probability distribution of the demand D. That is, the demand D is viewed now as a *random variable* with a known, or at least well estimated, probability distribution measured by the corresponding cumulative distribution function (cdf) $G(w) := \mathbb{P}(D \leq w)$. Note that since the demand cannot be negative, it follows that $G(w) = 0$ for any $w < 0$. By the Law of Large Numbers the average profit over a long period of time tends to the expected value

$$\mathbb{E}[F(x, D)] = \int_0^\infty F(x, w)\, \mathrm{d}G(w).$$

Therefore, from the statistical point of view it makes sense to optimize the objective function on *average*, i.e., to maximize the expected profit $\mathbb{E}[F(x, D)]$. This leads to the following stochastic programming problem[1]

$$\max_{x \geq 0} \{f(x) := \mathbb{E}[F(x, D)]\}. \tag{1.2}$$

Note that we treat here x as a continuous rather than integer variable. This makes sense if the quantity of newspapers x is reasonably large.

In the present case it is not difficult to solve the above optimization problem in a closed form. Let us observe that for any $D \geq 0$, the function $F(\cdot, D)$ is concave (and piecewise linear). Therefore, the expected value function $f(\cdot)$ is also concave. Suppose for a moment that $G(\cdot)$ is continuous at a point $x > 0$. Then

$$f(x) = \int_0^x [(r-c)x + (s-r)w] \, dG(w) + \int_x^\infty (s-c)x \, dG(w).$$

Using integration by parts it is possible to calculate then that

$$f(x) = (s-c)x - (s-r)\int_0^x G(w) \, dw. \tag{1.3}$$

The function $f(\cdot)$ is concave, and hence continuous, and therefore formula (1.3) holds even if $G(\cdot)$ is discontinuous at x. It follows that $f(\cdot)$ is differentiable at x iff (that is, if and only if) $G(\cdot)$ is continuous at x, in which case

$$f'(x) = s - c - (s-r)G(x). \tag{1.4}$$

Consider the inverse $G^{-1}(\alpha) := \min\{x : G(x) \geq \alpha\}$ function[2] of the cdf G, which is defined for $\alpha \in (0, 1)$. Since $f(\cdot)$ is concave, a necessary and sufficient condition for $x^* > 0$ to be an optimal solution of problem (1.2) is that $f'(x^*) = 0$, provided that $f(\cdot)$ is differentiable at x^*. Note that because $r < c < s$, it follows that $0 < (s-c)/(s-r) < 1$. Consequently, an optimal solution of (1.2) is given by

$$x^* = G^{-1}\left(\frac{s-c}{s-r}\right). \tag{1.5}$$

This holds even if $G(\cdot)$ is discontinuous at x^*. It is interesting to note that $G(0)$ is equal to the probability that the demand D is zero, and

[1] The notation "$:=$" means equal by definition.

[2] Recall that $G^{-1}(\alpha)$ is called the *α-quantile* of the cdf G.

hence if this probability is positive and $(s-c)/(s-r) \leq G(0)$, then the optimal solution $x^* = 0$.

Clearly the above approach explicitly depends on the knowledge of the probability distribution of the demand D. In practice the corresponding cdf $G(\cdot)$ is never known exactly and could be approximated (estimated) at best. In the present case the optimal solution is given in a closed form and therefore its dependence on $G(\cdot)$ can be easily evaluated. It is well known that α-quantiles are robust (stable) with respect to small perturbations of the corresponding cdf $G(\cdot)$, provided that α is not too close to 0 or 1. In general, it is important to investigate sensitivity of a considered stochastic programming problem with respect to the assumed probability distributions.

The following deterministic optimization approach is also often used for decision making under uncertainty. The random variable D is replaced by its mean $\mu = \mathbb{E}[D]$, and then the following deterministic optimization problem is solved:

$$\operatorname*{Max}_{x \geq 0} \; F(x, \mu). \tag{1.6}$$

A resulting optimal solution $\overline{x}$ is sometimes called the expected value solution. In the present example, the optimal solution of this deterministic optimization problem is $\overline{x} = \mu$. Note that the mean solution $\overline{x}$ can be very different from the solution x^* given in (1.5). It is well known that the quantiles are much more stable to variations of the cdf G than the corresponding mean value. Therefore, the optimal solution x^* of the stochastic optimization problem is more robust with respect to variations of the probability distributions than an optimal solution $\overline{x}$ of the corresponding deterministic optimization problem. This should be not surprising since the deterministic problem (1.6) can be formulated in the framework of the stochastic programming problem (1.2) by considering the trivial distribution of D being identically equal to μ.

For any x, $F(x, D)$ is concave in D. Therefore the following Jensen's inequality holds:

$$F(x, \; \mu) \geq \mathbb{E}[F(x, \; D)].$$

Hence

$$\max_{x \geq 0} \; F(x, \mu) \geq \max_{x \geq 0} \; \mathbb{E}[F(x, D)].$$

Thus the optimal value of the deterministic optimization problem is biased upward relative to the optimal value of the stochastic optimization problem. This should be also not surprising since the optimization problem

(1.6) is "too optimistic" in the sense that it does not take into account possible variability of the demand D.

Another point which is worth mentioning is that by solving (1.2) the newsvendor tries to optimize the profit on average. However, for a particular realization of the demand D, on a particular day, the profit $F(x^*, D)$ could be very different from the corresponding expected value $f(x^*)$. This may happen if $F(x^*, D)$, considered as a random variable, has a large variability which could be measured by its variance $\mathbb{V}\mathrm{ar}[F(x^*, D)]$. Therefore, if the newsvendor wants to hedge against such variability he may consider the following optimization problem

$$\max_{x \geq 0} \{f_\kappa(x) := \mathbb{E}[F(x, D)] - \kappa \mathbb{V}\mathrm{ar}[F(x, d)]\}. \tag{1.7}$$

The coefficient $\kappa \geq 0$ represents the weight given to the conservative part of the decision. If κ is "large", then the above optimization problem tries to find a solution with minimal profit variance, while if $\kappa = 0$, then problem (1.7) coincides with problem (1.2). Since

$$\mathbb{V}\mathrm{ar}[F(x, D)] = \mathbb{E}[F(x, D)^2] - [\mathbb{E}F(x, D)]^2,$$

from a mathematical point of view problem (1.7) is similar to the expected value problem (1.2). Note, however, that the additional (variance) term in (1.7) destroys the convexity of the optimization problem (see Section 4 for a further discussion).

The newsvendor may be also interested in making at least a specified amount of money, b, on a particular day. Then it would be reasonable to consider the problem of purchasing the minimum number of newspapers, x, under the condition that the probability of making at least b is not less than $1 - \alpha$, where $\alpha \in (0, 1)$ is fixed. Such a problem can be formulated in the form

$$\text{Min } x \tag{1.8}$$

$$\text{s.t.} \quad \mathbb{P}\{F(x, D) \geq b\} \geq 1 - \alpha. \tag{1.9}$$

The newsvendor can solve this problem, too (remember that he is really smart). It is clear that the following inequality should be satisfied

$$(s - c)x \geq b, \tag{1.10}$$

since otherwise there is no way of making b. For a fixed x satisfying this condition, the profit $F(x, D)$ is a nondecreasing function of the demand D.

Therefore

$$\mathbb{P}\{F(x, D) \geq b\} = \mathbb{P}\{D \geq d(x, b)\},$$

where (after straightforward calculations)

$$d(x, b) = \frac{b + (c - r)x}{s - r}.$$

It follows from (1.9) that $d(x, b) \leq G^{-1}(\alpha)$, which can be written as

$$b + (c - r)x \leq (s - r)G^{-1}(\alpha). \tag{1.11}$$

It is clear that the solution can exist iff the constraints (1.10)–(1.11) are consistent, that is, if

$$b \leq (s - c)G^{-1}(\alpha). \tag{1.12}$$

Therefore, we obtain that problem (1.8)–(1.9) is feasible iff (1.12) holds, in which case it has the optimal solution

$$\hat{x} = \frac{b}{s - c}. \tag{1.13}$$

1.2 The basic model

Let us formalize optimization problems of the type discussed in the newsvendor example. To this end we use the following notation and terminology. By $\mathcal{X}$ we denote the space of decision variables. In most applications considered in this book $\mathcal{X}$ can be identified with a finite dimensional vector space $\mathbb{R}^n$. It is assumed that there is a given set $X \subset \mathcal{X}$ of feasible (or permissible) decisions and an (objective) function $F(x, \omega)$ of decision vector $x \subset X$ and random element ω. In an abstract setting we consider ω as an element of a sample space Ω equipped with a sigma algebra $\mathcal{F}$. In typical applications considered in this book, the involved random data is formed by a finite number of parameters. Consequently, the objective function is given in the form $F(x, \omega) := V(x, \xi(\omega))$, where $\xi(\omega)$ is a finite dimensional random vector and $V(x, \xi)$ is a function of two vector variables x and ξ.

Of course, the mathematical programming problem of minimization (or maximization) of $F(x, \omega)$ subject to $x \in X$ depends on ω and does not make much sense. For different realizations of the random parameters one would obtain different optimal solutions without any insight which one is "better"

than the others. A way of dealing with that is to optimize the objective function on *average*. This leads to the following mathematical programming problem

$$\operatorname*{Min}_{x\in X}\{f(x) := \mathbb{E}[F(x,\omega)]\}. \tag{1.14}$$

The above formulation of a stochastic programming problem assumes implicitly that the expected value is taken with respect to a known probability distribution (measure) P on $(\Omega, \mathcal{F})$ and that the expected value operator

$$\mathbb{E}[F(x,\omega)] = \int_{\Omega} F(x,\omega)\, \mathrm{d}P(\omega) \tag{1.15}$$

is well defined. We refer to the function $f(x)$, defined in (1.14), as the *expectation* or *expected value* function. Note that we will have to deal with extended real valued functions. That is, the function $F(x,\omega)$ (as well as its expectation) is allowed to take values $+\infty$ or $-\infty$. The precise meaning of the involved concepts is discussed in the Appendix (Section 5).

1.3 Modeling the constraints

In (1.14) we have assumed that we have an explicit description of the feasible set X. For example, the feasible set X can be written in a standard mathematical programming formulation as follows

$$X := \{x \in X_0 \colon g_i(x) \le 0,\ i = 1,\dots,m\}, \tag{1.16}$$

where X_0 is a convex subset of $\mathcal{X} := \mathbb{R}^n$ and $g_i(x)$ are real valued functions.

When the uncertain quantities enter the 'raw' constraints of our background model,

$$G_i(x,\omega) \le 0, \quad i = 1,\dots,m, \tag{1.17}$$

we need to specify what we mean by 'feasibility'. Some values of x may satisfy (1.17) for some ω and violate these conditions for other ω. Often it is unrealistic to require that constraints (1.17) should hold for all $\omega \in \Omega$. In our newsvendor example, for instance, the requirement to make at least profit b can hardly be satisfied for all realizations of the demand D.

Several approaches can be used to introduce a meaningful notion of feasibility in this context. One of them is to consider the expected values,

$$g_i(x) := \mathbb{E}[G_i(x,\omega)], \quad i = 1,\dots,m, \tag{1.18}$$

as constraint functions in (1.16).

Expected value constraints usually occur in situations when we have, in fact, several objectives, and we put some of them into the constraints, as in the example below.

Example 2 (Reservoir Capacity). Consider the system of two reservoirs (Fig. 1), whose objective is to retain the flood in the protected area. The flood is produced by two random inflows, ξ_1 and ξ_2. Flood danger occurs once a year, say, and ξ_1, ξ_2 appear simultaneously. The damage from flood of size $y \geq 0$ is modeled as a convex nondecreasing function $L(y)$, where $L(0) = 0$. Our objective is to determine the reservoir capacities, x_1 and x_2, so that the expected damage from the flood is below some specified limit b, and the cost of the reservoirs, $f(x_1, x_2)$ is minimized.

The size of the flood is random and is given by the expression

$$y = \max\{0, \xi_1 + \xi_2 - x_1 - x_2, \xi_2 - x_2\}.$$

Our problem takes on the form

$$\begin{aligned} &\text{Min } f(x_1, x_2) \\ &\text{s.t.} \quad \mathbb{E}[L(\max\{0, \xi_1 + \xi_2 - x_1 - x_2, \xi_2 - x_2\})] \leq b. \\ &\qquad x_1 \geq 0, x_2 \geq 0. \end{aligned} \tag{1.19}$$

It would be an error to replace the random inflows in this problem by their expected values, μ_1 and μ_2. By Jensen's inequality we have

$$\begin{aligned} &L(\max\{0, \mu_1 + \mu_2 - x_1 - x_2, \mu_2 - x_2\}) \\ &\quad \leq \mathbb{E}[L(\max\{0, \xi_1 + \xi_2 - x_1 - x_2, \xi_2 - x_2\})], \end{aligned}$$

and the difference may be large, even for a linear function $L(\cdot)$. As a result, the expected losses from a flood may be much higher than foreseen by a naive deterministic model.

Another way to define the feasible set is to use constraints on the probability of satisfying (1.17):

$$\mathbb{P}\{G_i(x, \omega) \leq 0\} \geq 1 - \alpha, \quad i = 1, \ldots, m, \tag{1.20}$$

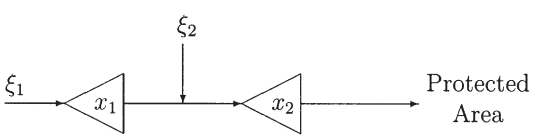

Fig. 1. The water reservoir system.

with some fixed $\alpha \in (0,1)$ (as in our newsvendor example). Such constraints are called *probabilistic* or *chance* constraints.[3]

For a set A we denote by $\mathbf{1}_A(\cdot)$ its characteristic function,

$$\mathbf{1}_A(t) := \begin{cases} 1, & \text{if } t \in A, \\ 0, & \text{if } t \notin A. \end{cases} \tag{1.21}$$

Then (1.20) can be written as the expected value constraints

$$\mathbb{E}[\mathbf{1}_{(-\infty,0)}(G_i(x,\omega))] \geq 1-\alpha, \quad i = 1,\ldots,m. \tag{1.22}$$

Note, however, that the discontinuity of the characteristic function makes such constraints very specific and different from the 'standard' expected value constraints.

Following is an example where probabilistic constraints appear in a natural way.

Example 3 (Value at Risk). Suppose that there are n investment opportunities, with random returns $R_1,\ldots,R_n$ in the next year. We have a practically unlimited initial capital and our aim is to invest some of it in such a way that the expected value of our investment after a year is maximized, under the condition that the chance of losing no more than some fixed amount $b>0$ is at least $1-\alpha$, where $\alpha \in (0,1)$. Such a requirement is called the *Value at Risk* constraint.

Let $x_1,\ldots,x_n$ be the amounts invested in the n opportunities. The net increase of the value of our investment after a year is $G(x,R) = \sum_{i=1}^n R_i x_i$. Our problem takes on the form of a probabilistic constrained stochastic program:

$$\begin{aligned} &\text{Max} \sum_{i=1}^n \mu_i x_i \\ &\text{s.t.} \quad \mathbb{P}\left\{\sum_{i=1}^n R_i x_i \geq -b\right\} \geq 1-\alpha, \\ &\qquad x \geq 0, \end{aligned} \tag{1.23}$$

where $\mu_i = \mathbb{E}[R_i]$. Note that for the sake of simplicity we do not impose here the constraint $x_1+\cdots+x_n = W_0$, where W_0 is the total invested amount, as compared with the example of financial planning (Example 7) discussed later.

[3] In the extreme case when $\alpha = 0$, conditions (1.20) mean that constraints $G_i(x,\omega) \leq 0$, $i=1,\ldots,m$, should hold for a.e. $\omega \in \Omega$.

If the returns have a joint normal distribution with the covariance matrix Σ, the distribution of the profit (or loss) is normal, too, with the expected value $\mu^T x$, and variance $x^T \Sigma x$. Consequently, $(G(x, R) - \mu^T x)/\sqrt{x^T \Sigma x}$ has the standard normal distribution (i.e., normal distribution with mean zero and variance one). Our probabilistic constraint is therefore equivalent to the inequality

$$\frac{b + \mu^T x}{\sqrt{x^T \Sigma x}} \geq z_\alpha,$$

where z_α is the $(1-\alpha)$-quantile of the standard normal variable. If $\alpha \leq 1/2$ then $z_\alpha \geq 0$. After elementary manipulations we obtain the following convex programming equivalent of problem (1.23)

$$\begin{aligned} &\text{Max } \mu^T x \\ &\text{s.t.} \quad z_\alpha \sqrt{x^T \Sigma x} - \mu^T x \leq b, \\ &\qquad x \geq 0. \end{aligned} \tag{1.24}$$

If we ignore the nonnegativity constraint on x we can solve this problem analytically. Indeed, $x = 0$ is a feasible solution and both functions are positively homogeneous in x, so either the probabilistic constraint has to be satisfied as an equality or the problem is unbounded. Let $\lambda \geq 0$ be the Lagrange multiplier associated with this constraint. We obtain the equation

$$(1 + \lambda)\mu - \frac{\lambda z_\alpha \Sigma x}{\sqrt{x^T \Sigma x}} = 0.$$

From here we deduce that there must exist a scalar t such that $\Sigma x = t\mu$. We assume that the matrix Σ is nonsingular and $\mu \neq 0$. Substitution to the constraint yields (after simple calculations) $t = b/\rho(z_\alpha - \rho)$ and $\lambda = (z_\alpha/\rho - 1)^{-1}$, with $\rho := \sqrt{\mu^T \Sigma^{-1} \mu}$ (note that Σ^{-1} is positive definite and hence $\mu^T \Sigma^{-1} \mu$ is positive). If $\rho \geq z_\alpha$, then the problem is unbounded, i.e., its optimal value is $+\infty$. If $\rho < z_\alpha$, then the vector

$$\hat{x} := \frac{b}{\rho(z_\alpha - \rho)} \Sigma^{-1} \mu$$

is the solution to the problem without sign restrictions on x. If, in addition, $\Sigma^{-1}\mu \geq 0$, then the vector $\hat{x}$ solves our original problem. Otherwise, numerical methods of convex programming are needed to find the optimal solution.

In many practical situations, though, the returns are not jointly normally distributed, and even the single Value at Risk constraint, like the one analyzed here, may create significant difficulties.

Let us now assume that our planning horizon is T years, and let $R_1(t), \ldots, R_n(t)$ be the random investment returns in years $t = 1, \ldots, T$. We want to maximize the expected value of our investment after T years, under the condition that with probability at least $1-\alpha$ the value of our investment will never drop by more than b from the initial amount invested. We do not want to re-allocate our investment, we just want to invest once and then watch our wealth grow (hopefully).

Let $x_1, \ldots, x_n$ be the amounts invested in the n opportunities. The net change in the value of our investment in year t is

$$G(x, R, t) = \sum_{i=1}^{n} S_i(t) x_i,$$

where $S_i(t) := \Pi_{\tau=1}^{t}(1 + R_i(\tau)) - 1$ is the compounded return of investment i up to year t. Denoting $\mu_i := \mathbb{E}[S_i(T)]$, our problem takes on the form:

$$\begin{aligned} &\operatorname*{Max}_{x \geq 0} \sum_{i=1}^{n} \mu_i x_i \\ &\text{s.t.} \quad \mathbb{P}\big\{G(x, R, t) \geq -b,\ t = 1, \ldots, T\big\} \geq 1 - \alpha. \end{aligned} \tag{1.25}$$

This is an example of a problem with a *joint probabilistic constraint*, which is different from imposing the constraints $\mathbb{P}\{G(x, R, t) \geq -b\} \geq 1-\alpha,\ t = 1, \ldots, T$, requiring that for each year the probability of losing no more than b is $1-\alpha$ or higher. A joint probabilistic constraint can be formally treated as a constraint for one function, defined as the worst case among the individual constraints. In our example we may define $G(x, R) := \min_{1 \leq t \leq T} G(x, R, t)$ and require that

$$\mathbb{P}\{G(x, R) \geq -b\} \geq 1 - \alpha. \tag{1.26}$$

Such constraints may be difficult to handle, both theoretically and computationally.

2 Two-stage models

2.1 The linear model

We can view the decision problem which the newsvendor faces in Example 1 as two stage. In the morning, before a realization of the demand D is known,

he has to decide about the quantity x of newspapers which he purchases for that day. By the end of the day when value of D becomes known, he optimizes his behavior by selling as many newspapers as possible. Although simple, his second stage decision can be also formulated as an optimization problem. His second stage decision variables can be defined as the quantity y which he sells at price s, and the quantity z which he returns at price r. Then, given a value of the first stage decision variable x and a realization of the demand D, the second stage problem consists of maximizing the profit and can be written as follows

$$\begin{aligned} &\operatorname*{Max}_{y,z} \; sy + rz \\ &\text{subject to} \\ &y \le D, \; y + z \le x, y \ge 0, z \ge 0. \end{aligned}$$

The optimal solution of the above problem is $y^* = \min\{x, D\}$, $z^* = \max\{x - D, 0\}$, and its optimal value is the profit $F(x, D)$ defined in (1.1).

This is the basic idea of a two stage process. At the first stage, before a realization of the corresponding random variables becomes known, one chooses the first stage decision variables to optimize the expected value of an objective function which in turn is the optimal value of the second stage optimization problem. A *two-stage stochastic linear program* can be written as follows

$$\begin{aligned} &\operatorname{Min}_x \, c^T x + \mathbb{E}[Q(x, \xi(\omega))] \\ &\text{s.t.} \quad Ax = b, x \ge 0, \end{aligned} \tag{2.2}$$

where $Q(x, \xi)$ is the optimal value of the second stage problem

$$\begin{aligned} &\operatorname{Min}_y \, q^T y \\ &\text{s.t.} \quad Tx + Wy = h, \; y \ge 0. \end{aligned} \tag{2.3}$$

Here x and y are vectors of first and second stage decision variables, respectively. The second stage problem depends on the data $\xi := (q, h, T, W)$, some (all) elements of which can be random. Therefore we view $\xi = \xi(\omega)$ as a *random vector*. The expectation in (2.2) is taken with respect to the probability distribution of $\xi(\omega)$, which is supposed to be known. The matrices T and W are called the *technology* and *recourse* matrices, respectively. If the matrix W is fixed (not random), the above two-stage problem is called the problem with *fixed recourse*. In a sense the second stage problem (2.3) can be viewed as a penalty term for violation of the constraint $Tx = h$, hence is the name "with recourse".

For any x and ξ the function $Q(x, \xi)$, although not given explicitly, is a well defined extended real valued function: it takes the value $+\infty$ if the feasible set of the second stage problem (2.3) is empty, and the value $-\infty$ if the second

stage problem is unbounded from below. As it is discussed in Section 5.2, it should be verified that the expected value in (2.2) is well defined. It is worthwhile to note at this point that problem (2.2) is a particular case of the stochastic programming problem (1.14) with $F(x,\omega) := c^T x + Q(x,\xi(\omega))$ and $X := \{x\colon Ax = b, x \geq 0\}$.

By the definition of the function $Q(x,\xi)$ we have that it can be written in the form $Q(x,\xi) = \mathcal{Q}(h - Tx)$, where

$$\mathcal{Q}(\chi) := \inf\{q^T y\colon Wy = \chi, y \geq 0\}. \tag{2.4}$$

By the duality theory of linear programming the optimal value $\mathcal{Q}(\chi)$ of the linear program in the right hand side of (2.4) is equal to $\sup\{\pi^T\chi\colon W^T\pi \leq q\}$, unless both systems: $Wy = \chi, y \geq 0$ and $W^T\pi \leq q$, are infeasible. Consequently,

$$Q(x,\xi) = \sup\{\pi^T(h - Tx)\colon W^T\pi \leq q\}. \tag{2.5}$$

The feasible set $\{\pi\colon W^T\pi \leq q\}$ of the dual problem is convex polyhedral. Therefore, for any realization of random data ξ, the function $Q(\cdot,\xi)$ is convex piecewise linear. Chapter "Optimality and Quality in Stochastic Programming" of this book provides a detailed analysis of the properties of $Q(\cdot,\xi)$ and of its expected value.

2.2 *The case of discrete distributions*

There are equivalent formulations of the two-stage linear recourse problem (2.2)–(2.3) which are useful in different situations. In order to simplify the presentation and to defer technical details let us assume now that the random data have a discrete distribution with a finite number K of possible realizations $\xi_k = (q_k, h_k, T_k, W_k)$, called *scenarios*, with the corresponding probabilities p_k. In that case $\mathbb{E}[Q(x,\xi)] = \sum_{k=1}^K p_k Q(x,\xi_k)$ where

$$Q(x,\xi_k) = \inf\{q_k^T y_k\colon T_k x + W_k y_k = h_k, y_k \geq 0\}. \tag{2.6}$$

Consequently, we can write (2.2)–(2.3) in the form

$$\begin{aligned} \operatorname*{Min}_{x,y_1,\ldots,y_K}\ & c^T x + \sum_{k=1}^K p_k q_k^T y_k \\ \text{s.t.}\ & Ax = b, \\ & T_k x + W_k y_k = h_k, \quad k = 1,\ldots,K, \\ & x \geq 0, y_k \geq 0, \quad k = 1,\ldots,K. \end{aligned} \tag{2.7}$$

That is, the two-stage problem can be formulated as one large linear programming problem.

Example 4 (Capacity Expansion). Consider a directed graph with node set $\mathcal{N}$ and arc set $\mathcal{A}$. With each arc $a \in \mathcal{A}$ we associate a decision variable x_a and call it the *capacity* of a. There is a cost c_a for each unit of capacity of arc a.

For each pair of nodes $(m, n) \in \mathcal{N} \times \mathcal{N}$ we have a random demand D^{mn} for shipments from m to n. These shipments have to be sent through the network and they can be arbitrarily split into pieces taking different paths. We denote by y_a^{mn} the amount of the shipment from m to n sent through arc a. There is a unit cost q_a for shipments on each arc a.

Our objective is to assign arc capacities and to organize shipments in such a way that the expected total cost, comprising the capacity cost and the shipping cost, is minimized. The condition is that the capacities have to be assigned *before* the actual demands D_{mn} become known, while the shipments can be arranged *after* that.

We recognize in this model a linear two-stage stochastic programming model with first stage variables x_a, $a \in \mathcal{A}$, and second stage variables y_a^{mn}, $a \in \mathcal{A}, (m, n) \in \mathcal{N} \times \mathcal{N}$.

Let us define the second stage problem. For each node i denote by $\mathcal{A}_+(i)$ and $\mathcal{A}_-(i)$ the sets of arcs entering and leaving node i. The second stage problem is the multicommodity network flow problem

$$
\begin{aligned}
&\text{Min} \sum_{m,n \in \mathcal{N}} \sum_{a \in \mathcal{A}} q_a y_a^{mn} \\
&\text{s.t.} \quad \sum_{a \in \mathcal{A}_+(i)} y_a^{mn} - \sum_{a \in \mathcal{A}_-(i)} y_a^{mn} = \begin{cases} -D^{mn}, & \text{if } i = m, \\ D^{mn}, & \text{if } i = n, \\ 0, & \text{otherwise,} \end{cases} \\
&\qquad \sum_{m,n \in \mathcal{N}} y_a^{mn} \le x_a, \quad a \in \mathcal{A}, \\
&\qquad y_a^{mn} \ge 0, \quad a \in \mathcal{A},\ i, m, n \in \mathcal{N}.
\end{aligned} \tag{2.8}
$$

This problem depends on the random demand vector D and on the arc capacities, x. Its optimal value will be denoted $Q(x, D)$.The first stage problem has the form

$$
\operatorname*{Min}_{x \ge 0} \sum_{a \in \mathcal{A}} c_a x_a + \mathbb{E}[Q(x, D)].
$$

In this example only some right hand side entries in the second stage constraints are random. All the matrices and cost vectors are deterministic. Nevertheless, the size of this problem, even for discete distributions of the demands, may be enormous. If the number of nodes is ν, the demand vector has $\nu(\nu - 1)$ components. If they are independent, and each of them has r possible realizations, we have to deal with $K = r^{\nu(\nu-1)}$ scenarios. For each of

them the second stage vector has $|\nu(\nu-1)|\mathcal{A}||$ components and there are $\nu^2(\nu-1)+|\mathcal{A}|$ constraints (excluding nonnegativity constraints). As a result, the large scale linear programming formulation has $|\mathcal{A}|+\nu(\nu-1)|\mathcal{A}|r^{\nu(\nu-1)}$ variables and $(\nu^2(\nu-1)+|\mathcal{A}|)r^{\nu(\nu-1)}$ constraints. These are large numbers, even for moderately sized networks and distributions with only few possibilities.

A more complex situation occurs when the arcs are subject to failures and they may lose random fractions θ_a of their capacities. Then the capacity constraint in the second stage problem has a slightly different form:

$$\sum_{m,n\in\mathcal{N}} y_a^{mn} \le (1-\theta_a)x_a, \quad a \in \mathcal{A},$$

and we have a two-stage problem with a random 'technology' matrix. Its complexity, of course, is even higher than before.

2.3 Scenario formulation and nonanticipativity

Let us relax problem (2.7) by replacing the first stage decision vector x by K possibly different vectors x_k. We obtain the problem

$$\begin{aligned} \operatorname*{Min}_{\substack{x_1,\dots,x_K\\ y_1,\dots,y_K}} & \sum_{k=1}^{K} p_k(c^T x_k + q_k^T y_k) \\ \text{s.t.} \quad & Ax_k = b, \\ & T_k x_k + W_k y_k = h_k, \\ & x_k \ge 0, y_k \ge 0, \quad k = 1,\dots,K. \end{aligned} \tag{2.9}$$

Problem (2.9) is separable in the sense that it can be split into K smaller problems, one for each scenario, and therefore it is much easier for a numerical solution. However, (2.9) is not suitable for modeling a two stage process. This is because the first stage decision variables x_k in (2.9) are now allowed to depend on a realization of the random data at the second stage. This can be fixed by introducing the additional constraints

$$x_k = x_j, \quad \text{for all } 1 \le k < j \le K. \tag{2.10}$$

Together with the additional constraints (2.10), problem (2.9) becomes equivalent to (2.7).

Constraints (2.10) are called *nonanticipativity* constraints. They ensure that the first stage decision variables do not depend on the second stage realization of the random data. Such nonanticipativity constraints will be especially important in multistage modeling which we will discuss later.

In fact, some of the constraints in (2.10) are redundant; for example, it is sufficient to require that $x_k = x_{k+1}$ for $k = 1, \ldots, K-1$. There are many other ways to express these conditions, but they all define the same linear subspace of the space of decision variables of (2.9). A way to express the non-anticipativity condition is to require that

$$x_k = \sum_{i=1}^{K} p_i x_i, \quad k = 1, \ldots, K, \tag{2.11}$$

which is convenient for extensions to the general case.

2.4 General formulation

As it was discussed above the essence of two stage modeling is that there are two distinct parts of the decision vector. The value of the first vector $x \in \mathcal{X}$, with $\mathcal{X} = \mathbb{R}^n$, has to be chosen *before* any realization of the unknown quantities, summarized in the data vector $\xi = \xi(\omega)$, are observed. The value of the second part, y, can be chosen *after* the realization of ξ becomes known and generally depends on the realization of ξ and on the choice of x. Consequently, at the first stage one has to solve the expectation optimization problem

$$\operatorname*{Min}_{x \in \mathcal{X}} \mathbb{E}[F(x, \omega)]. \tag{2.12}$$

In the case of two-stage linear problem (2.2),

$$F(x, \omega) := c^T x \; + \; Q(x, \xi(\omega))$$

with $Q(x, \xi)$ being the optimal value of the second stage optimization problem (2.3) (viewed as an *extended* real valued function). In such formulation an explicit dependence on the second stage decision variables y is suppressed. It will be convenient to discuss that formulation first.

As in the example of problem (2.9), we may relax the expectation problem (2.12) by allowing the first stage decision variables to depend on the random data and then to correct that by enforcing nonanticipativity constraints. Denote by $\mathcal{M} = \mathcal{M}(\Omega, \mathcal{F}, \mathcal{X})$ the space of measurable mappings[4] $x(\cdot): \Omega \to \mathcal{X}$ such that the expectation $\mathbb{E}[F(x(\omega), \omega)]$ is well defined. Then the relaxed problem can be formulated in the form

$$\operatorname*{Min}_{x(\cdot) \in \mathcal{M}} \mathbb{E}[F(x(\omega), \omega)]. \tag{2.13}$$

[4] We write here $x(\cdot)$, instead of x, in order to emphasize that $x(\cdot)$ is not a vector, but rather a vector valued function of ω.

Denote

$$\vartheta(\omega) := \inf_{x \in \mathcal{X}} F(x, \omega)$$

the optimal value function of problem (2.12).

Note that optimization in (2.13) is performed over all mappings $x(\omega)$ in the functional space $\mathcal{M}$. In particular, if $\Omega := \{\omega_1, \ldots, \omega_K\}$ is finite, with respective probabilities $p_1, \ldots, p_k$, then $x(\omega)$ can be identified with $(x_1, \ldots, x_K)$, where $x_k := x(\omega_k)$. In that case problem (2.13) can be written in the form

$$\operatorname*{Min}_{x1, \ldots, x_K} \sum_{k=1}^{K} p_k F(x_k, \omega_k). \tag{2.14}$$

Proposition 5. *Suppose that*: (i) *the function* $F(x, \omega)$ *is random lower semicontinuous,*[5] (ii) *either* $\mathbb{E}[\vartheta(\omega)_+] < +\infty$ *or* $\mathbb{E}[(-\vartheta(\omega))_+] < +\infty$. *Then*

$$\inf_{x(\cdot) \in \mathcal{M}} \mathbb{E}[F(x(\omega), \omega)] = \mathbb{E}\left[\inf_{x \in \mathcal{X}} F(x, \omega)\right]. \tag{2.15}$$

Proof. Since $F(x, \omega)$ is random lsc we have by Theorem 19 that $\vartheta(\omega)$ is measurable. Together with the assumption (ii) this implies that the expectation in the right hand side of (2.15) is well defined. For any $x(\cdot) \in \mathcal{M}(\Omega, \mathcal{F}, \mathcal{X})$ we have that $F(x(\omega), \omega) \geq \vartheta(\omega)$ for all $\omega \in \Omega$, and hence the left hand side of (2.15) is always greater than or equal to the right hand side of (2.15). Conversely, if $\vartheta(\omega) > -\infty$ for a.e. $\omega \in \Omega$, then for any given $\varepsilon > 0$ and a.e. $\omega \in \Omega$ there exists an ε-optimal solution $\tilde{x}(\omega)$. Moreover, since $F(x, \omega)$ is random lsc, $\tilde{x}(\omega)$ can be chosen to be measurable, i.e., $\tilde{x} \in \mathcal{M}(\Omega, \mathcal{F}, \mathcal{X})$. It follows that

$$\mathbb{E}[F(\tilde{x}(\omega), \omega)] \leq \mathbb{E}[\vartheta(\omega)] + \varepsilon.$$

Since ε is an arbitrary positive number, this implies that the left hand side of (2.15) is less than or equal to the right hand side of (2.15). Finally, if the event "$\vartheta(\omega) = -\infty$" happens with positive probability, then both sides of (2.15) are equal to $-\infty$. □

[5] See Section 5.3 of the Appendix for the definition and discussion of random lower semicontinuous functions.

We also have that problem (2.12) is equivalent to

$$\operatorname*{Min}_{x(\cdot)\in\mathcal{M}} \mathbb{E}[F(x(\omega),\omega)] \tag{2.16}$$

$$\text{s.t.} \quad x(\omega) = \mathbb{E}[x(\omega)], \quad \forall \omega \in \Omega. \tag{2.17}$$

Constraints (2.17) give an extension of constraints (2.11), and represent the *nonanticipativity* condition.[6] Since problem (2.13) is a relaxation of (2.16)–(2.17), and because of (2.15), we obtain that

$$\inf_{x\in X} \mathbb{E}[F(x,\omega)] \geq \mathbb{E}\left[\inf_{x\in\mathcal{X}} F(x,\omega)\right]. \tag{2.18}$$

The above inequality also follows directly from the obvious inequality $F(x,\omega) \geq \vartheta(\omega)$ for all $x \in \mathcal{X}$ and $\omega \in \Omega$.

Let us give now a formulation where the second stage decision variables appear explicitly:

$$\operatorname*{Min}_{x\in X} \mathbb{E}[V(x,\xi(\omega))], \tag{2.19}$$

where $V(x,\xi)$ is the optimal value of the second stage problem

$$\begin{aligned} &\operatorname*{Min}_{y\in Y} \ F(x,y,\xi) \\ &\text{s.t.} \quad G_i(x,y,\xi) \leq 0, \quad i = 1,\ldots,m. \end{aligned} \tag{2.20}$$

Here X is a subset of $\mathbb{R}^{n_1}$, Y is a subset of $\mathbb{R}^{n_2}$, and

$$\begin{aligned} &F\colon \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \mathbb{R}^d \longrightarrow \mathbb{R}, \\ &G_i\colon \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \mathbb{R}^d \longrightarrow \mathbb{R}, \quad i = 1,\ldots,m, \end{aligned}$$

are the objective and the constraint functionals, respectively.

Alternatively, in an abstract form the above two stage stochastic programming problem can be formulated as follows

$$\operatorname*{Min}_{x\in\mathcal{X},\ y(\cdot)\in\mathcal{Y}} \mathbb{E}[F(x,y(\omega),\xi(\omega))] \tag{2.21}$$

$$\text{s.t.} \quad G_i(x,y(\omega),\xi(\omega)) \leq 0, \quad i = 1,\ldots,m, \tag{2.22}$$

$$x \in X, \tag{2.23}$$

[6] Since the expected value of two random variables which may differ on a set of measure zero is the same, it actually suffices to verify the constraints (2.17) for P-almost every $\omega \in \Omega$.

$$y(\omega) \in Y, \tag{2.24}$$

where $\mathcal{X} := \mathbb{R}^{n_1}$ and $\mathcal{Y}$ is a space of measurable functions from Ω to $\mathbb{R}^{n_2}$. In that formulation $y(\omega)$ is viewed as a *random* vector in $\mathbb{R}^{n_2}$. Note, however, an important difference between random vectors $\xi(\omega)$ and $y(\omega)$. Vector $\xi(\omega)$ represents the random data of the problem with a given (known) distribution, while $y(\omega)$ denotes the second stage decision variables. We have explicitly marked the dependence of y on the elementary event ω to stress the recourse nature of these variables. The inequalities (2.22) and the inclusion (2.24) are understood in the *almost sure* sense, i.e., they have to hold for P-almost every[7] $\omega \in \Omega$. Recall that the probability measure P on $(\Omega, \mathcal{F})$ generates the corresponding probability distribution of $(\xi(\omega), y(\omega))$ viewed as a random vector. Therefore, "for P-almost every $\omega \in \Omega$" means that the event happens for almost every realization of the random vector (ξ, y).

The difficulty in the formulation (2.21)–(2.24) is the fact that the second stage decisions y are allowed to be *functions* of the elementary event ω. We need to specify from which classes of functions these decisions have to be chosen, i.e., to define the functional space $\mathcal{Y}$. The mappings $y\colon \Omega \to \mathbb{R}^{n_2}$, have to be measurable with respect to the sigma algebra $\mathcal{F}$ and such that the expectation in (2.21) makes sense. Otherwise we shall not be able to talk in a meaningful way about the expectation of the objective functional and the 'almost sure' satisfaction of the constraints. Moreover, in fact y is a function of ξ. Therefore, we can identify the probability space $(\Omega, \mathcal{F}, P)$ with the probability space $(\mathbb{R}^d, \mathcal{B}, \mathbb{P})$ of the random vector ξ, and view $y(\cdot)$ as an element of a space of measurable mappings from $\mathbb{R}^d$ into $\mathbb{R}^{n_2}$. In particular, in the case of finitely many realizations $\xi_1, \ldots, \xi_K$, we can identify the sample space with the set $\Omega := \{1, \ldots, K\}$ equipped with the sigma algebra of all its subsets. In that case it suffices to consider mappings $y\colon \{1, \ldots, K\} \to \mathbb{R}^{n_2}$, which could be identified with vectors $y_1, \ldots, y_K \in \mathbb{R}^{n_2}$. As a result, the decision space in the case of finitely many realizations is just $\mathbb{R}^{n_1} \times \underbrace{\mathbb{R}^{n_2} \times \cdots \times \mathbb{R}^{n_2}}_{K \text{ times}}$.

The constraints (2.22)–(2.24) can be absorbed into the objective function by defining

$$\overline{F}(x, y, \xi) := \begin{cases} F(x, y, \xi), & \text{if } x \in X, y \in Y, G_i(x, y, \xi) \leq 0, i = 1, \ldots, m, \\ +\infty, & \text{otherwise.} \end{cases}$$

[7] Written: "a.e. $\omega \in \Omega$".

Then problem (2.21)–(2.24) can be written in the form

$$\underset{x\in\mathcal{X},\, y(\cdot)\in\mathcal{Y}}{\text{Min}} \ \mathbb{E}\big[\, \overline{F}(x,\ y(\omega),\ \xi(\omega))\big]. \tag{2.25}$$

In a way similar to the proof of Proposition 5 it is possible to show that the two formulations (2.19)–(2.20) and (2.21)–(2.24) are equivalent if for every $x \in \mathcal{X}$, the function $\overline{F}(x, \cdot, \cdot)$ is random lsc and the expectation of the optimal value function $\inf_{y\in\mathbb{R}^{n_2}} \overline{F}(x, y, \xi(\omega))$ is well defined.

Let us now consider both parts of the decision vector, x and y as random elements. We obtain the problem

$$\begin{aligned} &\underset{x(\cdot),\, y(\cdot)}{\text{Min}} \ \mathbb{E}[F(x(\omega),\ y(\omega),\ \xi(\omega))] \\ &\text{s.t.} \quad G_i(x(\omega),\ y(\omega),\ \xi(\omega)) \le 0, \quad i = 1,\ \ldots,\ m, \\ &\qquad x(\omega) \in X,\ y(\omega) \in Y. \end{aligned}$$

All constraints here are assumed to hold P-almost surely, i.e., for a.e. $\omega \in \Omega$. The above problem is an analogue of (2.13) with optimization performed over mappings $(x(\cdot), y(\cdot))$ in an appropriate functional space, and as in the finite scenario case, is a relaxation of the problem (2.21)–(2.24). To make it equivalent to the original formulation we must add the nonanticipativity constraint which can be written, for example, in the form (2.17).

For example, consider the two-stage linear program (2.2)–(2.3). We can write it in the form

$$\begin{aligned} &\underset{x,\, y(\cdot)}{\text{Min}} \ \mathbb{E}\big[c^T x \ +\ q(\omega)^T y(\omega)\big] \\ &\text{s.t.} \quad T(\omega)x \ +\ W(\omega)y(\omega) = h(\omega),\ \text{a.e. } \omega \in \Omega, \\ &\qquad Ax = b,\ x \ge 0, \\ &\qquad y(\omega) \ge 0,\ \text{a.e. } \omega \in \Omega, \end{aligned}$$

with $y(\cdot)$ being a mapping from Ω into $\mathbb{R}^{n_2}$. In order for the above problem to make sense the mapping $y(\omega)$ should be measurable and the corresponding expected value should be well defined. Suppose for a moment that vector q is not random, i.e., it does not depend on ω. Then we can assume that $y(\omega)$ is an element of the space $\mathcal{L}_1^{n_2}(\Omega, \mathcal{F}, P)$ of $\mathcal{F}$-measurable mappings[8] $y\colon \Omega \to \mathbb{R}^{n_2}$

[8] In fact an element of $\mathcal{L}_1^{n_2}(\Omega, \mathcal{F}, P)$ is a class of mappings which may differ from each other on sets of P-measure zero.

such that $\int_\Omega ||y(\omega)||\, dP(\omega) < +\infty$. If $q(\omega)$ is random we can consider a space of measurable mappings $y(\cdot)$ such that $\int_\Omega |q(\omega)^T y(\omega)|\, dP(\omega) < +\infty$.

2.5 Value of perfect information

Consider a two stage stochastic programming problem in the form (2.19) with $V(x, \xi)$ being the optimal value of the second stage problem (2.20). If we have a perfect information about the data ξ, i.e., the value of ξ is known at the time when the first stage decision should be made, then the optimization problem becomes the deterministic problem

$$\operatorname*{Min}_{x \in X} V(x, \xi), \tag{2.26}$$

and can be written in the following equivalent form

$$\begin{aligned} &\operatorname{Min}_{x \in X,\, y \in Y} F(x, y, \xi) \\ &\text{s.t.} \quad G_i(x, y, \xi) \leq 0, \quad 1, \ldots, m. \end{aligned} \tag{2.27}$$

Of course, the optimal solution $\overline{x}(\xi)$ (if it exists) and the optimal value $\upsilon(\xi)$ of problem (2.26) depend on the realization ξ of the data. The average of $\upsilon(\xi)$ over all possible realizations of the random data $\xi = \xi(\omega)$, i.e., the expected value

$$\mathbb{E}[\upsilon(\xi)] = \mathbb{E}\left[\inf_{x \in X} V(x, \xi(\omega))\right], \tag{2.28}$$

is called the *wait-and-see* solution.

We have that for any $x \in X$ and any ξ the inequality $V(x, \xi) \geq \upsilon(\xi)$ holds, and hence

$$\mathbb{E}[V(x, \xi(\omega)] \geq \mathbb{E}\left[\inf_{x \in X} V(x, \xi(\omega))\right]. \tag{2.29}$$

Therefore, as it was mentioned earlier (see (2.18)), it follows that

$$\inf_{x \in X} \mathbb{E}[V(x, \xi(\omega)] \geq \mathbb{E}\left[\inf_{x \in X} V(x, \xi(\omega))\right]. \tag{2.30}$$

That is, the optimal value of the stochastic programming problem (2.19) is always greater than or equal to $\mathbb{E}[\upsilon(\xi)]$. Suppose further that problem (2.19) has an optimal solution $\hat{x}$. We have that $V(\hat{x}, \xi) - \upsilon(\xi)$ is nonnegative for all ξ,

and hence its expected value is zero iff $V(\hat{x},\xi)-\upsilon(\xi)=0$ w.p.1. That is, the equality in (2.30) holds, iff

$$V(\hat{x},\xi(\omega))=\inf_{x\in X}V(x,\xi(\omega)) \quad \text{for a.e. } \omega\in\Omega. \tag{2.31}$$

In particular, the equality in (2.30) holds, if there exists an optimal solution of (2.26) which does not depend on ξ w.p.1.

The difference $V(\hat{x},\xi)-\upsilon(\xi)$ is the value of perfect information of knowing the realization ξ. Consequently,

$$\text{EVPI}:=\inf_{x\in X}\mathbb{E}[V(x,\xi(\omega))]-\mathbb{E}\left[\inf_{x\in X}V(x,\xi(\omega))\right] \tag{2.32}$$

represents the *expected value of perfect information*. It follows from (2.30) that EVPI is always nonnegative, and EVPI $=0$ iff condition (2.31) holds.

3 Multistage models

3.1 The linear case

The two-stage model is a special case of a more general structure, called the multi-stage stochastic programming model, in which the decision variables and constraints are divided into groups corresponding to stages $t=1,\ldots,T$. The fundamental issue in such a model is the *information structure*: what is known at stage t when decisions associated with this period are made? We first give a general description of such multistage models and then discuss examples in Section 3.4.

Let $x_1,\ldots,x_T$ be decision vectors corresponding to time periods (stages) $1,\ldots,T$. Consider the following linear programming problem

$$\begin{array}{lccccccccc}
\text{Min} & c_1^Tx_1 & + & c_2^Tx_2 & + & c_3^Tx_3 & + & \ldots & + & c_T^Tx_T & \\
\text{s.t.} & A_{11}x_1 & & & & & & & & & = b_1,\\
& A_{21}x_1 & + & A_{22}x_2 & & & & & & & = b_2,\\
& & & A_{32}x_2 & + & A_{33}x_3 & & & & & = b_3,\\
& \ldots\ldots & & & & & & & & & \\
& & & & & & & A_{T,T-1}x_{T-1} & + & A_{TT}x_T & = b_T,\\
& x_1\geq 0, & & x_2\geq 0, & & x_3\geq 0, & & \ldots & & x_T\geq 0. &
\end{array} \tag{3.1}$$

We view it as a multiperiod stochastic program where c_1, A_{11} and b_1 are known, but some (all) of the entries of the cost vectors $c_2,\ldots,c_T$, matrices

$A_{t,t-1}$ and A_{tt}, $t = 2, \ldots, T$, and right hand side vectors $b_2, \ldots, b_T$ are random. At each stage some of these quantities become known, and we have the following sequence of actions:

$$\begin{aligned}
&\text{decision } (x_1)\\
&\text{observation } \xi_2 := (c_2, A_{21}, A_{22}, b_2)\\
&\text{decision } (x_2)\\
&\vdots\\
&\text{observation } \xi_T := (c_T, A_{T,T-1}, A_{TT}, b_T)\\
&\text{decision } (x_T).
\end{aligned}$$

Our objective is to design the decision process in such a way that the expected value of the total cost is minimized while optimal decisions are allowed to be made at *every* time period $t = 1, \ldots, T$.

Let us denote by ξ_t the data which become known at time period t. In the setting of the multiperiod problem (3.1), ξ_t is assembled from the components of c_t, $A_{t,t-1}$, A_{tt}, b_t, some (all) of which can be random, and the data $\xi_1 = (c_1, A_{11}, b_1)$ at the first stage of problem (3.1) which is assumed to be known. For $1 \leq t_1 \leq t_2 \leq T$, denote by

$$\xi_{[t_1,t_2]} := (\xi_{t_1}, \ldots, \xi_{t_2})$$

the history of the process from time t_1 to time t_2. In particular, $\xi_{[1,t]}$ represents the information available up to time t. The important condition in the above multistage process is that every decision vector x_t may depend on the information available at time t (that is, $\xi_{[1,t]}$), but *not* on the results of observations to be made at later stages. This differs multistage stochastic programs from deterministic multiperiod problems, in which all the information is assumed to be available at the beginning.

There are several possible ways how multistage stochastic programs can be formulated in a precise mathematical form. In one such formulation $x_t = x_t(\xi_{[1,t]})$, $t = 2, \ldots, T$, is viewed as a function of $\xi_{[1,t]} = (\xi_1, \ldots, \xi_t)$, and the minimization in (3.1) is performed over appropriate functional spaces (as it was discussed in Section 2.4 in the case of two-stage programming). If the number of scenarios is finite, this leads to a formulation of the linear multistage stochastic program as one large (deterministic) linear programming problem. We discuss that further in the following Section 3.2. It is also useful to connect dynamics of the multistage process starting from the end as follows.

Let us look at our problem from the perspective of the last stage T. At that time the values of all problem data, $\xi_{[1,T]}$, are already known, and the values

of the earlier decision vectors, $x_1, \ldots, x_{T-1}$, have been chosen. Our problem is, therefore, a simple linear programming problem

$$\begin{aligned} \underset{x_T}{\text{Min}} \ & c_T^T x_T \\ \text{s.t.} \ & A_{T,T-1}x_{T-1} + A_{TT}x_T = b_T, \\ & x_T \geq 0. \end{aligned} \tag{3.2}$$

The optimal value of this problem depends on the earlier decision vector x_{T-1} and data $\xi_T = (c_T, A_{T,T-1}, A_{T,T}, b_T)$, and is denoted by $Q_T(x_{T-1}, \xi_T)$.

At stage $T-1$ we know x_{T-2} and $\xi_{[1,T-1]}$. We face, therefore, the following two-stage stochastic programming problem

$$\begin{aligned} \underset{x_{T-1}}{\text{Min}} \ & c_{T-1}^T x_{T-1} + \mathbb{E}\big[Q_T(x_{T-1}, \xi_T) \mid \xi_{[1,T-1]}\big] \\ \text{s.t.} \ & A_{T-1,T-2}x_{T-2} + A_{T-1,T-1}x_{T-1} = b_{T-1}, \\ & x_{T-1} \geq 0. \end{aligned} \tag{3.3}$$

The optimal value of the above problem depends on x_{T-2} and data $\xi_{[1,T-1]}$, and is denoted $Q_{T-1}(x_{T-2}, \xi_{[1,T-1]})$.

Generally, at stage $t = 2, \ldots, T-1$, we have the problem

$$\begin{aligned} \underset{x_t}{\text{Min}} \ & c_t^T x_t + \mathbb{E}\big[Q_{t+1}(x_t, \xi_{[1,t+1]}) \mid \xi_{[1,t]}\big] \\ \text{s.t.} \ & A_{t,t-1}x_{t-1} + A_{t,t}x_t = b_t, \\ & x_t \geq 0. \end{aligned} \tag{3.4}$$

Its optimal value is denoted $Q_t(x_{t-1}, \xi_{[1,t]})$ and is called the *cost-to-go* function.

Note that, since ξ_1 is not random, the conditional distribution of ξ_{t+1} given $\xi_{[1,t]}$ is the same as the conditional distribution of ξ_{t+1} given $\xi_{[2,t]}$, $t = 2, \ldots, T-1$. Therefore, it suffices to take the conditional expectation in (3.4) (in (3.3)) with respect to $\xi_{[2,t]}$ (with respect to $\xi_{[2,T-1]}$), only.

On top of all these problems is the problem to find the first decisions, x_1,

$$\begin{aligned} \underset{x_1}{\text{Min}} \ & c_1^T x_1 + \mathbb{E}[Q_2(x_1, \xi_2)] \\ \text{s.t.} \ & A_{11}x_1 = b_1, \\ & x_1 \geq 0. \end{aligned} \tag{3.5}$$

Note that all subsequent stages $t = 2, \ldots, T$ are absorbed in the above problem (3.5) into the function $Q_2(x_1, \xi_2)$ through the corresponding

expected values. Note also that since ξ_1 is not random, the optimal value $Q_2(x_1, \xi_2)$ does not depend on ξ_1. In particular, if $T = 2$, then (3.5) coincides with the formulation (2.2) of a two-stage linear problem.

We arrived in this way at the following *nested formulation*:

$$\operatorname*{Min}_{\substack{A_{11}x_1 = b_1 \\ x_1 \geq 0}} c_1^T x_1 + \mathbb{E}\left[\min_{\substack{A_{21}x_1 + A_{22}x_2 = b_2 \\ x_2 \geq 0}} c_2^T x_2 + \mathbb{E}\left[\cdots + \mathbb{E}\left[\min_{\substack{A_{T,T-1}x_{T-1} + A_{TT}x_T = b_T \\ x_T \geq 0}} c_T^T x_T\right]\right]\right].$$

Recall that the random process $\xi_1, \ldots, \xi_T$ is said to be *Markovian*, if for each $t = 2, \ldots, T-1$ the conditional distribution of ξ_{t+1} given $\xi_{[1,t]} = (\xi_1, \ldots, \xi_t)$ is the same as the conditional distribution of ξ_{t+1} given ξ_t. If the process $\xi_1, \ldots, \xi_T$ is Markovian, the model is simplified considerably. In the Markovian case, for given ξ_{T-1}, the conditional expectation in problem (3.3) does not depend on $\xi_1, \ldots, \xi_{T-2}$, and hence the optimal value of (3.3) depends only on x_{T-2} and ξ_{T-1}. Similarly, at stage $t = 2, \ldots, T-1$, the optimal value of problem (3.4) is then a function of x_{t-1} and ξ_t, and can be denoted by $Q_t(x_{t-1}, \xi_t)$. We shall call then ξ_t the *information state* of the model. In particular, the process $\xi_1, \ldots, \xi_T$ is Markovian if the random vectors ξ_t, $t = 2, \ldots, T$, are mutually independent. In that case the conditional expectation in problem (3.3) does not depend on $\xi_{[1,T-1]}$, and hence the optimal value $Q_{T-1}(x_{T-2}, \xi_{T-1})$ of (3.3) depends on ξ_{T-1} only through the linear constraint of that problem, and similarly, at stages $t = T-2, \ldots$, the optimal value $Q_t(x_{t-1}, \xi_t)$ depends on ξ_t only through the linear constraint of (3.4).

The assumption that the blocks $A_{t1}, \ldots, A_{t,t-2}$ in the constraint matrix are zeros, allowed us to express the optimal value Q_t of (3.4) as the function of the immediately preceding decision, x_{t-1}, rather than all earlier decisions $x_1, \ldots, x_{t-1}$. Suppose now that we deal with an underlying model with a full lower block triangular constraint matrix:

$$\begin{array}{lllllllllll}
\text{Min} & c_1^T x_1 & + & c_2^T x_2 & + & c_3^T x_3 & + & \ldots & + & c_T^T x_T & \\
\text{s.t.} & A_{11}x_1 & & & & & & & & & = b_1, \\
& A_{21}x_1 & + & A_{22}x_2 & & & & & & & = b_2, \\
& A_{31}x_1 & + & A_{32}x_2 & + & A_{33}x_3 & & & & & = b_3, \\
& \ldots\ldots & & & & & & & & & \\
& A_{T1}x_1 & + & A_{T2}x_2 & + & \ldots & + & A_{T,T-1}x_{T-1} & + & A_{TT}x_T & = b_T, \\
& x_1 \geq 0, & & x_2 \geq 0, & & x_3 \geq 0, & & \ldots & & x_T \geq 0. &
\end{array} \tag{3.6}$$

Then, of course, each subproblem (3.4) depends on the entire history of our decisions, $x_{[1,t-1]} := (x_1, \dots, x_{t-1})$. It takes on the form

$$\begin{aligned} \min_{x_t}\ & c_t^T x_t + \mathbb{E}\big[Q_{t+1}(x_{[1,t]}, \xi_{[1,t+1]}) \mid \xi_{[1,t]}\big] \\ \text{s.t. } & A_{t1}x_1 + \cdots + A_{t,t-1}x_{t-1} + A_{t,t}x_t = b_t, \\ & x_t \ge 0. \end{aligned} \tag{3.7}$$

Its optimal value is denoted $Q_t(x_{[1,t-1]}, \xi_{[1,t]})$.

Sometimes it is convenient to convert such a lower triangular formulation into the staircase formulation from which we started our presentation. This can be accomplished by introducing additional variables r_t which summarize the relevant history of our decisions. We shall call these variables the *model state* variables (to distinguish from information states discussed before). The relations that describe the next values of the state variables as a function of the current values of these variables, current decisions and current random parameters are called *model state equations*.

For the general problem (3.6) the vectors $x_{[1,t]} = (x_1, \dots, x_t)$ are sufficient model state variables. They are updated at each stage according to the state equation $x_{[1,t]} = (x_{[1,t-1]}, x_t)$ (which is linear), and the constraint in (3.7) can be formally written as

$$[A_{t1} A_{t2} \dots A_{t,t-1}] x_{[1,t-1]} + A_{t,t}x_t = b_t.$$

Although, it looks a little awkward in this general case, in many problems it is possible to define model state variables of reasonable size. As an example let us consider the structure

$$\begin{array}{lllllllll} \text{Min} & c_1^T x_1 & + & c_2^T x_2 & + & c_3^T x_3 & + \ \cdots & + & c_T^T x_T \\ \text{s.t.} & A_{11}x_1 & & & & & & & = b_1, \\ & B_1 x_1 & + & A_{22}x_2 & & & & & = b_2, \\ & B_1 x_1 & + & B_2 x_2 & + & A_{33}x_3 & & & = b_3, \\ & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ & B_1 x_1 & + & B_2 x_2 & + & \dots & + \ B_{T-1}x_{T-1} & + & A_{TT}x_T = b_T, \\ & x_1 \ge 0, & & x_2 \ge 0, & & x_3 \ge 0, & \dots & & x_T \ge 0, \end{array}$$

in which all blocks A_{it}, $i = 2, \dots, T$ are identical and observed at time t. Then we can define the state variables r_t, $t = 1, \dots, T$ recursively by the state

equation $r_t = r_{t-1} + B_t x_t$, $t = 1, \ldots, T-1$, where $r_0 = 0$. Subproblem (3.7) simplifies substantially:

$$\begin{aligned} \underset{x_t, r_t}{\text{Min}} \; & c_t^T x_t + \mathbb{E}\big[Q_{t+1}(r_t, \xi_{[1,t+1]}) \mid \xi_{[1,t]}\big] \\ \text{s.t.} \; & r_{t-1} + A_{tt} x_t = b_t, \\ & r_t = r_{t-1} + B_t x_t, \\ & x_t \geq 0. \end{aligned}$$

Its optimal value depends on r_{t-1} and is denoted $Q_t(r_{t-1}, \xi_{[1,t]})$.

Trucking Example 9 (discussed in Section 3.4) uses such model state variables: the capacities r_t available at all locations at the end of day t. We do not need to remember all decisions made in the past, we only need to know the numbers of trucks at each location today.

It should be clear, too, that the simple sign constraints $x_t \geq 0$ can be replaced in our model by a general constraint $x_t \in X_t$, where X_t is a convex polyhedron defined by some linear equations and inequalities (local for stage t). The set X_t may be random, too, but has to become known at stage t.

3.2 *The case of finitely many scenarios*

Suppose that in our basic problem (3.1) there are only finitely many, say K, different values the problem data can take. We shall call them *scenarios*.

With each scenario k is associated probability p_k and the corresponding sequence of decisions[9] $x^k = (x_1^k, x_2^k, \ldots, x_T^k)$. Of course, it would not be appropriate to try to find the optimal values of these decisions by solving the relaxed version of (3.1):

$$\begin{aligned} \text{Min} \; & \textstyle\sum_{k=1}^K p_k [(c_1)^T x_1^k + (c_2^k)^T x_2^k + (c_3^k)^T x_3^k + \quad \cdots \quad + (c_T^k)^T x_T^k] \\ \text{s.t.} \quad & A_{11} x_1^k = b_1, \\ & A_{21}^k x_1^k + A_{22}^k x_2^k = b_2^k, \\ & A_{32}^k x_2^k + A_{33}^k x_3^k = b_3^k, \\ & \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ & A_{T,T-1}^k x_{T-1}^k + A_{TT}^k x_T^k = b_T^k, \\ & x_1^k \geq 0, \quad x_2^k \geq 0, \quad x_3^k \geq 0, \quad \ldots \quad x_T^k \geq 0, \\ & k = 1, \ldots, K. \end{aligned} \tag{3.8}$$

[9] To avoid ugly collisions of subscripts we change our notation a little and we put the index of the scenario, k, as a superscript.

The reason is the same as in the two-stage case: in the problem above all parts of the decision vector are allowed to depend on *all* parts of the random data, while in reality each part x_t is allowed to depend only on the data known up to stage t. In particular, problem (3.8) may suggest different values of x_1 for each scenario k, but we need only one value.

It is clear that we need the nonanticipativity constraints

$$x_1^k = x_1^j \quad \text{for all } 1 \le k < j \le K, \tag{3.9}$$

similarly to (2.10). But this is not sufficient, in general. Consider the second part of the decision vector, x_2. It is allowed to depend only on $\xi_{[1,2]} = (\xi_1, \xi_2)$, so it has to have the same value for all scenarios k for which $\xi^k_{[1,2]}$ is identical. We must therefore, satisfy the equations

$$x_2^k = x_2^j \quad \text{for all } k, j \text{ for which } \xi^k_{[1,2]} = \xi^j_{[1,2]}. \tag{3.10}$$

Generally, at stage $t = 1, \ldots, T$, the scenarios that have the same history $\xi_{[1,t]}$ cannot be distinguished, so we need to enforce the nonanticipativity constraints

$$x_t^k = x_t^j \quad \text{for all } k, j \text{ for which } \xi^k_{[1,t]} = \xi^j_{[1,t]}, \quad t = 1, \ldots, T. \tag{3.11}$$

Problem (3.8) together with the nonanticipativity constraints (3.11) becomes equivalent to our original formulation (3.1).

Let us observe that if in the problem (3.8) only the constraints (3.9) are enforced, then from the mathematical point of view the obtained problem becomes a two-stage stochastic linear program with K scenarios. In that two-stage program the first stage decision vector is x_1, the second stage decision vector is $(x_2, \ldots, x_K)$, the technology matrix is A_{21} and the recourse matrix is the block matrix

$$\begin{bmatrix} A_{22} & 0 & \cdots\cdots & 0 & 0 \\ A_{32} & A_{33} & \cdots\cdots & 0 & 0 \\ & & \cdots\cdots\cdots\cdots & & \\ 0 & 0 & \cdots\cdots & A_{T,T-1} & A_{TT} \end{bmatrix}.$$

Since the obtained two-stage problem is a relaxation of the multistage problem (3.1), its optimal value gives a lower bound for the optimal value of

problem (3.1) and in that sense can be useful. Note, however, that this model does not make much sense, since it assumes that at the end of the process when all realizations of the random data become known, one can go back in time and make all decisions $x_2, \ldots, x_K$.

It is useful to depict the possible sequences of data $\xi_{[1,t]}$ in a form of a *scenario tree*. It has nodes organized in levels which correspond to stages $1, 2, \ldots, T$. At level 1 we have only one root node, and we associate with it the value of ξ_1 (which is known at stage 1). At level 2 we have at least as many nodes as many different realizations of ξ_2 may occur. Each of them is connected with the root node by an arc. For each node i of level 2 (which corresponds to a particular realization ξ_2^i of ξ_2) we create at least as many nodes at level 3 as different values of ξ_3 may follow ξ_2^i, and we connect them with the node i, etc. Generally, nodes at level t correspond to possible values of ξ_t that may occur. Each of them is connected to a unique node at level $t-1$, called the *ancestor* node, which corresponds to the identical first $t-1$ parts of the process $\xi_{[1,t]}$, and is also connected to nodes at level $t+1$, which correspond to possible continuations of history $\xi_{[1,t]}$.

Note that, in general, realizations ξ_t^i are vectors and it may happen that some of the values ξ_t^i, associated with nodes at a given level t, are equal to each other. Nevertheless, such equal values may be represented by different nodes since they may correspond to different histories of the process. Note also that if for every $t = 1, \ldots, T$ all realizations ξ_t^i are *different* from each other, then the random process $\xi_1, \ldots, \xi_T$ is Markovian because of the tree structure of the process. Indeed, in that case the conditional probability of ξ_t to be at state ξ_t^i depends on the previous history of the process only through the ancestor node at level $t-1$.

In order to illustrate the above ideas let us discuss the following simple example.

Example 6 (Scenario Tree). An example of the scenario tree is depicted in Fig. 2. Numbers along the arcs represent conditional probabilities of moving from one node to the next. The associated process $\xi_t = (c_t, A_{t,t-1}, A_{tt}, b_t)$, $t = 1, \ldots, T$, with $T = 4$, is defined as follows. All involved variables are assumed to be one dimensional, with $c_t, A_{t,t-1}, A_{tt}$, $t = 2, 3, 4$, being fixed and only right hand side variables b_t being random. The numerical values (realizations) of the random process $b_1, \ldots, b_T$ are indicated by the bold numbers at the nodes of the tree. Numerical values of $c_t, A_{t,t-1}, A_{tt}$ will be specified later. That is, at level $t = 1$, b_1 has unique value 36. At level $t = 2$, b_2 has two values 15 and 50 with respective probabilities 0.4 and 0.6. At level $t = 3$ we have 5 nodes with which are associated the following numerical values (from left to right) $10, 20, 12, 20, 70$. That is, b_3 can take 4 different values with respective probabilities $\mathbb{P}\{b_3 = 10\} = 0.4 \cdot 0.1$, $\mathbb{P}\{b_3 = 20\} = 0.4 \cdot 0.4 + 0.6 \cdot 0.4$, $\mathbb{P}\{b_3 = 12\} = 0.4 \cdot 0.5$ and $\mathbb{P}\{b_3 = 70\} = 0.6 \cdot 0.6$. At

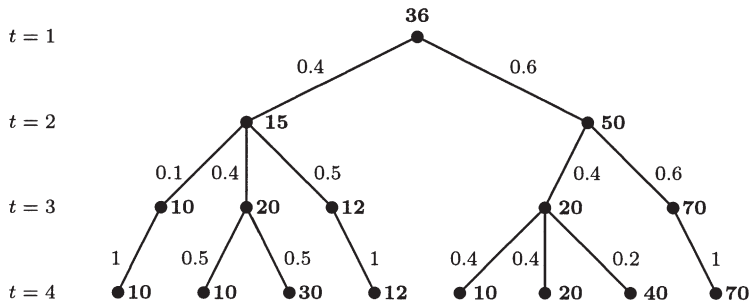

Fig. 2. Scenario tree. Nodes represent information states. Paths from the root to leaves represent scenarios. Numbers along the arcs represent conditional probabilities of moving to the next node. Bold numbers represent numerical values of the process.

level $t = 4$, the numerical values associated with eight nodes are defined, from left to right, as $10, 10, 30, 12, 10, 20, 40, 70$. The respective probabilities can be calculated by using the corresponding conditional probabilities. For example,

$$\mathbb{P}\{b_4 = 10\} = 0.4 \cdot 0.1 \cdot 1.0 \; + \; 0.4 \cdot 0.4 \cdot 0.5 \; + \; 0.6 \cdot 0.4 \cdot 0.4.$$

Note that although some of the realizations of b_3, and hence of ξ_3, are equal to each other, they are represented by different nodes. This is necessary in order to identify different histories of the process corresponding to different scenarios. The same remark applies to b_4 and ξ_4. Altogether, there are eight scenarios in this tree. Figure 3 illustrates the way in which sequences of decisions are associated with scenarios from Fig. 2.

The process b_t (and hence the process ξ_t) in the above example is not Markovian. For instance,

$$\mathbb{P}\{b_4 = 10 \mid b_3 = 20, b_2 = 15, b_1 = 36\} = 0.5,$$

while

$$\mathbb{P}\{b_4 = 10 \mid b_3 = 20\} = \frac{\mathbb{P}\{b_4 = 10, b_3 = 20\}}{\mathbb{P}\{b_3 = 20\}}$$

$$= \frac{0.5 \cdot 0.4 \cdot 0.4 \; + \; 0.4 \cdot 0.4 \cdot 0.6}{0.4 \cdot 0.4 \; + \; 0.4 \cdot 0.6} = 0.44 \neq 0.5.$$

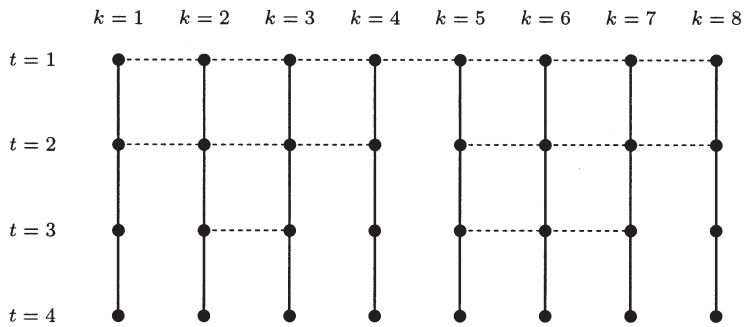

Fig. 3. Sequences of decisions for scenarios from Fig. 2. Horizontal dotted lines represent the equations of nonanticipativity.

On the other hand, the process b_t in this example is a *martingale*.[10] For instance,

$$\mathbb{E}[b_2 \mid b_1 = 36] = \mathbb{E}[b_2] = 15 \cdot 0.4 + 50 \cdot 0.6 = 36,$$
$$\mathbb{E}[b_3 \mid b_2 = 15, b_1 = 36] = 10 \cdot 0.1 + 20 \cdot 0.4 + 12 \cdot 0.5 = 15, \text{ etc.}$$

Suppose now that $c_T = 1$ and $A_{T,T-1} = A_{TT} = 1$. Then the cost-to-go function $Q_T(x_{T-1}, \xi_T)$ is given by the optimal value of the problem

$$\operatorname*{Min}_{x_T} x_T \quad \text{subject to} \quad x_{T-1} + x_T = b_T,\ x_T \geq 0,$$

and hence

$$Q_T(x_{T-1}, \xi_T) = \begin{cases} b_T - x_{T-1}, & \text{if } x_{T-1} \leq b_T, \\ +\infty, & \text{otherwise.} \end{cases}$$

Note again that b_T has six possible realizations.

Suppose further that $c_{T-1} = 1$ and $A_{T-1,T-2} = A_{T-1,T-1} = 1$. Then the cost-to-go function $Q_{T-1}(x_{T-2}, \xi_{[1,T-1]})$ is the optimal value of the problem

$$\operatorname*{Min}_{x_{T-1}} x_{T-1} + \mathbb{E}[Q_T(x_{T-1}, \xi_T) \mid \xi_{[1,T-1]}]$$
$$\text{subject to} \quad x_{T-2} + x_{T-1} = b_{T-1},\ x_{T-1} \geq 0.$$

[10] Recall that a random process Z_t, $t = 1, \ldots,$ is called a martingale if the equalities $\mathbb{E}[Z_{t+1} | Z_{[1,t]}] = Z_t$, $t = 1, \ldots,$ hold with probability one.

The history $b_{[1,3]}$ of the process b_t, and hence the history $\xi_{[1,3]}$ of the process ξ_t, is in one-to-one correspondence with the nodes of the tree at level $t = 3$. It has 5 possible realizations $\xi^i_{[1,3]}$, $i = 1, \ldots, 5$, numbered from left to right, i.e., for $i = 1$ it corresponds to the realization $b_1 = 36$, $b_2 = 15$, $b_3 = 10$ of $b_{[1,3]}$. We have that

$$\mathbb{E}\big[Q_T(x_{T-1}, \xi_4)|\xi_{[1,3]} = \xi^1_{[1,3]}\big] = Q_T(x_{T-1}, \xi^1_4) = \begin{cases} 10 - x_{T-1}, & \text{if } x_{T-1} \le 10, \\ +\infty, & \text{otherwise,} \end{cases}$$

where $\xi^1_4 = \big(1, 1, 1, b^1_4\big)$ and $b^1_4 = 10$. Consequently,

$$Q_{T-1}\big(x_{T-2}, \xi^1_{[1,3]}\big) = \begin{cases} 10, & \text{if } 0 \le x_{T-2} \le 10, \\ +\infty, & \text{otherwise.} \end{cases}$$

Similarly,

$$\mathbb{E}\big[Q_T\big(x_{T-1}, \xi_4\big)|\xi_{[1,3]} = \xi^2_{[1,3]}\big] = \frac{1}{2} Q_T(x_{T-1}, \xi^2_4) + \frac{1}{2} Q_T(x_{T-1}, \xi^3_4),$$

and hence

$$Q_{T-1}\big(x_{T-2}, \xi^2_{[1,3]}\big) = \begin{cases} 20, & \text{if } 10 \le x_{T-2} \le 20, \\ +\infty, & \text{otherwise,} \end{cases}$$

etc. By continuing the above backward calculations (which, of course, depend on numerical values of c_2, A_{21}, A_{22} and c_1, A_{11}) one can either show that the problem is infeasible or find the optimal value and an optimal solution of the first stage problem.

It is also possible to solve this multistage problem by formulating it as a linear programming problem of the form (3.8) subject to the corresponding nonanticipativity constraints. Such linear program will have $4 \times 8 = 32$ decision variables, 16 nonanticipativity constraints and four linear equality constraints.

Consider now a scenario tree and corresponding process $\xi_1, \ldots, \xi_T$. With each scenario of the tree is associated a probability p_k, $k = 1, \ldots, K$. These probabilities are related to the time structure of the multistage process and can be constructed as follows. In order to deal with the nested structure of problems (3.4) we need to specify the conditional distribution of ξ_t given[11]

[11] Since ξ_1 is not random, for $t = 2$ the distribution of ξ_2 is independent of ξ_1.

$\xi_{[1,t-1]}$, $t = 2, \ldots, T$. Consider a node $i \in \mathcal{N}$ and its ancestor $a = a(i)$ in the scenario tree. Denote by ρ_{ai} the probability of moving from the node a to the node i. For instance, in the tree of Fig. 2 it is possible to move from the root node to two nodes at stage $t = 2$, say i_1 and i_2, with the corresponding probabilities $\rho_{1i_1} = 0.4$ and $\rho_{1i_2} = 0.6$. Clearly the numbers ρ_{ai} should be nonnegative and for a given $a \in \mathcal{N}$ the sum of ρ_{ai} over all continuations $i \in \mathcal{N}$ of the node a should be equal to one. Each probability ρ_{ai} can be viewed as the conditional probability of the process being in the node i given its history up to the ancestor node $a = a(i)$. Note also that probabilities ρ_{ai} are in one-to-one correspondence with the arcs of the scenario tree. Every scenario can be defined by its nodes $i_1, i_2, \ldots, i_T$, arranged in the chronological order, i.e., node i_2 (at level $t = 2$) is connected to the root $i_1 = 1$, node i_3 is connected to the node i_2, etc. The probability of that scenario is then given by the product $\rho_{i_1 i_2}, \rho_{i_2 i_3}, \ldots, \rho_{i_{T-1} i_T}$. The conditional probabilities ρ_{ai} describe the probabilistic structure of the considered problem and could be specified together with the corresponding scenario tree.

It is possible to derive these conditional probabilities from scenario probabilities p_k as follows. Let us denote by $\mathcal{B}^{(i)}$ the set of scenarios passing through node i (at level t) of the scenario tree, and let $p^{(i)} := \mathbb{P}[\mathcal{B}^{(i)}]$. If $i_1, i_2, \ldots, i_t$, with $i_1 = 1$ and $i_t = i$, is the history of the process up to node i, then the probability $p^{(i)}$ is given by the product

$$p^{(i)} = \rho_{i_1 i_2}, \rho_{i_2 i_3}, \ldots, \rho_{i_{t-1} i_t}$$

of the corresponding conditional probabilities. In another way we can write this in the recursive form $p^{(i)} = \rho_{ai} p^{(a)}$, where $a = a(i)$ is the ancestor of the node i. This equation defines the conditional probability ρ_{ai} from the probabilities $p^{(i)}$ and $p^{(a)}$. Note that if $a = a(i)$ is the ancestor of the node i, then $\mathcal{B}^{(i)} \subset \mathcal{B}^{(a)}$ and hence $p^{(i)} \leq p^{(a)}$. Consequently if $p^{(a)} > 0$, then $\rho_{ai} = p^{(i)}/p^{(a)}$. Otherwise $\mathcal{B}^{(a)}$ is empty, i.e., no scenario is passing through the node a, and hence no scenario is passing through the node i.

Recall that a stochastic process Z_t, $t = 1, 2, \ldots$, that can take a finite number $\{z_1, \ldots, z_m\}$ of different values, is said to be a *Markov chain* if

$$\mathbb{P}\{Z_{t+1} = z_j | Z_t = z_i,\ Z_{t-1} = z_{i_{t-1}}, \ldots, Z_1 = z_{i_1}\} = p_{ij}, \tag{3.12}$$

for all states $z_{i_{t-1}}, \ldots, z_{i_1}, z_i, z_j$ and all $t = 1, \ldots$. In some instances it is natural to model the data process as a Markov chain with the corresponding state space[12] $\{\zeta^1, \ldots, \zeta^m\}$ and probabilities p_{ij} of moving from state ζ^i to state ζ^j, $i, j = 1, \ldots, m$. We can model such process by a scenario tree. At stage $t = 1$ there is one root node to which is assigned one of the values from the state space, say ζ^i. At stage $t = 2$ there are m nodes to which are assigned values

[12] In our modeling, values $\zeta^1, \ldots, \zeta^m$ can be numbers or vectors.

$\zeta^1, \ldots, \zeta^m$ with the corresponding probabilities $p_{i1}, \ldots, p_{im}$. At stage $t = 3$ there are m^2 nodes, such that each node at stage $t = 2$, associated with a state ζ^a, $a = 1, \ldots, m$, is the ancestor of m nodes at stage $t = 3$ to which are assigned values $\zeta^1, \ldots, \zeta^m$ with the corresponding conditional probabilities $p_{a1}, \ldots, p_{am}$. At stage $t = 4$ there are m^3 nodes, etc. At each stage t of such T-stage Markov chain process there are m^{t-1} nodes, the corresponding random vector (variable) ξ_t can take values $\zeta^1, \ldots, \zeta^m$ with respective probabilities which can be calculated from the history of the process up to time t, and the total number of scenarios is m^{T-1}. We have here that the random vectors (variables) $\xi_1, \ldots, \xi_T$ are independently distributed iff $p_{ij} = p_{i'j}$ for any $i, i', j = 1, \ldots, m$, i.e., the conditional probability p_{ij} of moving from state ζ^i to state ζ^j does not depend on i.

In the above formulation of the Markov chain the corresponding scenario tree represents the total history of the process with the number of scenarios growing exponentially with the number of stages. Now if we approach the problem by writing the cost-to-go functions $Q_t(x_{t-1}, \xi_t)$, going backwards, then we do not need to keep the track of the history of the process. That is, at every stage t the cost-to-go function $Q_t(\cdot, \xi_t)$ only depends on the current state (realization) $\xi_t = \zeta^i$, $i = 1, \ldots, m$, of the process. On the other hand, if we want to write the corresponding optimization problem (in the case of a finite number of scenarios) as one large linear programming problem, we still need the scenario tree formulation. This is the basic difference between the stochastic and dynamic programming approaches to the problem. That is, the stochastic programming approach does not necessarily rely on the Markovian structure of the considered process. This makes it more general at the price of considering a possibly very large number of scenarios.

There are many ways to express the nonanticipativity constraints (3.11) which may be convenient for different solution methods. One way is particularly elegant from the theoretical point of view:

$$x_t^k = \frac{\sum_{j \in \mathcal{A}_t(k)} p_j x_t^j}{\sum_{j \in \mathcal{A}_t(k)} p_j}, \quad k = 1, \ldots, K, \ t = 1, \ldots, T, \tag{3.13}$$

where $\mathcal{A}_t(k) := \{j : \xi^j_{[1,t]} = \xi^k_{[1,t]}\}$ is the set of scenarios that share with scenario k the history up to stage t. The expression at the right hand side of the above relation is the conditional expectation of x_t under the condition that $\xi_{[1,t]} = \xi^k_{[1,t]}$, where x_t is viewed as a random variable which can take values x_t^j with probabilities p_j, $j = 1, \ldots, K$. We can therefore rewrite (3.13) as

$$x_t = \mathbb{E}\big[x_t \mid \xi_{[1,t]}\big], \quad t = 1, \ldots, T. \tag{3.14}$$

This formulation of the nonanticipativity constraints can be conveniently extended to the case of a general distribution of the data $\xi_{[1,T]}$.

The nonanticipativity conditions (3.14) can be analytically eliminated from the multistage model. As before, denote by $\mathcal{N}$ the set of nodes of the scenario tree (with root 1), and let $i \in \mathcal{N}$ be a node at level t. Recall that $\mathcal{B}^{(i)}$ denotes the set of scenarios passing through node i and $a(i)$ denotes the ancestor of node i. We have that x_t has to be constant for scenarios $k \in \mathcal{B}^{(i)}$. Let us denote the value of x_t associated with node i by $x^{(i)}$. Similarly, let $c^{(i)}$, $D_i^{(i)}$, $W^{(i)}$ and $h^{(i)}$ be the values c_t, $A_{t,t-1}^k$, A_{tt}^k and b_t^k, respectively, corresponding to node i. We can rewrite then the corresponding linear programming problem as follows

$$\begin{aligned} &\text{Min} \sum_{i \in \mathcal{N}} p^{(i)} (c^{(i)})^T x^{(i)} \\ &\text{s.t.} \quad D^{(i)} x^{a(i)} + W^{(i)} x^{(i)} = h^{(i)}, \quad i \in \mathcal{N} \backslash \{1\}, \\ &\qquad W^{(1)} x^{(1)} = h^{(1)}, \\ &\qquad x^{(i)} \geq 0, \quad i \in \mathcal{N}. \end{aligned}$$

3.3 The general model

In the general multistage model, similarly to the linear case, we have a sequence of data vectors $\xi_1 \in \mathbb{R}^{d_1}, \xi_2 \in \mathbb{R}^{d_2}, \ldots, \xi_T \in \mathbb{R}^{d_T}$, and a sequence of decisions: $x_1 \in \mathbb{R}^{n_1}, x_2 \in \mathbb{R}^{n_2}, \ldots, x_T \in \mathbb{R}^{n_T}$. We assume that ξ_1 is already known and random vectors $\xi_2, \ldots, \xi_T$ are observed at the corresponding time periods. The decision process has then the form:

$$\begin{aligned} &\text{decision } (x_1) \rightsquigarrow \text{observation } (\xi_2) \rightsquigarrow \text{decision } (x_2) \rightsquigarrow \\ &\quad \cdots\cdots \rightsquigarrow \text{observation } (\xi_T) \rightsquigarrow \text{decision } (x_T). \end{aligned}$$

The values of the decision vector x_t, chosen at stage t, may depend on the information $\xi_{[1,t]}$ available up to time t, but not on the results of future observations. We can formulate this requirement using nonanticipativity constraints. That is, we view each $x_t = x_t(\cdot)$ as an element of the space of measurable mappings from Ω to $\mathbb{R}^{n_t}$, and hence consider $x_t(\omega)$ as a random (vector valued) process of time t. It has to satisfy the following additional condition, called the *nonanticipativity constraint*,

$$x_t = \mathbb{E}[x_t \mid \xi_{[1,t]}], \quad t = 1, \ldots, T. \tag{3.15}$$

If $\mathcal{F}_t$ is the sigma algebra generated by[13] $\xi_{[1,t]}$, then $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \ldots \subset \mathcal{F}_T \subset \mathcal{F}$, and condition (3.15) ensures that $x_t(\omega)$ is measurable with respect $\mathcal{F}_t$.

[13] $\mathcal{F}_t$ is the minimal subalgebra of the sigma algebra $\mathcal{F}$ such that $\xi_1(\omega), \ldots, \xi_t(\omega)$ are $\mathcal{F}_t$-measurable. Since ξ_1 is not random, $\mathcal{F}_1$ contains only two sets $\emptyset$ and Ω. We can assume that $\mathcal{F}_T = \mathcal{F}$.

One can use this measurability requirement as a definition of the nonanticipativity constraint.

To describe the objective and other constraints, let us denote the decisions associated with stages $1, \ldots, T$, as before, by $x_{[1,t]} := (x_1, \ldots, x_t)$. We have the objective functional

$$F: \mathbb{R}^{n_1 + \cdots + n_T} \times \mathbb{R}^{d_1 + \cdots + d_T} \longrightarrow \mathbb{R},$$

and constraint functionals

$$G_{ti}: \mathbb{R}^{n_1 + \cdots + n_t} \times \mathbb{R}^{d_1 + \cdots + d_t} \longrightarrow \mathbb{R}, \quad t = 2, \ldots, T, \ i = 1, \ldots, m_t.$$

The multistage stochastic programming problem is abstractly formulated as follows

$$\begin{aligned} &\text{Min } \mathbb{E}\big[F(x_{[1,T]}(\omega), \xi_{[1,T]}(\omega))\big] \\ &\text{s.t.} \quad G_{ti}(x_{[1,t]}(\omega), \xi_{[1,t]}(\omega)) \le 0, \quad i = 1, \ldots, m_t, \ t = 1, \ldots, T, \\ &\qquad x_t(\omega) \in X_t, \quad t = 1, \ldots, T, \\ &\qquad x_t = \mathbb{E}\big[x_t \mid \xi_{[1,t]}\big], \quad t = 1, \ldots, T. \end{aligned} \tag{3.16}$$

In the above formulation X_t is a convex closed subset of $\mathbb{R}^{n_t}$, and all constraints are assumed to hold almost surely.

The nested formulation can be developed similarly to the linear case. At stage T we know $\xi_{[1,T]}$ and $x_{[1,T-1]}$ and we have the problem

$$\begin{aligned} &\min_{x_T} \ F\big(x_{[1,T-1]}, x_T, \xi_{[1,T]}\big) \\ &\text{s.t.} \quad G_{Ti}(x_{[1,T-1]}, x_T, \xi_{[1,T]}) \le 0, \quad i = 1, \ldots, m_T, \\ &\qquad x_T \in X_T. \end{aligned} \tag{3.17}$$

Its optimal value is denoted $Q_T\big(x_{[1,T-1]}, \xi_{[1,T]}\big)$. Generally, at stage $t = T-1, \ldots, 1$ we have the problem

$$\begin{aligned} &\min_{x_t} \ \mathbb{E}\big[Q_{t+1}(x_{[1,t-1]}, x_t, \xi_{[1,t+1]}) \mid \xi_{[1,t]}\big] \\ &\text{s.t.} \quad G_{ti}\big(x_{[1,t-1]}, x_t, \xi_{[1,t]}\big) \le 0, \quad i = 1, \ldots, m_t, \\ &\qquad x_t \in X_t. \end{aligned} \tag{3.18}$$

Its optimal value is denoted $Q_t\big(x_{[1,t-1]}, \xi_{[1,t]}\big)$.

If F and G_{ti} are random lsc functions and the sets X_t are closed and bounded, then all Q_t are random lsc functions, too. This can be proved by

recursively applying Theorem 20 (from the Appendix) to problems (3.18) at stages $T, T-1, \ldots, 1$. By the forward induction, for $t = 1, \ldots, T$, we can also prove that each problem (3.18) has its data measurable with respect to $\mathcal{F}_t$ and has, by the measurable selection theorem (Theorem 16 in the Appendix), a solution which is $\mathcal{F}_t$-measurable (if a solution exists at all). Therefore, under natural assumptions, the multistage stochastic program (3.16) is a well defined model.

3.4 Examples of multistage models

Example 7 (Financial Planning). Suppose that there are n investment opportunities, with random returns $R_t = (R_{1t}, \ldots, R_{nt})$ in time periods $t = 1, \ldots, T$. One of possible investments is just cash. Our objective is to invest the given amount W_0 at time $t = 0$ so as to maximize the expected utility of our wealth at the last period T. The utility of wealth W is represented by a concave nondecreasing function $U(W)$. In our investment strategy we are allowed to rebalance our portfolio after each period, but without injecting additional cash into it.

Let $x_{10}, \ldots, x_{n0}$ denote the initial amounts invested in assets $1, \ldots, n$ at time $t = 0$. Clearly, they have to be nonnegative and to satisfy the condition

$$\sum_{i=1}^{n} x_{i0} = W_0. \tag{3.19}$$

We can put an equation sign here, because one of our assets is cash.

After the first period, our wealth may change, due to random returns from the investments, and at time $t = 1$ it will be equal to

$$W_1 = \sum_{i=1}^{n} (1 + R_{i1}) x_{i0}. \tag{3.20}$$

If we stop at that time, our problem becomes the stochastic programming problem

$$\begin{aligned} &\operatorname*{Max}_{x_0 \in \mathbb{R}^n} \mathbb{E}\left[U\left(\sum_{i=1}^{n} (1 + R_{i1}) x_{i0} \right) \right] \\ &\text{s.t.} \quad \sum_{i=1}^{n} x_{i0} = W_0, \\ &\qquad x_{i0} \geq 0, \quad i = 1, \ldots, n. \end{aligned} \tag{3.21}$$

In particular, if $U(W) \equiv W$, i.e., we want to maximize the expected wealth, then the objective function in the above problem (3.21) becomes

$\sum_{i=1}^{n}(1 + \mathbb{E}[R_{i1}])x_{i0}$, and hence problem (3.21) becomes a deterministic optimization program. It has the trivial optimal solution of investing everything into the asset with the maximum expected return.

Suppose, on the other hand, that $U(W)$ is defined as

$$U(W) := \begin{cases} (1 + q)(W - a), & \text{if } W \geq a, \\ (1 + r)(W - a), & \text{if } W \leq a, \end{cases} \tag{3.22}$$

with $r > q > 0$ and $a > 0$. We can view the involved parameters as follows: a is the amount that we have to pay at time $t = 1$, q is the interest at which we can invest the additional wealth $W - a$, provided that $W > a$, and r is the interest at which we will have to borrow if W is less than a. For the above utility function, problem (3.21) can be formulated as the following two-stage stochastic linear program

$$\begin{aligned} &\operatorname*{Max}_{x_0 \in \mathbb{R}^n} \ \mathbb{E}[Q(x_0, R_1)], \\ &\text{s.t.} \quad \sum_{i=1}^{n} x_{i0} = W_0, \\ &\qquad x_{i0} \geq 0, \quad i = 1, \ldots, n, \end{aligned} \tag{3.23}$$

where $Q(x_0, R_1)$ is the optimal value of the second stage program

$$\begin{aligned} &\operatorname*{Max}_{y,\, z \in \mathbb{R}} \ (1 + q)y - (1 + r)z \\ &\text{s.t.} \quad \sum_{i=1}^{n} (1 + R_{i1})x_{i0} = a + y - z, \\ &\qquad y \geq 0, \ z \geq 0. \end{aligned} \tag{3.24}$$

Suppose now that $T > 1$. In that case we can rebalance the portfolio at time $t = 1$, by specifying the amounts $x_{11}, \ldots, x_{n1}$ invested in the assets in the second period. Note that we already know the actual returns in the first period, so it is reasonable to use this information in the rebalancing decisions. Thus, our second stage decisions are actually functions of $R_1 = (R_{11}, \ldots, R_{n1})$, and they can be written as $x_{11}(R_1), \ldots, x_{n1}(R_1)$. We also must remember about our balance of wealth:

$$\sum_{i=1}^{n} x_{i1}(R_1) = W_1 \tag{3.25}$$

and the condition of nonnegativity. In general, the wealth after period t is equal to

$$W_t = \sum_{i=1}^{n} (1 + R_{it}) x_{i,t-1}(R_{[1,t-1]}), \tag{3.26}$$

where $R_{[1,t]} := (R_1, \ldots, R_t)$.

Our next decisions, $x_{1t}, \ldots, x_{nt}$ may depend on $R_1, \ldots, R_t$. They have to be nonnegative and satisfy the balance constraint,

$$\sum_{i=1}^{n} x_{it}(R_{[1,t]}) = W_t. \tag{3.27}$$

At the end, the wealth after period T is

$$W_T = \sum_{i=1}^{n} (1 + R_{iT}) x_{i,T-1}(R_{[1,T-1]}). \tag{3.28}$$

Our objective is to maximize the expected utility of this wealth,

$$\text{Max } \mathbb{E}[U(W_T)]. \tag{3.29}$$

It is a multistage stochastic programming problem, where stages are numbered from $t = 0$ to $t = T - 1$, and decisions x_t at each stage are allowed to depend on the history $R_1, \ldots, R_t$ of returns prior to this stage.

Of course, in order to complete the description of the above multistage stochastic programming problem, we need to define the probability structure of the random process $R_1, \ldots, R_T$. This can be done in many different ways. For example, one can construct a particular scenario tree defining time evolving of the process. If at every stage the random return of each asset is allowed to have just two continuations independently of other assets, then the total number of scenarios is 2^{nT}. It also should be ensured that $1 + R_{it} > 0$, $i = 1, \ldots, n$, $t = 1, \ldots, T$, for all possible realizations of the random data.

Let us consider the above multistage problem backwards, as it was discussed in Section 3.1. At the last stage $t = T - 1$ all realizations of the random process $R_1, \ldots, R_{T-1}$ are known and x_{T-2} has been chosen. Therefore, we have to solve the problem

$$\begin{aligned} &\text{Max } \mathbb{E}\left\{ U\left[\sum_{i=1}^{n} (1 + R_{iT}) x_{i,T-1} \right] \middle| R_{[1,T-1]} \right\} \\ &\text{s.t.} \quad \sum_{i=1}^{n} x_{i,T-1} = \sum_{i=1}^{n} (1 + R_{i,T-1}) x_{i,T-2}, \\ &\qquad x_{i,T-1} \geq 0, \quad i = 1, \ldots, n. \end{aligned} \tag{3.30}$$

Its optimal value is denoted $Q_{T-1}(x_{T-2}, R_{[1,T-1]})$. At stage $t = T-2$ realizations of the random process $R_1, \ldots, R_{T-2}$ are known and x_{T-3} has been chosen. We have then to solve the following two-stage stochastic program

$$\begin{aligned} &\text{Max } \mathbb{E}\big[Q_{T-1}(x_{T-2}, R_{[1,T-1]})|R_{[1,T-2]}\big] \\ &\text{s.t.} \quad \sum_{i=1}^{n} x_{i,T-2} = \sum_{i=1}^{n} (1 + R_{i,T-2})x_{i,T-3}, \\ &\qquad x_{i,T-2} \geq 0, \quad i = 1, \ldots, n. \end{aligned} \tag{3.31}$$

Its optimal value is denoted $Q_{T-2}(x_{T-3}, R_{[1,T-2]})$, etc. At stage $t = 0$ we have to solve the following program

$$\begin{aligned} &\text{Max } \mathbb{E}[Q_1(x_0, R_1)] \\ &\text{s.t.} \quad \sum_{i=1}^{n} x_{i0} = W_0, \\ &\qquad x_{i0} \geq 0, \quad i = 1, \ldots, n. \end{aligned} \tag{3.32}$$

Note that in the present case the cost-to-go function $Q_{T-1}(x_{T-2}, R_{[1,T-1]})$ depends on $x_{T-2} = (x_{1,T-2}, \ldots, x_{n,T-2})$ only through $W_{T-1} = \sum_{i=1}^{n} (1 + R_{i,T-1})x_{i,T-2}$. That is, if $\tilde{Q}_{T-1}(W_{T-1}, R_{[1,T-1]})$ is defined as the optimal value of the problem

$$\begin{aligned} &\text{Max } \mathbb{E}\left\{ U\left[\sum_{i=1}^{n} (1 + R_{iT})x_{i,T-1} \right] \middle| R_{[1,T-1]} \right\} \\ &\text{s.t.} \sum_{i=1}^{n} x_{i,T-1} = W_{T-1}, \quad x_{i,T-1} \geq 0, \quad i = 1, \ldots, n, \end{aligned} \tag{3.33}$$

then

$$Q_{T-1}(x_{T-2}, R_{[1,T-1]}) = \tilde{Q}_{T-1}\left(\sum_{i=1}^{n} (1 + R_{i,T-1})x_{i,T-2}, R_{[1,T-1]} \right).$$

Similarly, $Q_{T-2}(x_{T-3}, R_{[1,T-2]})$ depends on x_{T-3} only through W_{T-2}, and so on.

We may also note that the need for multistage modeling occurs here mainly because of the nonlinearity of the utility function $U(\cdot)$. Indeed, if $U(W) \equiv W$, and the returns in different stages are independent random vectors, it is

sufficient to maximize the expected wealth after each period, in a completely myopic fashion, by solving for $t = 0, \ldots, T-1$ the single stage models

$$\begin{aligned} &\underset{x_t}{\text{Max}}\ \mathbb{E}\left[\sum_{i=1}^{n}(1+R_{i,t+1})x_{i,t}|R_{[1,t]}\right] \\ &\text{s.t.}\quad \sum_{i=1}^{n} x_{it} = W_t, \quad x_t \geq 0, \end{aligned} \tag{3.34}$$

where W_t and $R_1, \ldots, R_t$ are already known. This, in turn, becomes a deterministic model with the objective coefficients

$$\alpha_{it}(R_{[1,t]}) := 1 + \mathbb{E}[R_{i,t+1}|R_{[1,t]}].$$

Such a model has a trivial optimal solution of investing everything in the asset with the maximum expected return in the next period.

A more realistic situation occurs in the presence of transaction costs. These are losses associated with the changes in the numbers of units (stocks, bonds) held. In such a situation multistage modeling is necessary, too, even if we use the expected wealth objective.

Let us observe now that the above problem can be also modeled as a T-period two-stage problem. To that end suppose that one makes a decision at the beginning of the process without thinking of rebalancing the portfolio. That is, our decision variables are initial amounts $x_1, \ldots, x_n$ invested in assets $1, \ldots, n$ at time $t = 0$. After T periods of time each asset i will be worth $\left[\prod_{t=1}^{T}(1+R_{it})\right]x_i$, and hence the total wealth will be $\sum_{i=1}^{n}\left[\prod_{t=1}^{T}(1+R_{it})\right]x_i$. The corresponding stochastic program can be then written as follows

$$\begin{aligned} &\underset{x\in R^n}{\text{Max}}\ \mathbb{E}\left[U\left(\sum_{i=1}^{n}\left[\prod_{t=1}^{T}(1+R_{it})\right]x_i\right)\right] \\ &\text{s.t.}\quad \sum_{i=1}^{n} x_i = W_0, \\ &\qquad x_i \geq 0, \quad i = 1, \ldots, n. \end{aligned} \tag{3.35}$$

Problem (3.35) is a two-stage stochastic program. It gives an extension of the two-stage problem (3.21) for T periods of time. If the utility function is given in the form (3.22), then problem (3.35) can be formulated as a linear two-stage stochastic program in a way similar to (3.23)–(3.24).

The difference between the two-stage (3.35) and multistage (3.29) programs is that in the two-stage model the value x_{it} of asset i at time t is

defined by the recursive equation[14] $x_{it} = (1 + R_{it})x_{i,t-1}$, which implies that $x_{it} = \left[\prod_{s=1}^{t}(1 + R_{is})\right]x_{i0}$. Consequently, x_{it} is completely determined by the initial value $x_{i0} = x_i$ and a realization of the random process $R_{i1}, \ldots, R_{it}$. On the other hand in the multistage model values x_{it} are rebalanced at every period of time subject to the constraints (3.26)–(3.27). Therefore the multistage problem (3.29) can be viewed as a relaxation of the two-stage problem (3.35), and hence has a larger optimal value.

We discuss further the above example in section "An Example of Financial Planning" of chapter "Monte Carlo Sampling Methods".

The following example also demonstrates that in some cases the same practical problem can be modeled as a multistage or two-stage multiperiod program.

Example 8 (Queueing Process). Consider stochastic process I_t, $t = 1, 2, \ldots$, governed by the recursive equation

$$I_t = [I_{t-1} + x_t - D_t]_+, \tag{3.36}$$

with initial value I_0. Here D_t are random numbers and x_t represent decision variables. The above process I_t can describe the waiting time of t-th customer in a $G/G/1$ queue, where D_t is the interarrival time between the $(t-1)$-th and t-th customers and x_t is the service time of $(t-1)$-th customer. Alternatively, we may view I_t as an inventory of a certain product at time t, with D_t and x_t representing the demand and production (or reordering), respectively, of the product at time t. Equation (3.36) assumes that the excess demand (over $I_{t-1} + x_t$) is not backordered, but simply lost.

Suppose that the process is considered over a finite horizon at periods $t = 1, \ldots, T$. Our goal then is to minimize (or maximize) the expected value of an objective function involving $I_1, \ldots, I_T$. For instance, one may be interested in maximizing a profit which at time t is given by $c_t \min[I_{t-1} + x_t, D_t] - h_t I_t$, where c_t and h_t are positive parameters representing the marginal profit and the holding cost, respectively, of the product at period t. The negative of the total profit is then given by

$$F(x, D) := \sum_{t=1}^{T} \left\{h_t I_t - c_t \min[I_{t-1} + x_t, D_t]\right\}.$$

Here $x = (x_1, \ldots, x_T)$ is a vector of decision variables, $D = (D_1, \ldots, D_T)$ is a random vector of the demands at periods $t = 1, \ldots, T$. By using the recursive

[14] This defines an implementable and feasible policy for the multistage problem (3.29), see section "Multistage Models" of Chapter "Optimality and Quality in Stochastic Programming" for the definition of implementable and feasible policies.

equation (3.36) it is straightforward to show that $F(x, D)$ can be also written in the form

$$F(x, D) = \sum_{t=1}^{T} q_t I_t - \sum_{t=1}^{T} c_t x_t - c_1 I_0,$$

where $q_t := h_t - c_{t+1} + c_t$, $t = 1, \ldots, T-1$, and $q_T := c_T + h_T$. We assume that all numbers q_t are positive, this certainly holds if $c_1 = \cdots = c_T$. By (3.36) we have that I_t is a convex function of $x_1, \ldots, x_t$. Since q_t are positive, it follows that the function $F(\cdot, D)$ is convex for any realization of D.

We can formulate a corresponding stochastic programming problem in several ways. First, suppose that the production cannot be changed during the process as some realizations of the demands become known. That is, a decision about production quantities $x_1, \ldots, x_T$ should be made before any realization of the demands $D_1, \ldots, D_T$ is available, and is not changed at times $t = 1, \ldots, T$. This leads to the problem of minimization of the expectation $\mathbb{E}[F(x, D)]$, which is taken with respect to the probability distribution of the random vector D. Although, we have here a multiperiod process, the above formulation can be viewed as a two-stage problem. In fact it can be formulated as a linear two-stage stochastic program as follows:[15]

$$\operatorname*{Min}_{x \geq 0} \left\{ -c^T x + \mathbb{E}[Q(x, D)] \right\}, \tag{3.37}$$

where $c = (c_1, \ldots, c_T)$ and $Q(x, D)$ is the optimal value of the problem

$$\begin{aligned} \operatorname*{Min}_{y \geq 0} \ & \sum_{t=1}^{T} q_t y_t \\ \text{s.t.} \ & y_{t-1} + x_t - D_t \leq y_t, \quad t = 1, \ldots, T, \\ & y_0 = I_0. \end{aligned} \tag{3.38}$$

Note that $I_t = I_t(x, D)$ is equal to y_t^*, $t = 1, \ldots, T$, where y^* is the optimal solution of (3.38).

Suppose now that the random vector D can take a finite number or realizations (scenarios) $D^1, \ldots, D^K$ with the corresponding probabilities $p_1, \ldots, p_K$. For example, if components D_t of the demand vector form a Markov chain with m possible realizations at each period, then the total

[15] Since I_0 does not depend on x, the term $c_1 I_0$ is omitted.

number of scenarios $K = m^T$. We can write then the two stage problem (3.37)–(3.38) as the linear problem (compare with (2.7)):

$$\begin{aligned} \text{Min} \; & -\sum_{t=1}^{T} c_t x_t + \sum_{k=1}^{K} p_k \left(\sum_{t=1}^{T} q_t y_t^k \right) \\ \text{s.t.} \; & y_{t-1}^k + x_t - D_t^k \le y_t^k, \quad t = 1, \ldots, T, \\ & x_t \ge 0, \; y_0^k = I_0, \; y_t^k \ge 0, \; t = 1, \ldots, T, \; k = 1, \ldots, K. \end{aligned} \tag{3.39}$$

Note that the optimal values of y_t^k in (3.39) represent $I_t(x, D^k)$. Since $I_t(x, D^k)$ depend only on the realization D^k up to time t, the nonanticipativity constraints with respect to y_t^k hold in (3.39) automatically.

On the other hand, depending on the flexibility of the production process, one can update production quantities at every time period $t = 1, \ldots, T$ using known realizations of the demand up to time t. This can be formulated as a multistage stochastic program where an optimal decision is made at every period of time based on available realizations of the random data. Consider the following relaxation of (3.39):

$$\begin{aligned} \text{Min} \; & \sum_{k=1}^{K} p_k \left[\sum_{t=1}^{T} \left(q_t y_t^k - c_t x_t^k \right) \right] \\ \text{s.t.} \; & y_{t-1}^k + x_t^k - D_t^k \le y_t^k, \quad t = 1, \ldots, T, \\ & x_t^k \ge 0, y_0^k = I_0, \; y_t^k \ge 0, \; t = 1, \ldots, T, \; k = 1, \ldots, K. \end{aligned} \tag{3.40}$$

By adding to the above problem (3.40) the nonanticipativity constraints associated with the scenario tree of the considered T-period process, we obtain the linear programming formulation of the corresponding multistage stochastic program.

Example 9 (Trucking). A trucking company serves n locations. For simplicity we assume that it takes exactly one day for a truck to go from one location to another, independently whether it is loaded or not. At the beginning of each day t, the company observes for each pair of locations, i and j, a random demand D_{ijt} for cargo to be shipped from i to j on day t. If they have a sufficient number of trucks at location i at this moment, they may take an order and ship the cargo. The revenue for shipping a unit of cargo from i to j is q_{ij}. The part of the demand that is not served is simply lost, and it does not result in any revenue or cost. It is important to stress that the numbers of trucks at different locations result from earlier decisions of moving the trucks and are therefore parts of the policy. The company may also move empty trucks between different locations (in anticipation of strong demand somewhere else). The cost of moving one unit of capacity from i to j is c_{ij},

independently whether it is loaded or empty (this is not a simplification, because we can always adjust the q_{ij}'s). Currently, the company has the capacity r_{i0} at each location i. Their objective is to maximize the expected profit in the next T days.

We recognize this problem as a multistage stochastic programming problem. With each day (stage) $t = 1, \ldots, T$ we associate the following decision variables:

y_{ijt} - the total capacity moved from i to j, where $i, j = 1, \ldots, n$,
z_{ijt} - the amount of cargo moved from i to j, where $i, j = 1, \ldots, n$,
r_{it} - the capacity available at i at the end of day t, where $i = 1, \ldots, n$.

Note that $D_{iit} = 0$ and $z_{iit} = 0$ by definition, and y_{iit} is the capacity waiting at i for the next day.

The problem takes on the form

$$\begin{aligned} \operatorname*{Max}_{y,z,r} \ & \mathbb{E}\left[\sum_{t=1}^{T}\sum_{i,j=1}^{n}(q_{ij}z_{ijt} - c_{ij}y_{ijt})\right] \\ \text{s.t.} \quad & z_{ijt} \le D_{ijt}, \quad i,j = 1,\ldots,n, \quad t = 1,\ldots,T, \\ & z_{ijt} \le y_{ijt}, \quad i,j = 1,\ldots,n, \quad t = 1,\ldots,T, \\ & r_{i,t-1} + \sum_{k=1}^{n} y_{kit} - \sum_{j=1}^{n} y_{ijt} = r_{it}, \quad i = 1,\ldots,n, \quad t = 1,\ldots,T, \\ & r \ge 0,\ y \ge 0,\ z \ge 0. \end{aligned} \tag{3.41}$$

In a more refined version we may want to put some additional constraints on the capacity r_{iT} available at each location i at the end of the planning period.

In the above problem the demand $D(t) = [D_{ij}(t)]_{i,j=1,\ldots,n}$ is a random vector valued process. The decisions y_{ijt} and z_{ijt} and the resulting numbers of trucks r_{it} at different locations may depend on all past and current demand values $D(\tau)$, $\tau \le t$, but not on the future values of the demand vector. Therefore, at stage t, we cannot exactly predict how many trucks we shall need at each location at stage $t + 1$; we can only use past data and our knowledge of the joint distribution of all demands to re-position our truck fleet. For a specified scenario tree of the demand process $D(t)$, the corresponding multistage problem can be written as a large linear program.

3.5 Relations to dynamic programming

There exist close relations between multistage stochastic programming models and classical models of dynamic programming and optimal control.

To illustrate these relations, consider the linear dynamical system described by the state equation

$$s_{t+1} = A_t s_t + B_t u_t + C_t e_t, \quad t = 1, \ldots, T,$$

in which s_t denotes the state of the system at time t, u_t is the control vector, and e_t is a random 'disturbance' at time t. The matrices A_t, B_t and C_t are known. The random vectors e_t, $t = 1, \ldots, T$, are assumed to be independent. At time t we observe the current state value, s_t, but *not* the disturbances e_t. Our objective is to find a *control law*, $\hat{u}_t(\cdot)$, $t = 1, \ldots, T$, so that the actual values of the control variables can be determined through the *feedback* rule:

$$u_t = \hat{u}_t(s_t), \quad t = 1, \ldots, T-1.$$

We want to do it in such a way that the expected value of the performance index,

$$\mathbb{E}\left[\sum_{t=1}^{T-1} F_t(s_t, u_t) + F_T(s_T)\right]$$

is minimized. In a more involved formulation, there may be additional constraints on the control variables, or mixed state–control constraints:

$$g_{ti}(s_t, u_t) \leq 0, \quad i = 1, \ldots, m_t, \quad t = 1, \ldots, T-1.$$

For the sake of simplicity we assume that they are all incorporated into the definition of the partial objectives, that is, $F_t(s_t, u_t) = +\infty$ if these constraints are not satisfied.

The crucial characteristics of the optimal control model is that we look for a solution in the form of a function of the state vector. We are allowed to focus on such a special form of the control rule due to the independence of the disturbances at different stages. If the disturbances are dependent in certain ways, augmentation of the state space may reduce the model to the case of independent e_t's.

The key role in the optimal control theory is played by the *cost-to-go* function

$$V_t(s_t) := \inf \mathbb{E}\left[\sum_{\tau=t}^{T-1} F_\tau(s_\tau, u_\tau) + F_T(s_T)\right],$$

where the minimization is carried out among all possible feedback laws applied at stages $t, \ldots, T-1$. The functions $V_t(\cdot)$ give the *dynamic programming equation*:

$$V_t(s_t) = \inf_{u_t}(F_t(s_t, u_t) + \mathbb{E}[V_{t+1}(A_t s_t + B_t u_t + C_t e_t)]), \quad t = T-1, \ldots, 1.$$

The optimal feedback rule is the minimizer of the above expression.

Except for very special cases, such as linear–quadratic or time optimal control, the form of the optimal feedback rule may be very involved. Usually, some functional form of the rule is assumed and parametric optimization employed to find the best rule within a chosen class. Discretization of the state space is a common approach here.

To transform the above model into a stochastic programming model we just need to make the substitutions:

$$\begin{aligned} &x_t = (u_t, s_t), \quad t = 1, \ldots, T-1, \\ &x_T = s_T, \\ &\xi_t = C_{t-1} e_{t-1}, \quad t = 2, \ldots, T. \end{aligned}$$

The function $V_t(\cdot)$ can be formally expressed as the optimal value of

$$\begin{aligned} &\operatorname*{Min}_{s_t, u_t} \left(F_t(s_t, u_t) + \mathbb{E}[V_{t+1}(A_t s_t + B_t u_t + C_t e_t)]\right) \\ &\text{s.t.} \quad s_t = A_{t-1} s_{t-1} + B_{t-1} u_{t-1} + \xi_t. \end{aligned}$$

Thus, we can define

$$Q_t(s_{t-1}, u_{t-1}, \xi_t) = V_t(A_{t-1} s_{t-1} + B_{t-1} u_{t-1} + \xi_t)$$

to perfectly match both models.

The opposite of that is also true. A multistage stochastic programming model with model state variables and independent random parameters ξ_t can be transformed into a control problem, as in the following example.

Example 10 (Trucking (continued)). Let us consider Example 9 in which the demand vectors D_{ijt}, $i, j = 1, \ldots, n$ are independent for $t = 1, \ldots, T$. We can formally define:

$$s_t := [r_{t-1}, D_t], \quad u_t := [y_t, z_t], \quad e_t := D_{t+1}.$$

Then the next state s_{t+1} is a function of s_t, u_t and e_t:

$$r_{it} = r_{i,t-1} + \sum_{k=1}^{n} y_{kit} - \sum_{j=1}^{n} y_{ijt},$$
$$D_{t+1} = e_t.$$

At each time $t = 1, \ldots, T$ we have mixed state–control constraints:

$$z_{ijt} \le D_{ijt}, \quad i,j = 1, \ldots, n,$$
$$z_{ijt} \le y_{ijt}, \quad i,j = 1, \ldots, n.$$

The objective functional has the form:

$$F_t(s_t, u_t) = \sum_{i,j=1}^{n} (q_{ij} z_{ijt} - c_{ij} y_{ijt}),$$

and depends on controls alone. So, if the demands in different days are independent, the optimal solution has the form of a feedback rule:

$$y_t = \hat{y}_t(r_{t-1}, D_t),$$
$$z_t = \hat{z}_t(r_{t-1}, D_t).$$

The form of these functions is rather involved, though.

As we shall see it later, the stochastic programming formulation tries to exploit as much as possible some advantageous properties of the functions $V_t(\cdot)$ or $Q_t(\cdot)$, such as convexity, or polyhedral structure, which are hard to exploit in the dynamic programming setting. Also, the stochastic programming model does not assume the independence of the random disturbances. It does require, though, in the scenario tree formulation the discretization of the disturbances distributions.

4 Robust and min–max approaches to stochastic optimization

4.1 Robust models

Consider the two-stage stochastic linear program (2.2)–(2.3). In that problem the optimal value $Q(x, \xi(\omega))$ of the second stage problem is optimized on average. Of course, for a particular realization ξ of the random data $\xi(\omega)$ the corresponding value $Q(x, \xi)$ can be quite different from the expected value $\mathbb{E}[Q(x, \xi(\omega))]$. An ''unlucky'' realization of $\xi(\omega)$ may have disastrous

consequences for the user of stochastic programming. For instance, in Example 1 the newsvendor may loose all his savings on an unlucky day, so that he will have to borrow from the mob on murderous interest to continue his business next day. In order to avoid such disastrous consequences one may try to be more conservative and to reach a compromise between the average (i.e, the mean) and a risk associated with variability of $Q(x,\xi)$. It appears then natural to add the term $\kappa\mathbb{V}\mathrm{ar}[Q(x,\xi)]$ to the objective of the optimization problem, where coefficient $\kappa \geq 0$ represents a compromise between the expectation and variability of the objective. Unfortunately, this destroys two important properties of the two-stage linear program (2.2)–(2.3), namely its convexity and second stage optimality.

In order to see that let us suppose for the sake of simplicity that there is a finite number of scenarios and hence the problem can be formulated in the form (2.7). By adding the term $\kappa\mathbb{V}\mathrm{ar}[Q(x,\xi)]$ to the objective function in (2.2) we obtain the problem

$$\begin{aligned} &\operatorname*{Min}_{x} \; c^T x + \psi(Q(x,\xi_1),\ldots,Q(x,\xi_K)) \\ &\text{s.t.} \quad Ax = b,\ x \geq 0, \end{aligned} \tag{4.1}$$

where

$$\psi(z) := \sum_{k=1}^{K} p_k z_k + \kappa\left[\sum_{k=1}^{K} p_k z_k^2 - \left(\sum_{k=1}^{K} p_k z_k\right)^2\right].$$

Now for $\kappa > 0$ the objective function of the above problem is not necessarily convex even though the functions $Q(\cdot,\xi_i)$, $i = 1,\ldots,K$, are all convex, and the second stage optimality does not hold in the sense that problem (4.1) is not equivalent to the problem

$$\begin{aligned} &\operatorname*{Min}_{x,y_1,\ldots,y_k} \; c^T x + \psi\big(q_1^T y_1,\ldots,q_K^T y_K\big) \\ &\text{s.t.} \quad Ax = b, \\ &\qquad T_k x + W_k y_k = h_k, \\ &\qquad x \geq 0,\ y_k \geq 0,\ k = 1,\ldots,K. \end{aligned} \tag{4.2}$$

In order to preserve the property of second stage optimality we may change the function $\psi(z)$ to a componentwise nondecreasing function. Recall that a function $\psi : \mathbb{R}^K \to \mathbb{R}$ is said to be *componentwise nondecreasing* if $\psi(z) \geq \psi(z')$ for any $z, z' \in \mathbb{R}^K$ such that $z \geq z'$.

Proposition 11. *Suppose that problem (4.2) is feasible and function $\psi(z)$ is componentwise nondecreasing. Then problems (4.1) and (4.2) have the same optimal value, and if, moreover, problem (4.2) has an optimal solution, then problems (4.1) and (4.2) have the same set of first stage optimal solutions.*

Proof. Since (4.2) is feasible it follows that there exists a feasible x such that all $Q(x,\xi_k)$, $k=1,\ldots,K$, are less than $+\infty$, and hence the optimal value of problem (4.1) is also less than $+\infty$. By (2.6) we have that $Q(x,\xi_k)$ is given by the optimal value of a linear programming problem. Therefore, if $Q(x,\xi_k)$ is finite, then the corresponding linear programming problem has an optimal solution. It follows that if all $Q(x,\xi_k)$ are finite, then $\psi(Q(x,\xi_1),\ldots,Q(x,\xi_K))$ is equal to $\psi(q_1^Ty_1,\ldots,q_K^Ty_K)$ for some y_k, $k=1,\ldots,K$, satisfying the constraints of problem (4.2) and hence the optimal value of (4.1) is greater than or equal to the optimal value of (4.2). Conversely, for a given x, $Q(x,\xi_k)$ is less than or equal to $q_k^Ty_k$, $k=1,\ldots,K$, for any $y_1,\ldots,y_k$ feasible for (4.2). Since $\psi(z)$ is componentwise nondecreasing, it follows that the optimal value of (4.2) is greater than or equal to the optimal value of (4.1), and hence these two optimal values are equal to each other. Moreover, it follows that if $x^*,y_1^*,\ldots,y_K^*$ is an optimal solution of problem (4.2), then x^* is an optimal solution of problem (4.1), and vice versa. □

We also have that if $\psi(z)$ is componentwise nondecreasing and convex, then since functions $Q(\cdot,\xi_k)$, $k=1,\ldots,K$ are convex, the corresponding composite function and hence the objective function of problem (4.1) are convex.

Of course, for $\psi(z):=\sum_{k=1}^K p_kz_k$ problem (4.2) coincides with the two-stage linear problem (2.7). Another possibility is to use a separable function $\psi(z)=\sum_{k=1}^K \psi_k(z_k)$ with one of the following two choices of functions ψ_k:

$$\psi_k(z_k):=p_kz_k+\kappa p_k(z_k-\alpha)_+, \tag{4.3}$$

$$\psi_k(z_k):=p_kz_k+\kappa p_k[(z_k-\alpha)_+]^2, \tag{4.4}$$

for some $\kappa\geq 0$ and $\alpha\in\mathbb{R}$. Note that for both above choices of ψ_k, the corresponding function $\psi(z)$ is componentwise nondecreasing and convex.

If the parameter α in (4.4) is equal to $\mathbb{E}[Q(x,\xi)]$ and the distribution of $Q(x,\xi)$ is symmetrical around its mean, then

$$\psi(Q(x,\xi_1),\ldots,Q(x,\xi_K))=\mathbb{E}[Q(x,\xi)]+(\kappa/2)\mathbb{V}\mathrm{ar}[Q(x,\xi)].$$

Of course, the mean (expected value) of $Q(x,\xi)$ depends on x; in practical applications it would have to be iteratively adjusted during an optimization procedure. An advantage of using ψ_k given in (4.3) is that then the function

$\psi(z)$ is piecewise linear, and hence (4.2) can be formulated as a linear programming problem. The above approach to stochastic programming is called *robust* by some authors.

The model (4.2) with (4.3) or (4.4) is an example of a *mean–risk model*. For a random outcome $F(x,\xi)$, these models use an objective which is composed of two parts: the expected outcome (the *mean*) $\mathbb{E}[F(x,\xi)]$, and a scalar composite measure of the size and frequency of undesirable outcome values, the *risk* $\rho(F(x,\xi))$. The risk measure $\rho(Z)$ is understood here as a function of the *entire distribution* of the random variable Z. For example, our formulas (4.3) and (4.4) correspond to risk measures

$$\rho_1(Z;\alpha) := \mathbb{E}[(Z-\alpha)_+] \quad \text{and} \quad \rho_2(Z;\alpha) := \mathbb{E}\big[\big((Z-\alpha)_+\big)^2\big],$$

respectively, which represent the expected excess (or square excess) over the target level α. More sophisticated are *semideviation measures*, which use, instead of the fixed target level α, the expected value of the random outcome. The simplest and most convenient in applications is the *absolute semideviation*:

$$\rho_1(Z) := \mathbb{E}[(Z-\mathbb{E}Z)_+]. \tag{4.5}$$

The presence of the expected value of the outcome in the definition of the measure makes the resulting risk term

$$\rho_1(F(x,\xi)) = \mathbb{E}\big[(F(x,\xi)-\mathbb{E}[F(x,\xi)])_+\big],$$

a nonconvex function of x, even if $F(\cdot,\xi)$ is convex. Nevertheless, the corresponding mean–risk model

$$\operatorname{Min}\big\{\mathbb{E}[F(x,\xi)]+\kappa\mathbb{E}\big[(F(x,\xi)-\mathbb{E}[F(x,\xi)])_+\big]\big\}$$

remains a convex problem, provided that the coefficient κ in front of the risk term is confined to $[0,1]$. This can be seen from the representation:

$$\begin{aligned}&\mathbb{E}[F(x,\xi)]+\kappa\mathbb{E}\big[(F(x,\xi)-\mathbb{E}[F(x,\xi)])_+\big]\\ &\quad=(1-\kappa)\mathbb{E}[F(x,\xi)]+\kappa\mathbb{E}[\max\{\mathbb{E}[F(x,\xi)],F(x,\xi)\}],\end{aligned}$$

in which the convexity of all terms is evident.

Example 12. Let us consider Example 3 again, but instead of bounding our risk of loss by the probabilistic constraint $\mathbb{P}\{\sum_{i=1}^n R_i x_i \geq -b\} \geq 1-\alpha$, let us modify the objective by subtracting a risk measure

$$\rho\left(\sum_{i=1}^n R_i x_i\right). \tag{4.6}$$

For example, similarly to (4.3), we may use

$$\rho(Z) := \mathbb{E}[(\alpha - Z)_+], \tag{4.7}$$

in which case (4.6) represents the expected shortfall below some target profit level α. If $\alpha = -b < 0$, our measure represents the expected loss in excess of b.

Supposing that our initial capital (wealth) is W, we may formulate the following mean–risk optimization problem

$$\begin{aligned} &\operatorname*{Max}_{x\geq 0} \left\{\sum_{i=1}^n \mu_i x_i - \kappa\rho\left(\sum_{i=1}^n R_i x_i\right)\right\} \\ &\text{s.t.} \quad \sum_{i=1}^n x_i \leq W. \end{aligned} \tag{4.8}$$

Problems of this type are usually solved as a family parametrized by $\kappa \geq 0$. Their solutions can be graphically depicted in the form of the *efficient frontier*: the collection of mean–risk pairs corresponding to the optimal solutions of (4.8) for all $\kappa \geq 0$.

If the risk measure (4.7) is used, the term $\kappa\mathbb{E}\big[\big(\alpha - \sum_{i=1}^n R_i x_i\big)_+\big]$ can be interpreted as the expected cost of a loan to cover the shortfall below α, where κ is the interest rate. In this case problem (4.8) has a convenient linear programming formulation, provided that the distribution of the returns is discrete. It is very similar to the model for the semideviation risk measure discussed below.

As we deal with a maximization problem, the semideviation risk measure (4.5) should be modified to represent the shortfall below the mean:

$$\rho_1(Z) := \mathbb{E}[(\mathbb{E}Z - Z)_+].$$

Then the mean–risk model (4.7) takes on the form

$$\begin{aligned} &\operatorname*{Max}_{x\geq 0} \left\{(1-\kappa)\sum_{i=1}^n \mu_i x_i + \kappa\mathbb{E}\left[\min\left(\sum_{i=1}^n \mu_i x_i, \sum_{i=1}^n R_i x_i\right)\right]\right\} \\ &\text{s.t.} \quad \sum_{i=1}^n x_i \leq W. \end{aligned}$$

For a discrete distribution of R we can convert the above mean–risk model into a linear programming problem. Indeed, let $k = 1, \ldots, K$ denote scenarios, and let R_{ik} be the realization of the return of security i in scenario k. The probabilities of scenarios are $p_1, \ldots, p_K$, $\sum_{k=1}^{K} p_k = 1$. Introducing new variables μ (representing the mean), and r_k, $k = 1, \ldots, K$ (representing the worst case of return and its expected value) we obtain the problem

$$
\begin{aligned}
&\operatorname*{Max}_{x \geq 0,\, \mu,\, r} \left\{ (1-\kappa)\mu + \kappa \sum_{k=1}^{K} p_k r_k \right\} \\
&\text{s.t.} \quad \sum_{i=1}^{n} \mu_i x_i = \mu, \\
&\qquad r_k \leq \mu, \quad k = 1, \ldots, K, \\
&\qquad r_k \leq \sum_{i=1}^{n} R_{ik} x_i, \quad k = 1, \ldots, K, \\
&\qquad \sum_{i=1}^{n} x_i \leq W.
\end{aligned}
$$

It can be solved by standard linear programming techniques.

4.2 *Min–max stochastic programming*

In practical applications probability distributions of the involved uncertain parameters are never known exactly and can be estimated at best. Even worse, quite often the probabilities are assigned on an *ad hoc* basis by a subjective judgment. Suppose now that there is a set $\mathbb{S}$ of probability distributions, defined on a sample space $(\Omega, \mathcal{F})$, which in some reasonable sense give a choice of the underlying probability distributions. For instance, in Example 3 one may foresee that the random investment returns will generally increase, stay flat or even decrease over the next T years. By specifying means, representing a possible trend, and variability of the investment returns one may assign a finite number of possible probability distributions for the random data. Alternatively, certain properties, like first- and maybe second-order moments, unimodality or specified marginal distributions of the random data can be postulated. Typically, this leads to an infinite set $\mathbb{S}$ of considered probability distributions.

There are two basic ways of dealing with such cases of several distributions. One can assign *a priori* probability distribution over $\mathbb{S}$, and hence reduce the problem to a unique distribution. Suppose, for example, that $\mathbb{S}$ is finite, say $\mathbb{S} := \{P_1, \ldots, P_\ell\}$. Then by assigning probability ρ_i to P_i, $i = 1, \ldots, \ell$, one obtains the unique (posteriori) distribution $P := \sum_{i=1}^{\ell} \rho_i P_i$.

The distribution P represents an averaging over possible distributions P_i. Again a choice of the a priori distribution $\{\rho_1, \ldots, \rho_\ell\}$ is often subjective.

An alternative approach is to hedge against the worst distribution by formulating the following min–max analogue of stochastic programs (1.14), (2.12):

$$\operatorname*{Min}_{x \in \mathcal{X}} \operatorname*{Max}_{P \in \mathbb{S}} \mathbb{E}_P[F(x, \omega)]. \tag{4.9}$$

For the above problem to make sense it is assumed, of course, that for every $P \in \mathbb{S}$ the expectation $\mathbb{E}_P[F(x, \omega)]$ is well defined for all $x \in \mathcal{X}$.

In order to see a relation between these two approaches let us assume for the sake of simplicity that the set $\mathbb{S} = \{P_1, \ldots, P_\ell\}$ is finite. Then problem (4.9) can be written in the following equivalent way

$$\begin{aligned} &\operatorname*{Min}_{(x,z) \in X \times \mathbb{R}} \; z \\ &\text{s.t.} \quad f_i(x) \leq z, \quad i = 1, \ldots, \ell, \end{aligned} \tag{4.10}$$

where $f_i(x) := \mathbb{E}_{P_i}[F(x, \omega)]$. Suppose further that problem (4.10), and hence problem (4.9), is feasible and for every $\omega \in \Omega$ the function $F(\cdot, \omega)$ is convex. It follows from convexity of $F(\cdot, \omega)$ that the functions $f_i(\cdot)$ are also convex, and hence problem (4.10) is a convex programming problem. Then, by the duality theory of convex programming, there exist Lagrange multipliers $\lambda_i \geq 0$, $i = 1, \ldots, \ell$, such that $\sum_{i=1}^{\ell} \lambda_i = 1$ and problem (4.10) has the same optimal value as the problem

$$\operatorname*{Min}_{x \in \mathcal{X}} \left\{ f(x) := \sum_{i=1}^{\ell} \lambda_i f_i(x) \right\}$$

and the set of optimal solutions of (4.10) is included in the set of optimal solutions of the above problem. Since $f(x) = \mathbb{E}_{P^*}[F(x, \omega)]$, where $P^* := \sum_{i=1}^{\ell} \lambda_i P_i$, we obtain that problem (4.9) is equivalent to the stochastic programming problem

$$\operatorname*{Min}_{x \in \mathcal{X}} \mathbb{E}_{P^*}[F(x, \omega)].$$

This shows that, under the assumption of convexity, the min–max approach automatically generates an a priori distribution given by the corresponding Lagrange multipliers. Of course, in order to calculate these Lagrange multipliers one still has to solve the min–max problem. Existence of such Lagrange multipliers, and hence of the a priori distribution, can be also shown for an infinite set $\mathbb{S}$ under the assumption of convexity and mild regularity conditions.

5 Appendix

In this section we briefly discuss some basic concepts and definitions from probability and optimization theories, needed for the development of stochastic programming models. Of course, a careful derivation of the required results goes far beyond the scope of this book. The interested reader may look into standard textbooks for a thorough development of these topics.

5.1 Random variables

Let Ω be an abstract set. It is said that a set $\mathcal{F}$ of subsets of Ω is a *sigma algebra* (also called sigma field) if it is closed under standard set theoretic operations, the set Ω belongs to $\mathcal{F}$, and if $A_i \in \mathcal{F}$, $i \in \mathbb{N}$, then $\cup_{i\in\mathbb{N}} A_i \in \mathcal{F}$. The set Ω equipped with a sigma algebra $\mathcal{F}$ is called a *sample* or *measurable* space and denoted $(\Omega, \mathcal{F})$. A set $A \subset \Omega$ is said to be *$\mathcal{F}$-measurable* if $A \in \mathcal{F}$. It is said that the sigma algebra $\mathcal{F}$ is *generated* by its subset $\mathcal{A}$ if any $\mathcal{F}$-measurable set can be obtained from sets belonging to $\mathcal{A}$ by set theoretic operations and by taking the union of a countable family of sets from $\mathcal{A}$. That is, $\mathcal{F}$ is generated by $\mathcal{A}$ if $\mathcal{F}$ is the smallest sigma algebra containing $\mathcal{A}$.

If Ω coincides with a finite dimensional space $\mathbb{R}^m$, unless stated otherwise, we always equip it with its Borel sigma algebra $\mathcal{B}$. Recall that $\mathcal{B}$ is generated by the set of open (or closed) subsets of $\mathbb{R}^m$. A function $P : \mathcal{F} \to [0, 1]$ is called a probability measure on $(\Omega, \mathcal{F})$ if $P(\Omega) = 1$, and for every collection $A_i \in \mathcal{F}$, $i \in \mathbb{N}$, such that $A_i \cap A_j = \emptyset$ for all $i \neq j$, we have $P(\cup_{i\in\mathbb{N}} A_i) = \sum_{i\in\mathbb{N}} P(A_i)$. A sample space $(\Omega, \mathcal{F})$ equipped with a probability measfure P is called a *probability space* and denoted $(\Omega, \mathcal{F}, P)$. Recall that $\mathcal{F}$ is said to be P-complete if $A \subset B$, $B \in \mathcal{F}$ and $P(B) = 0$, implies that $A \in \mathcal{F}$, and hence $P(A) = 0$. Since it is always possible to enlarge the sigma algebra and extend the measure in such a way as to get complete space, we can assume without loss of generality that considered probability measures are *complete*. It is said that an event $A \in \mathcal{F}$ happens *P-almost surely* (a.s.) or *almost everywhere* (a.e.) if $P(A) = 1$, or equivalently $P(\Omega\backslash A) = 0$.

A mapping $V : \Omega \to \mathbb{R}^m$ is said to be *measurable* if for any Borel set $A \in \mathcal{B}$, its inverse image $V^{-1}(A) := \{\omega \in \Omega : V(\omega) \in A\}$ is $\mathcal{F}$-measurable.[16] A measurable mapping $V(\omega)$ from probability space $(\Omega, \mathcal{F}, P)$ into $\mathbb{R}^m$ is called a *random vector*. Note that the mapping V generates the probability measure (also called the probability distribution) $\mathbb{P}(A) := P(V^{-1}(A))$ on $(\mathbb{R}^m, \mathcal{B})$, which provides all relevant probabilistic information about the considered random vector. Clearly an event $A \in \mathcal{B}$ happens P-almost surely iff the corresponding event $V^{-1}(A) \in \mathcal{F}$ happens $\mathbb{P}$-almost surely. In particular, a measurable mapping (function) $Z : \Omega \to \mathbb{R}$ is called a *random variable*. Its probability

[16] In fact it suffices to verify $\mathcal{F}$-measurability of $V^{-1}(A)$ for any family of sets generating the Borel sigma algebra of $\mathbb{R}^m$.

distribution is completely defined by the cumulative distribution function (cdf) $F_Z(z) := \mathbb{P}\{Z \leq z\}$. Note that since the Borel sigma algebra of $\mathbb{R}$ is generated by the family of half line intervals $(-\infty, a]$, in order to verify measurability of $Z(\omega)$ it suffices to verify measurability of sets $\{\omega \in \Omega\colon Z(\omega) \leq z\}$ for all $z \in \mathbb{R}$. We denote random vectors (variables) by capital letters, like V, Z etc., or $\xi(\omega)$, and often suppress their explicit dependence on the elementary event ω. Also quite often we denote by the same symbol ξ a particular realization of the random vector $\xi = \xi(\omega)$. Usually, the meaning of such notation will be clear from the context and will not cause any confusion. The coordinate functions $V_1(\omega), \ldots, V_m(\omega)$ of the m-dimensional random vector $V(\omega)$ are called its components. While considering a random vector V we often talk about its probability distribution as the joint distribution of its components (random variables) $V_1, \ldots, V_m$.

Since we often deal with random variables which are given as optimal values of optimization problems we need to consider random variables $Z(\omega)$ which can also take values $+\infty$ or $-\infty$, i.e., functions $Z\colon \Omega \to \overline{\mathbb{R}}$, where $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty\} \cup \{+\infty\}$ denotes the set of extended real numbers. Such functions $Z\colon \Omega \to \overline{\mathbb{R}}$ are referred to as *extended real valued* functions. Operations between real numbers and symbols $\pm\infty$ are clear except for such operations as adding $+\infty$ and $-\infty$ which should be avoided. Measurability of an extended real valued function $Z(\omega)$ is defined in the standard way, i.e., $Z(\omega)$ is measurable if the set $\{\omega \in \Omega\colon Z(\omega) \leq z\}$ is $\mathcal{F}$-measurable for any $z \in \mathbb{R}$. A measurable extended real valued function is called an (extended) random variable. Note that here $\lim_{z \to +\infty} F_Z(z)$ is equal to the probability of the event $\{\omega \in \Omega\colon Z(\omega) < +\infty\}$ and can be less than one if the event $\{\omega \in \Omega\colon Z(\omega) = +\infty\}$ has a positive probability.

The *expected value* or *expectation* of an (extended) random variable $Z\colon \Omega \to \overline{\mathbb{R}}$ is defined by the integral

$$\mathbb{E}_P[Z] := \int_\Omega Z(\omega)\, \mathrm{d}P(\omega). \tag{5.1}$$

When there is no ambiguity as to what probability measure is considered, we omit the subscript P and simply write $\mathbb{E}[Z]$. For a nonnegative valued measurable function $Z(\omega)$ such that the event $\Upsilon := \{\omega \in \Omega\colon Z(\omega) = +\infty\}$ has zero probability the above integral is defined in the usual way and can take value $+\infty$. If probability of the event Υ is positive, then, by definition, $\mathbb{E}[Z] = +\infty$. For a general (not necessarily nonnegative valued) random variable we would like to define[17] $\mathbb{E}[Z] := \mathbb{E}[Z_+] - \mathbb{E}[(-Z)_+]$. In order to do that we have to ensure that we do not add $+\infty$ and $-\infty$. We say that the expected value $\mathbb{E}[Z]$ of an (extended real valued) random variable $Z(\omega)$ is *well defined* if it does not happen that both $\mathbb{E}[Z_+]$ and $\mathbb{E}[(-Z)_+]$ are $+\infty$, in which case $\mathbb{E}[Z] = \mathbb{E}[Z_+] - \mathbb{E}[(-Z)_+]$. That is, in order to verify that the expected

[17] Recall that $Z_+ := \max\{0, Z\}$.

value of $Z(\omega)$ is well defined one has to check that $Z(\omega)$ is measurable and either $\mathbb{E}[Z_+] < +\infty$ or $\mathbb{E}[(-Z)_+] < +\infty$. Note that if $Z(\omega)$ and $Z'(\omega)$ are two (extended) random variables such that their expectations are well defined and $Z(\omega) = Z'(\omega)$ for all $\omega \in \Omega$ except possibly on a set of measure zero, then $\mathbb{E}[Z] = \mathbb{E}[Z']$. It is said that $Z(\omega)$ is *P-integrable* if the expected value $\mathbb{E}[Z]$ is well defined and *finite*. The expected value of a random vector is defined componentwise.

If the random variable $Z(\omega)$ can take only a countable (finite) number of different values, say $z_1, z_2, \ldots,$ then it is said that $Z(\omega)$ has a *discrete* distribution (*discrete* distribution with a *finite support*). In such cases all relevant probabilistic information is contained in the probabilities $p_i := \mathbb{P}\{Z = z_i\}$. In that case $\mathbb{E}[Z] = \sum_i p_i z_i$.

5.2 Expectation functions

Consider now the expectation optimization problem (1.14) with $X := \mathbb{R}^n$. For a given x we can view $F(x) = F(x, \omega)$ as a random variable. We assume that the expectation function

$$f(x) = \mathbb{E}[F(x, \omega)]$$

is *well defined*, i.e., for every[18] $x \in \mathbb{R}^n$ the function $F(x, \cdot)$ is measurable, and either $\mathbb{E}[F(x)_+] < +\infty$ or $\mathbb{E}[(-F(x))_+] < +\infty$. The (effective) feasible set of the problem (1.4) is given by $X \cap (\operatorname{dom} f)$, where

$$\operatorname{dom} f := \{x \in \mathbb{R}^n : f(x) < +\infty\}$$

denotes the *domain* of f. It is said that f is *proper* if $f(x) > -\infty$ for all $x \in \mathbb{R}^n$ and $\operatorname{dom} f \neq \emptyset$.

From the theoretical point of view it is convenient to incorporate the constraints "$x \in X$" into the objective function. That is, for any $\omega \in \Omega$ define

$$\overline{F}(x, \omega) := \begin{cases} F(x, \omega), & \text{if } x \in X, \\ +\infty, & \text{if } x \notin X. \end{cases}$$

Then problem (1.4) can be written in the form

$$\operatorname*{Min}_{x \in \mathbb{R}^n} \mathbb{E}[\overline{F}(x, \omega)]. \tag{5.2}$$

[18] Since we are interested here in x belonging to the feasible set X, we can assume that $f(x)$ is well defined for $x \in X$.

Clearly, the domain of the expectation function $\mathbb{E}[\overline{F}(\cdot,\omega)]$ is $X \cap (\mathrm{dom} f)$, i.e., it coincides with the feasible set of problem (1.14). In the remainder of this section we assume that the objective function $F(x,\omega)$ is extended real valued and that the corresponding constraints are already absorbed into the objective function.

For $\varepsilon \geq 0$ we say that $x^* \in \mathcal{X}$ is an ε-optimal solution of the problem of minimization of $f(x)$ over $\mathcal{X}$ if

$$f(x^*) \leq \inf_{x \in \mathcal{X}} f(x) + \varepsilon.$$

If the problem is infeasible (that is, $f(x) = +\infty$ for every $x \in \mathcal{X}$), then any $x^* \in \mathcal{X}$ is ε-optimal. If the problem is feasible, and hence $\inf_{x \in \mathcal{X}} f(x) < +\infty$, then ε-optimality of x^* implies that $f(x^*) < +\infty$, i.e., that $x^* \in \mathrm{dom} f$. Note that by the nature of the minimization process, if $\inf_{x \in \mathcal{X}} f(x) > -\infty$, then for any $\varepsilon > 0$ there always exists an ε-optimal solution.

An extended real valued function $f: \mathbb{R}^n \to \overline{\mathbb{R}}$ is called *lower semicontinuous* (lsc) at a point x_0 if

$$\liminf_{x \longrightarrow x_0} f(x) \geq f(x_0).$$

It is said that f is *lower semicontinuous* if it is lsc at every point $x \in \mathbb{R}^n$. It is not difficult to show that f is lsc iff its epigraph

$$\mathrm{epi} f := \{(x,\alpha): f(x) \leq \alpha\}$$

is a closed subset of $\mathbb{R}^n \times \mathbb{R}$.

Theorem 13. *Let $f: \mathbb{R}^n \to \overline{\mathbb{R}}$ be a proper extended real valued function. Suppose that f is* lsc *and its domain* $\mathrm{dom} f$ *is bounded. Then the set* $\arg\min_{x \in \mathbb{R}^n} f(x)$ *of its optimal solutions is nonempty.*

Since f is proper, its domain is nonempty, and hence $\inf_{x \in \mathbb{R}^n} f(x) < +\infty$. Let us take a number $c > \inf_{x \in \mathbb{R}^n} f(x)$, and consider the level set $S := \{x: f(x) \leq c\}$. We have that the set S is nonempty, is contained in $\mathrm{dom} f$ and hence is bounded, and is closed since f is lsc. Consequently, the set S is compact, and clearly $\mathrm{argmin}_{x \in \mathbb{R}^n} f(x)$ coincides with $\mathrm{argmin}_{x \in S} f(x)$. Therefore, the above theorem states the well known result that a lsc real valued function attains its minimum over a nonempty compact subset of $\mathbb{R}^n$.

The expected value function $f(x) := \mathbb{E}[F(x,\omega)]$ inherits various properties of the functions $F(\cdot,\omega)$. If for P-almost $\omega \in \Omega$ the function $F(\cdot,\omega)$ is convex, then the expected value function $f(\cdot)$ is also convex. Indeed, if Ω is finite, then $f(\cdot)$ is a weighted sum of convex functions with nonnegative weights, and hence is convex. The case of a general distribution can be then proved by passing to a limit.

As it is shown in the next proposition the lower semicontinuity of the expected value function follows from the lower semicontinuity of $F(\cdot, \omega)$.

Proposition 14. *Suppose that:* (i) *for P-almost every* $\omega \in \Omega$ *the function* $F(\cdot, \omega)$ *is* lsc *at* x_0, (ii) $F(x, \cdot)$ *is measurable for every* x *in a neighborhood of* x_0, (iii) *there exists P-integrable function* $Z(\omega)$ *such that* $F(x, \omega) \geq Z(\omega)$ *for P-almost all* $\omega \in \Omega$ *and all* x *in a neighborhood of* x_0. *Then for all* x *in a neighborhood of* x_0 *the expected value function* $f(x) := \mathbb{E}[F(x, \omega)]$ *is well defined and* lsc *at* x_0.

Proof. It follows from assumptions (ii) and (iii) that $f(\cdot)$ is well defined in a neighborhood of x_0. Under assumption (iii), it follows by Fatou's lemma that

$$\liminf_{x \to x_0} \int_{\Omega} F(x, \omega)\, \mathrm{d}P(\omega) \geq \int_{\Omega} \liminf_{x \to x_0} F(x, \omega)\, \mathrm{d}P(\omega). \tag{5.3}$$

Together with (i) this implies lower semicontinuity of f at x_0. □

In particular, let us consider the probabilistic constraints (1.20). We can write these constraints in the form

$$\mathbb{E}\big[1_{(0,+\infty)}(G_i(x, \omega))\big] \leq \alpha, \quad i = 1, \ldots, m. \tag{5.4}$$

Suppose further that for P-almost every $\omega \in \Omega$ the functions $G_i(\cdot, \omega)$ are lsc, and for all x, $G_i(x, \cdot)$ are measurable. Then functions $\mathbf{1}_{(0,+\infty)}(G_i(\cdot, \omega))$ are also lsc for P-almost every $\omega \in \Omega$, and clearly are bounded. Consequently we obtain by Proposition 14 that the corresponding expected value functions in the left hand side of (5.4) are lsc. It follows that constraints (5.4), and hence the probabilistic constraints (1.20), define a closed subset of $\mathbb{R}^n$.

5.3 *Optimal values and optimal solutions*

We often have to deal with optimal value functions of min or max types. That is, consider an extended real valued function $h\colon \mathbb{R}^n \times \mathbb{R}^m \to \overline{\mathbb{R}}$ and the associated functions

$$\phi(x) := \inf_{y \in \mathbb{R}^m} h(x, y) \quad \text{and} \quad \psi(x) := \sup_{y \in \mathbb{R}^m} h(x, y). \tag{5.5}$$

Proposition 15. *The following holds.* (i) *Suppose that for every* $y \in \mathbb{R}^m$ *the function* $h(\cdot, y)$ *is* lsc. *Then the max-function* $\psi(x)$ *is* lsc. (ii) *Suppose that the function* $h(\cdot, \cdot)$ *is* lsc *and there exists a bounded set* $S \subset \mathbb{R}^m$ *such that* $\operatorname{dom} h(x, \cdot) \subset S$ *for all* $x \in \mathbb{R}^n$. *Then the min-function* $\phi(x)$ *is* lsc.

Proof. (i) The epigraph of the max-function $\psi(\cdot)$ is given by the intersection of the epigraphs of $h(\cdot, y)$, $y \in \mathbb{R}^m$. By lower semicontinuity of $h(\cdot, y)$, these epigraphs are closed, and hence their intersection is closed. It follows that $\psi(\cdot)$ is lsc.

(ii) Consider a point $x_0 \in \mathbb{R}^n$ and let $\{x_k\}$ be a sequence converging to x_0 along which the $\liminf_{x \to x_0} \phi(x)$ is attained. If $\lim_{k\to\infty} \phi(x_k) = +\infty$, then clearly $\lim_{k\to\infty} \phi(x_k) \geq \phi(x_0)$, and hence ϕ is lsc at x_0. Therefore, we can assume that $\phi(x_k) < +\infty$ for all k. Let ε be a given positive number and $y_k \in R^m$ be such that $h(x_k, y_k) \leq \phi(x_k) + \varepsilon$. Since $y_k \in \operatorname{dom} h(x_k, \cdot) \subset S$ and S is bounded, by passing to a subsequence if necessary we can assume that y_k converges to a point y_0. By lower semicontinuity of $h(\cdot, \cdot)$ we have then that $\lim_{k\to\infty} \phi(x_k) \geq h(x_0, y_0) - \varepsilon \geq \phi(x_0) - \varepsilon$. Since ε was arbitrary, it follows that $\lim_{k\to\infty} \phi(x_k) \geq \phi(x_0)$, and hence $\phi(\cdot)$ is lsc at x_0. This completes the proof. □

Let $F: \mathbb{R}^n \times \Omega \to \overline{\mathbb{R}}$ and let us now consider the optimal value

$$\vartheta(\omega) := \inf_{x \in \mathcal{X}} F(x, \omega) \tag{5.6}$$

and the corresponding set

$$X^*(\omega) := \arg \min_{x \in \mathcal{X}} F(x, \omega) \tag{5.7}$$

of optimal solutions. In order to deal with measurability of these objects we need the following concepts.

Let $\mathcal{G}$ be a mapping from Ω into the set of subsets of $\mathbb{R}^n$, i.e., $\mathcal{G}$ assigns to each $\omega \in \Omega$ a subset (possibly empty) $\mathcal{G}(\omega)$ of $\mathbb{R}^n$. We refer to $\mathcal{G}$ as a *multifunction* and write $\mathcal{G}: \Omega \rightrightarrows \mathbb{R}^n$. It is said that $\mathcal{G}$ is *closed valued* if $\mathcal{G}(\omega)$ is a closed subset of $\mathbb{R}^n$ for every $\omega \in \Omega$. A closed valued multifunction $\mathcal{G}$ is said to be *measurable*, if for every closed set $A \subset \mathbb{R}^n$ one has that the inverse image $\mathcal{G}^{-1}(A) := \{\omega \in \Omega: \mathcal{G}(\omega) \cap A \neq \phi\}$ is $\mathcal{F}$-measurable. Note that measurability of $\mathcal{G}$ implies that the domain

$$\operatorname{dom} \mathcal{G} := \{\omega \in \Omega: \mathcal{G}(\omega) \neq \emptyset\} = \mathcal{G}^{-1}(\mathbb{R}^n)$$

of $\mathcal{G}$ is an $\mathcal{F}$-measurable subset of Ω.

It is said that a mapping $G: \operatorname{dom} \mathcal{G} \to \mathbb{R}^n$ is a selection of $\mathcal{G}$ if $G(\omega) \in \mathcal{G}(\omega)$ for all $\omega \in \operatorname{dom} \mathcal{G}$. If, in addition, the mapping G is measurable, it is said that G is a measurable selection of $\mathcal{G}$.

Theorem 16 (Castaing Representation theorem). *A closed valued multifunction $\mathcal{G}: \Omega \rightrightarrows \mathbb{R}^n$ is measurable iff its domain is an $\mathcal{F}$-measurable subset of Ω and there exists a countable family $\{G_i\}_{i\in\mathbb{N}}$, of measurable selections of $\mathcal{G}$ such that for every $\omega \in \Omega$, the set $\{G_i(\omega): i \in \mathbb{N}\}$ is dense in $\mathcal{G}(\omega)$.*

It follows from the above theorem that if $\mathcal{G}: \Omega \rightrightarrows \mathbb{R}^n$ is a closed valued measurable multifunction, then there exists at least one measurable selection of $\mathcal{G}$.

Definition 17. It is said that the function $(x, \omega) \mapsto F(x, \omega)$ is random lower semicontinuous if the associated epigraphical multifunction $\omega \mapsto \operatorname{epi} F(\cdot, \omega)$ is closed valued and measurable.

Note that close valuedness of the epigraphical multifunction means that for every $\omega \in \Omega$, the epigraph $\operatorname{epi} F(\cdot, \omega)$ is a closed subset of $\mathbb{R}^n \times \mathbb{R}$, i.e., that $F(\cdot, \omega)$ is lsc.

Theorem 18. *Suppose that the sigma algebra $\mathcal{F}$ is P-complete. Then an extended real valued function $F: \mathbb{R}^n \times \Omega \to \overline{\mathbb{R}}$ is random lsc iff the following two properties hold:* (i) *for every $\omega \in \Omega$, the function $F(\cdot, \omega)$ is lsc,* (ii) *the function $F(\cdot, \cdot)$ is measurable with respect to the sigma algebra of $\mathbb{R}^n \times \Omega$ given by the product of the sigma algebras $\mathcal{B}$ and $\mathcal{F}$.*

A large class of random lower semicontinuous functions is given by the so-called *Carathéodory* functions, i.e., real valued functions $F: \mathbb{R}^n \times \Omega \to \mathbb{R}$ such that $F(x, \cdot)$ is $\mathcal{F}$-measurable for every $x \in \mathbb{R}^n$ and $F(\cdot, \omega)$ continuous for a.e. $\omega \in \Omega$.

Theorem 19. *Let $F: \mathbb{R}^n \times \Omega \to \overline{\mathbb{R}}$ be a random* lsc *function. Then the optimal value function $\vartheta(\omega)$ and the optimal solution multifunction $X^*(\omega)$ are both measurable.*

Note that it follows from lower semicontinuity of $F(\cdot, \omega)$ that the optimal solution multifunction $X^*(\omega)$ is closed valued. Note also that if $F(x, \omega)$ is random lsc and $\mathcal{G}: \Omega \rightrightarrows \mathbb{R}^n$ is a closed valued measurable multifunction, then the function

$$\overline{F}(x, \omega) := \begin{cases} F(x, \omega), & \text{if } x \in \mathcal{G}(\omega), \\ +\infty, & \text{if } x \notin \mathcal{G}(\omega), \end{cases}$$

is also random lsc. Consequently the corresponding optimal value $\omega \mapsto \inf_{x \in \mathcal{G}(\omega)} F(x, \omega)$ and the optimal solution multifunction $\omega \mapsto \operatorname{argmin}_{x \in \mathcal{G}(\omega)} F(x, \omega)$ are both measurable, and hence by the measurable selection theorem, there exists a measurable selection $\overline{x}(\omega) \in \operatorname{argmin}_{x \in \mathcal{G}(\omega)} F(x, \omega)$.

Theorem 20. *Let $F: \mathbb{R}^{n+m} \times \Omega \to \overline{\mathbb{R}}$ be a random* lsc *function and*

$$\vartheta(x, \omega) := \inf_{y \in \mathbb{R}^m} F(x, y, \omega) \tag{5.8}$$

be the associated optimal value function. Suppose that there exists a bounded set $S \subset \mathbb{R}^m$ such that $\operatorname{dom} F(x, \cdot, \omega) \subset S$ for all $(x, \omega) \in \mathbb{R}^n \times \Omega$. Then the optimal value function $\vartheta(x, \omega)$ is random lsc.

Let us finally observe that the above framework of random lsc functions is aimed at minimization problems. Of course, the problem of maximization of $\mathbb{E}[F(x,\omega)]$ is equivalent to minimization of $\mathbb{E}[-F(x,\omega)]$. Therefore, for maximization problems one would need the comparable concept of random upper semicontinuous functions.

6 Bibliographic notes

Stochastic programming with recourse originated in the works of Beale (1955) and Dantzig (1955). Basic properties of two-stage problems were investigated by Wets (1966), Walkup and Wets (1967, 1969) and Kall (1976). A comprehensive treatment of the theory and numerical methods for expectation models can be found in Birge and Louveaux (1997). Simulation-based approaches to stochastic optimization were discussed by various authors, see Chapter "Monte Carlo Sampling Methods".

Models involving constraints on probability were introduced by Charnes et al. (1958), Miller and Wagner (1965), and Prékopa (1970). Prékopa (1995) discusses in detail the theory and numerical methods for linear chance-constrained models. Applications to finance are discussed by Dowd (1997). Klein Haneveld (1986) introduced the concept of *integrated chance constraints*, which are the predecessors of *conditional value at risk* constraints of Uryasev and Rockafellar (2001).

A general discussion of interchangeability of minimization and integration operations can be found in Rockafellar and Wets (1998). Proposition 5 is a particular case of Theorem 14.60 in Rockafellar and Wets (1998).

Expected value of perfect information is a classical concept in decision theory (see, e.g., Raiffa (1968)). In stochastic programming this and related concepts were analyzed first by Madansky (1960). Other advances are due to Dempster (1981) and Birge (1982).

Early contributions to multistage stochastic programming models appeared in Marti (1975), Beale et al. (1980), Louveaux (1980), Birge (1985), Noël and Smeers (1986) and Dempster (1981). Varaiya and Wets (1989) discuss relations of multistage stochastic programming and stochastic control. For other examples and approaches to multistage modeling see Birge and Louveaux (1997).

Robust approaches to stochastic programming were initiated by Mulvey et al. (1995). Proposition 11 is based on the work of Takriti and Ahmed (2002). Mean–risk models in portfolio optimization were introduced by Markowitz (1952). For a general perspective, see Markowitz (1987) and Luenberger (1998). Mean–absolute deviation models for portfolio problems were introduced by Konno and Yamazaki (1991). Semideviations and other risk measures were analyzed by Ogryczak and Ruszczyński (1999, 2001, 2002). Min–max approach to stochastic programming was initiated in Žáčková (1966), Dupačová (1980), Dupačová (1987).

There are many good textbooks on probability and measure theory, e.g., Billingsley (1995), to which we refer for a thorough discussion of such basic concepts as random variables, probability space, etc. Also a proof of Fatou's lemma, used in the proof of Proposition 14, can be found there. For an additional discussion of the expected value function see section "Expectation Functions" of Chapter 2. Continuity and differentiability properties of the optimal value functions, of the form defined in equation (5.5), were studied extensively in the optimization literature (see, e.g., Bonnans and Shapiro (2000) and the references therein).

Measurable selection theorem (Theorem 16) is due to Castaing. A thorough discussion of measurable mappings and selections can be found in Castaing and Valadier (1977), Ioffe and Tihomirov (1979) and Rockafellar and Wets (1998). Random lower semicontinuous functions are called *normal integrands* (see Definition 14.27 in Rockafellar and Wets (1998)) by some authors. Proofs of theorems 18, 19 and 20 can be found in the section on normal integrands of Rockafellar and Wets (1998).

References

Beale, E.M.L. (1955). On minimizing a convex function subject to linear inequalities. *Journal of the Royal Statistical Society Series B* 17, 173–184.

Beale, E.M.L., J.J.H. Forrest, C.J. Taylor (1980). Multi-time-period stochastic programming, in: M.A.H. Dempster (ed.), *Stochastic Programming*, Academic Press, New York, pp. 387–402.

Billingsley, P. (1995). *Probability and Measure*, John Wiley & Sons, New York.

Birge, J.R. (1982). The value of the stochastic solution in stochastic linear programs with fixed recourse. *Mathematical Programming* 33, 314–325.

Birge, J.R. (1985). Decomposition and partitioning methods for multistage stochastic linear programs. *Operations Research*, 989–1007.

Birge, J.R., F.V. Louveaux (1997). *Introduction to Stochastic Programming*, Springer-Verlag, New York.

Bonnans, J.F., A. Shapiro (2000). *Perturbation Analysis of Optimization Problems*, Springer-Verlag, New York, NY.

Castaing, C., M. Valadier (1977). *Convex Analysis and Measurable Multifunctions*, Lecture Notes in Mathematics, Vol. 580, Springer-Verlag, Berlin.

Charnes, A., W.W. Cooper, G.H. Symonds (1958). Cost horizons and certainty equivalents: an approach to stochastic programming of heating oil. *Management Science* 4, 235–263.

Dantzig, G.B. (1955). Linear programming under uncertainty. *Management Science* 1, 197–206.

Dempster, M.A.H. (1981). The expected value of perfect information in the optimal evolution of stochastic systems, in: M. Arato, D. Vermes, A.V. Balakrishnan (eds.), *Stochastic Differential Systems, Lecture Notes in Control and Information Systems*, Vol. 36. Springer-Verlag, Berlin, pp. 25–41.

Dowd, K. (1997). *Beyond Value at Risk. The Science of Risk Management*, Wiley, New York.

Dupačová, J. (1980). On minimax decision rule in stochastic linear programming *Studies on Mathematical Programming*, A. Prékopa Akadémiai Kiadó, Budapest, pp. 47–60.

Dupačová, J. (1987). The minimax approach to stochastic programming and an illustrative application. *Stochastics* 20, 73–88.

Ioffe, A.D., V.M. Tihomirov (1979). *Theory of Extremal Problems*, North-Holland Publishing Company, Amsterdam.

Kall, P. (1976). *Stochastic Linear Programming*, Springer-Verlag, Berlin.

Klein Haneveld, W.K. (1986). *Duality in Stochastic Linear and Dynamic Programming, Lecture Notes in Economic and Mathematical Systems*, Vol. 274, Springer-Verlag, New York.

Konno, H., H. Yamazaki (1991). Mean–absolute deviation portfolio optimization model and its application to Tokyo stock market. *Management Science* 37, 519–531.

Louveaux, F.V. (1980). A solution method for multistage stochastic programs with recourse with applications to an energy investment problem. *Operations Research* 28, 889–902.

Luenberger, D.G. (1998). *Investment Science*, Oxford University Press, New York.

Madansky, A. (1960). Inequalities for stochastic linear programming problems. *Management Science* 6, 197–204.

Markowitz, H.M. (1952). Portfolio selection. *Journal of Finance* 7, 77–91.

Markowitz, H.M. (1987). *Mean–Variance Analysis in Portfolio Choice and Capital Markets*, Blackwell, Oxford.

Marti, K. (1975). Über zwei- und mehrstufige stochastische Kontrollprobleme, in: Vorträge der Wissenschaftlichen Jahrestagung der Gesellschaft für Angewandte Mathematik und Mechanik (Bochum, 1974). *Zeitschrift für Angewandte Mathematik und Mechanik* 55, T281–T282.

Miller, L.B., H. Wagner (1965). Chance-constrained programming with joint constraints. *Operations Research* 13, 930–945.

Mulvey, J.M., R.J. Vanderbei, S.A. Zenios (1995). Robust optimization of large-scale systems. *Operations Research* 42, 264–281.

Noël, M.-C., Y. Smeers (1986). On the use of nested decomposition for solving nonlinear multistage stochastic programs, in: *Stochastic Programming (Gargnano, 1983)*, *Lecture Notes in Control and Inform. Sci.*, Vol. 76, Springer-Verlag, Berlin, pp. 235–246.

Ogryczak, W., A. Ruszczynski (1999). From stochastic dominance to mean–risk models: semideviations as risk measures. *European Journal of Operational Research* 116, 33–50.

Ogryczak, W., A. Ruszczyński (2001). On consistency of stochastic dominance and mean–semideviation models. *Mathematical Programming* 89, 217–232.

Ogryczak, W., A. Ruszczynski (2002). Dual stochastic dominance and related mean–risk models. *SIAM Journal on Optimization* 13, 60–78.

Pflug, G.Ch. (1996). *Optimization of Stochastic Models. The Interface Between Simulation and Optimization*, Kluwer Academic Publishers, Boston, MA.

Prékopa, A. (1970). On probabilistic constrained programming: *Proceedings of the Princeton Symposium on Mathematical Programming*, Princeton University Press, Princeton, pp. 113–138.

Prékopa, A. (1995). *Stochastic Programming*, Kluwer, Dordrecht, Boston.

Raiffa, H. (1968). *Decision Analysis*, Addison–Wesley, Reading, MA.

Rockafellar, R.T., R.J.-B. Wets (1998). *Variational Analysis*, Springer-Verlag, Berlin.

Takriti, S., S. Ahmed (2002). On robust optimization of two-stage systems. *Mathematical Programming*. To appear.

Uryasev, S., R.T. Rockafellar (2001). Conditional value-at-risk: optimization approach). *Stochastic optimization: algorithms and applications* (Gainesville, FL, 2000), *Applied Optimization*, Vol. 54, Kluwer Academic Publishers, Dordrecht, pp. 411–435.

Varaiya, P., R.J.-B. Wets (1989). Stochastic dynamic optimization approaches and computation, in: M. Iri, K. Tanabe (eds.), *Mathematical Programming: Recent Developments and Applications,* Kluwer, Dordrecht, pp. 309–332.

Walkup, D., R.J.-B. Wets (1967). Stochastic programs with recourse. *SIAM Journal on Applied Mathematics* 15, 1299–1314.

Walkup, D., R.J.-B. Wets (1969). Stochastic programs with recourse II: on the continuity of the objective. *SIAM Journal on Applied Mathematics* 15, 1299–1314.

Wets, R.J.-B. (1966). Programming under uncertainty: the equivalent convex program. *SIAM Journal on Applied Mathematics* 14, 89–105.

Žácková, J. (1966). On minimax solutions of stochastic linear programming problems. *Čas. Pest. Mat.* 91, 423–430.

A. Ruszczyński and A. Shapiro, Eds., *Handbooks in OR & MS, Vol. 10*

Chapter 2

Optimality and Duality in Stochastic Programming

Andrzej Ruszczyński

Department of Management Science and Information Systems, Rutgers University, 94 Rockafeller Rd, Piscataway, NJ 08854, USA

Alexander Shapiro

School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA

Abstract

In this chapter we discuss basic mathematical properties of convex stochastic programming models. We develop expressions for the subdifferentials of the objective function in two- and multi-stage models. Then we present necessary and sufficient conditions of optimality, and duality relations for these problems.

Key words: Expected value function, two stage stochastic programming, multistage stochastic programming, optimality conditions, duality.

1 Expectation functions

In this section we discuss general properties of the *expectation* (also called *expected value*) functions of the form

$$f(x) := \mathbb{E}_P[F(x, \omega)]. \tag{1.1}$$

Here P is a probability measure defined on a measurable space $(\Omega, \mathcal{F})$ and $F: \mathbb{R}^n \times \Omega \to \overline{\mathbb{R}}$ is an extended real valued function, called an *integrand*. The function $f(x)$ is *well defined* on a set $X \subset \mathbb{R}^n$ if for every $x \in X$ the function $F(x, \cdot)$ is measurable, and either $\mathbb{E}[F(x, \omega)_+] < +\infty$ or $\mathbb{E}[(-F(x, \omega))_+] < +\infty$ (see Chapter 1, Appendix).

The expected value function $f(x)$ inherits various properties of the integrand $F(x, \omega)$. We already gave a preliminary discussion of that in Section 5.2 of

Chapter 1. In particular, it was shown in Proposition 14 that as a consequence of Fatou's lemma, $f(\cdot)$ inherits lower semicontinuity of $F(\cdot, \omega)$ under the condition that, locally in x, $F(x, \cdot)$ is bounded from below by a P-integrable function. The following proposition gives similar conditions for the continuity of $f(x)$ at a point $x_0 \in \mathbb{R}^n$.

Proposition 1. *Suppose that:* (i) *for P-almost every* $\omega \in \Omega$ *the function* $F(\cdot, \omega)$ *is continuous at* x_0, (ii) $F(x, \cdot)$ *is measurable for every* x *in a neighborhood of* x_0, (iii) *there exists P-integrable function* $Z(\omega)$ *such that* $|F(x, \omega)| \leq Z(\omega)$ *for P-almost every* $\omega \in \Omega$ *and all* x *in a neighborhood of* x_0. *Then for all* x *in a neighborhood of* x_0 *the expected value function* $f(x)$ *is well defined and continuous at* x_0.

Proof. It follows from assumptions (ii) and (iii) that $f(x)$ is well defined for all x in a neighborhood of x_0. Moreover, by the Lebesgue Dominated Convergence Theorem we can take the limit inside the integral, which together with (i) implies

$$\lim_{x \longrightarrow x_0} \int_\Omega F(x, \omega)\, \mathrm{d}P(\omega) = \int_\Omega \lim_{x \longrightarrow x_0} F(x, \omega)\, \mathrm{d}P(\omega) = \int_\Omega F(x_0, \omega)\, \mathrm{d}P(\omega). \tag{1.2}$$

This shows the continuity of $f(x)$ at x_0. □

Consider, for example, the characteristic function $F(x, \omega) := \mathbf{1}_{(-\infty, x]}(\xi(\omega))$, with $x \in \mathbb{R}$ and $\xi = \xi(\omega)$ being a real valued random variable. We have then that $f(x) = P(\xi \leq x)$, i.e., that $f(\cdot)$ is the cumulative distribution function of ξ. It follows that in this example the expected value function is continuous at a point x_0 iff the the probability of the event $\{\xi = x_0\}$ is zero. Note that $x = \xi(\omega)$ is the only point at which the function $F(\cdot, \omega)$ is discontinuous.

We discuss now the differentiability properties of the expected value function $f(x)$ defined in (1.1). We sometimes write $F_\omega(\cdot)$ for the function $F(\cdot, \omega)$ and denote by $F'_\omega(x_0, h)$ the directional derivative of $F_\omega(\cdot)$ at the point x_0 in the direction h. Definitions and basic properties of directional derivatives are given in Section 9.1 of the Appendix.

Proposition 2. *Suppose that:* (i) $F(x, \cdot)$ *is measurable for all* x *in a neighborhood of* x_0, (ii) $\mathbb{E}|F(x_0, \omega)| < +\infty$, (iii) *there exists a positive valued random variable* $C(\omega)$ *such that* $\mathbb{E}[C(\omega)] < +\infty$, *and for all* x_1, x_2 *in a neighborhood of* x_0 *and almost every* $\omega \in \Omega$ *the following inequality holds*

$$|F(x_1, \omega) - F(x_2, \omega)| \leq C(\omega)\|x_1 - x_2\|, \tag{1.3}$$

(iv) *for almost every* ω *the function* $F_\omega(\cdot)$ *is directionally differentiable at* x_0. *Then the expected value function* $f(x)$ *is Lipschitz continuous in a neighborhood of* x_0, *directionally differentiable at* x_0, *and*

$$f'(x_0, h) = \mathbb{E}\big[F'_\omega(x_0, h)\big], \quad \textit{for all } h. \tag{1.4}$$

If, in addition, the function $F(\cdot, \omega)$ *is differentiable at* x_0 *w.p.1, then* $f(x)$ *is differentiable at* $f(x)$ *and*

$$\nabla f(x_0) = \mathbb{E}[\nabla_x F(x_0, \omega)]. \tag{1.5}$$

Proof. It follows from (1.3) that for any x_1, x_2 in a neighborhood of x_0,

$$|f(x_1) - f(x_2)| \leq \int_\Omega |F(x_1, \omega) - F(x_2, \omega)| \, \mathrm{d}P(\omega) \leq C\|x_1 - x_2\|,$$

where $C := \mathbb{E}[C(\omega)]$. Since $f(x_0)$ is finite by assumption (ii), it follows that $f(x)$ is well defined, finite valued and Lipschitz continuous in a neighborhood of x_0. For $t \neq 0$ consider the ratio

$$R_t(\omega) := t^{-1}[F(x_0 + th, \omega) - F(x_0, \omega)].$$

By assumption (iii) we have that $|R_t(\omega)| \leq C(\omega)\|h\|$, and by assumption (iv) that

$$\lim_{t \downarrow 0} R_t(\omega) = F'_\omega(x_0, h) \quad \text{w.p.1.}$$

Therefore, it follows from the Lebesgue Dominated Convergence Theorem that

$$\lim_{t \downarrow 0} \int_\Omega R_t(\omega) \, \mathrm{d}P(\omega) = \int_\Omega \lim_{t \downarrow 0} R_t(\omega) \, \mathrm{d}P(\omega).$$

Together with assumption (iv) this implies formula (1.4).

Finally, if $F'_\omega(x_0, h)$ is linear in h for almost every ω, i.e., the function $F_\omega(\cdot)$ is differentiable at x_0 w.p.1, then (1.4) implies that $f'(x_0, h)$ is linear in h, and hence (1.5) follows. Note that since $f(x)$ is locally Lipschitz continuous, we only need to verify linearity of $f'(x_0, \cdot)$ in order to establish (Fréchet) differentiability of $f(x)$ at x_0. □

The above analysis shows that two basic conditions for interchangeability of the expectation and differentiation operators, i.e., for the validity of

formula (1.5), are: (a) the local Lipschitz continuity of the random function $F(\cdot, \omega)$; and (b) the differentiability of $F(\cdot, \omega)$ at the given point x_0 w.p.1. The following lemma shows that if, in addition to the assumptions of the above proposition, the directional derivative $F'_\omega(x_0, h)$ is *convex* in h w.p.1 (i.e., for almost every ω), then $f(x)$ is differentiable at x_0 if and only if $F(\cdot, \omega)$ is differentiable at x_0 w.p.1.

Lemma 3. *Let $\psi\colon \mathbb{R}^n \times \Omega \to \mathbb{R}$ be a function such that for almost every $\omega \in \Omega$ the function $\psi(\cdot, \omega)$ is convex and positively homogeneous, and the expected value function $\phi(h) := \mathbb{E}[\psi(h, \omega)]$ is well defined and finite valued. Then the expected value function $\phi(\cdot)$ is linear if and only if the function $\psi(\cdot, \omega)$ is linear w.p.1.*

Proof. We have here that the expected value function $\phi(\cdot)$ is convex and positively homogeneous. Moreover, it immediately follows from the linearity properties of the expectation operator that if the function $\psi(\cdot, \omega)$ is linear w.p.1, then $\phi(\cdot)$ is also linear.

Conversely, let $e_1, \ldots, e_n$ be a basis of the space $\mathbb{R}^n$. Since $\phi(\cdot)$ is convex and positively homogeneous, it follows that $\phi(e_i) + \phi(-e_i) \geq \phi(0) = 0$, $i = 1, \ldots, n$. Furthermore, since $\phi(\cdot)$ is finite valued, it is the support function of a convex compact set. This convex set is a singleton iff

$$\phi(e_i) + \phi(-e_i) = 0, \quad i = 1, \ldots, n. \tag{1.6}$$

Therefore, $\phi(\cdot)$ is linear iff condition (1.6) holds. Consider the sets

$$A_i := \{\omega \in \Omega\colon \psi(e_i, \omega) + \psi(-e_i, \omega) > 0\}.$$

Thus the set of $\omega \in \Omega$ such that $\psi(\cdot, \omega)$ is not linear coincides with the set $\cup_{i=1}^n A_i$. If $P\big(\cup_{i=1}^n A_i\big) > 0$, then at least one of the sets A_i has a positive measure. Let, for example, $P(A_1)$ be positive. Then $\phi(e_1) + \phi(-e_1) > 0$, and hence $\phi(\cdot)$ is not linear. This completes the proof. □

Regularity conditions which are required for formula (1.4) to hold are simplified further if the integrand $F(x, \omega)$ is *convex*, i.e., the function $F(\cdot, \omega)$ is convex for almost every $\omega \in \Omega$. In that case, by using the Monotone Convergence Theorem instead of the Lebesgue Dominated Convergence Theorem, it is possible to prove the following result.

Proposition 4. *Suppose that:* (i) *the expected value function $f(x)$ is well defined and finite valued in a neighborhood of a point x_0,* (ii) *for almost every $\omega \in \Omega$ the function $F_\omega(\cdot) := F(\cdot, \omega)$ is convex. Then $f(x)$ is convex, directionally differentiable at x_0 and formula (1.4) holds. Moreover, $f(x)$ is differentiable at x_0 if and only if $F_\omega(x)$ is differentiable at x_0 w.p.1, in which case formula (1.5) holds.*

Proof. The convexity of $f(x)$ follows immediately from the convexity of $F_\omega(\cdot)$ (we already discussed that in Chapter 1). Consider a direction $h \in \mathbb{R}^n$. By assumption (i) we have that $f(x_0)$ and, for some $t_0 > 0$, $f(x_0 + t_0 h)$ are finite. It follows from the convexity of $F_\omega(\cdot)$ that the ratio

$$R_t(\omega) := t^{-1}[F(x_0 + th, \omega) - F(x_0, \omega)]$$

is monotonically decreasing to $F'_\omega(x_0, h)$ as $t \downarrow 0$. Also we have that

$$\mathbb{E}|R_{t_0}(\omega)| \leq t_0^{-1}(\mathbb{E}|F(x_0 + t_0 h, \omega)| + \mathbb{E}|F(x_0, \omega)|) < +\infty.$$

Then it follows by the Monotone Convergence Theorem that

$$\lim_{t \downarrow 0} \mathbb{E}[R_t(\omega)] = \mathbb{E}\left[\lim_{t \downarrow 0} R_t(\omega)\right] = \mathbb{E}\big[F'_\omega(x_0, h)\big].$$

This proves formula (1.4). The last assertion follows then from Lemma 3. □

Remark 5. It is possible to give a version of the above result for a particular direction $h \in \mathbb{R}^n$. That is, suppose that: (i) the expected value function $f(x)$ is well defined in a neighborhood of a point x_0, (ii) $f(x_0)$ is finite, (iii) for almost every $\omega \in \Omega$ the function $F_\omega(\cdot) := F(\cdot, \omega)$ is convex, (iv) $\mathbb{E}[F(x_0 + t_0 h, \omega)] < +\infty$ for some $t_0 > 0$. Then $f'(x_0, h) < +\infty$ and formula (1.4) holds. Note also that if the above assumptions (i)–(iii) are satisfied and $\mathbb{E}[F(x_0 + th, \omega)] = +\infty$ for any $t > 0$, then clearly $f'(x_0, h) = +\infty$.

Often the expectation operator smoothes the integrand $F(x, \omega)$. Consider, for example, $F(x, \omega) := |x - \xi(\omega)|$ with $x \in \mathbb{R}$ and $\xi(\omega)$ being a real valued random variable. Suppose that $f(x) = \mathbb{E}[F(x, \omega)]$ is finite valued. We have here that $F(\cdot, \omega)$ is convex and $F(\cdot, \omega)$ is differentiable everywhere except $x = \xi(\omega)$. The corresponding derivative is given by $\partial F(x, \omega)/\partial x = 1$ if $x > \xi(\omega)$ and $\partial F(x, \omega)/\partial x = -1$ if $x < \xi(\omega)$. Therefore, $f(x)$ is differentiable at x_0 iff the event $\{\xi(\omega) = x_0\}$ has zero probability, in which case $df(x_0)/dx = \mathbb{E}[\partial F(x_0, \omega)/\partial x]$. If the event $\{\xi(\omega) = x_0\}$ has positive probability, then the directional derivatives $f'(x_0, h)$ exist but are not linear in h, that is,

$$f'(x_0, -1) + f'(x_0, 1) = 2P(\xi(\omega) = x_0) > 0.$$

We can also investigate differentiability properties of the expectation function by studying the subdifferentiability of the integrand. Suppose for the moment that the set Ω is finite, say $\Omega := \{\omega_1, \ldots \omega_K\}$ with $P\{\omega = \omega_k\} = p_k > 0$, and that the functions $F(\cdot, \omega)$, $\omega \in \Omega$, are proper. Then $f(x) = \sum_{k=1}^K p_k F(x, \omega_k)$ and $\operatorname{dom} f = \cap_{k=1}^K \operatorname{dom} F_k$, where $F_k(\cdot) := F(\cdot, \omega_k)$. The Moreau–Rockafellar

Theorem (Theorem 50) allows us to express the subdifferential of $f(x)$ as the sum of subdifferentials of $p_k F(x, \omega_k)$.

Theorem 6. *Suppose that:* (i) *the set* $\Omega = \{\omega_1, \ldots \omega_K\}$ *is finite,* (ii) *for every* $\omega_k \in \Omega$ *the function* $F_k(\cdot) := F(\cdot, \omega_k)$ *is proper and convex,* (iii) *the sets* $\mathrm{ri}(\mathrm{dom}\, F_k)$, $k = 1, \ldots, K$, *have a common point. Then for any* $x_0 \in \mathrm{dom}\, f$,

$$\partial f(x_0) = \sum_{k=1}^{K} p_k \partial F(x_0, \omega_k). \tag{1.7}$$

Note that the regularity assumption (iii) holds, in particular, if the interior of dom f is nonempty.

The subdifferentials at the right hand side of (1.7) are taken with respect to x, and the sum of these subdifferentials is understood to be the set of all points of the form $\sum_{k=1}^{K} p_k G_k$ with G_k being a selection (i.e., an element) of $\partial F(x_0, \omega_k)$. Note that $\partial F(x_0, \omega_k)$, and hence $\partial f(x_0)$, in (1.7) can be unbounded or empty. Suppose that all probabilities p_k are positive. It follows then from (1.7) that $\partial f(x_0)$ is a singleton iff all subdifferentials $\partial F(x_0, \omega_k)$, $k = 1, \ldots, K$, are singletons. That is, $f(\cdot)$ is differentiable at a point $x_0 \in \mathrm{dom}\, f$ iff all $F(\cdot, \omega_k)$ are differentiable at x_0.

In the case of a finite set Ω we did not have to worry about the measurability of the multifunction $\omega \mapsto \partial F(x, \omega)$. Consider now a general case where the measurable space does not need to be finite. Recall that the function $F(x, \omega)$ is said to be *random lower semicontinuous* if the multifunction $\omega \mapsto \mathrm{epi}\, F(\cdot, \omega)$ is closed valued and measurable (see the Appendix of Chapter 1 for a discussion of that concept).

Proposition 7. *Suppose that the function* $F(x, \omega)$ *is random lower semicontinuous and for a.e.* $\omega \in \Omega$ *the function* $F(\cdot, \omega)$ *is convex and proper. Then for any* $x \in \mathbb{R}^n$, *the multifunction* $\omega \mapsto \partial F(x, \omega)$ *is measurable.*

Proof. Consider the conjugate

$$F^*(z, \omega) := \sup_{x \in \mathbb{R}^n} \{z^T x - F(x, \omega)\}$$

of the function $F(\cdot, \omega)$. It is possible to show that the function $F^*(z, \omega)$ is also random lower semicontinuous. Moreover, by the Fenchel–Moreau Theorem, $F^{**} = F$ and by convex analysis (see (9.15))

$$\partial F(x, \omega) = \arg\max_{z \in \mathbb{R}^n} \{z^T x - F^*(z, \omega)\}.$$

It follows then by Theorem 19 from the Appendix of Chapter 1 that the multifunction $\omega \mapsto \partial F(x, \omega)$ is measurable. $\square$

Definition 8. For a given $x \in \mathbb{R}^n$, the integral $\int_\Omega \partial F(x, \omega)\, \mathrm{d}P(\omega)$ is defined as the set of all points of the form $\int_\Omega G(\omega)\, \mathrm{d}P(\omega)$, where $G(\omega)$ is P-integrable selection of $\partial F(x, \omega)$, i.e., $G(\omega) \in \partial F(x, \omega)$ for a.e. $\omega \in \Omega$, $G(\omega)$ is measurable and $\int_\Omega \|G(\omega)\|\, \mathrm{d}P(\omega)$ is finite.

Of course, if $\Omega = \{\omega_1, \ldots, \omega_K\}$ is finite, then we do not have to worry about the integrability of a selection $G(\omega) \in \partial F(x, \omega)$ and

$$\int_\Omega \partial F(x, \omega)\, \mathrm{d}P(\omega) = \sum_{k=1}^{K} p_k \partial F(x, \omega_k).$$

In general we have the following extension of formula (1.7).

Theorem 9. *Suppose that:* (i) *the function $F(x, \omega)$ is random lower semicontinuous,* (ii) *for a.e. $\omega \in \Omega$ the function $F(\cdot, \omega)$ is convex,* (iii) *the expectation function f is proper,* (iv) *the domain of f has a nonempty interior. Then for any $x_0 \in \mathrm{dom} f$,*

$$\partial f(x_0) = \int_\Omega \partial F(x_0, \omega)\, \mathrm{d}P(\omega) + N_{\mathrm{dom} f}(x_0). \tag{1.8}$$

Proof. Consider a point $z \in \int_\Omega \partial F(x_0, \omega)\, \mathrm{d}P(\omega)$. By the definition of that integral we have then that there exists a P-ntegrable selection $G(\omega) \in \partial F(x_0, \omega)$ such that $z = \int_\Omega G(\omega)\, \mathrm{d}P(\omega)$. Consequently, for a.e. $\omega \in \Omega$ the following holds

$$F(x, \omega) - F(x_0, \omega) \geq G(\omega)^T (x - x_0), \qquad \forall\, x \in \mathbb{R}^n.$$

By taking the integral of the both sides of the above inequality we obtain that z is a subgradient of f at x_0. This shows that

$$\int_\Omega \partial F(x_0, \omega)\, \mathrm{d}P(\omega) \subset \partial f(x_0). \tag{1.9}$$

In particular, it follows from (1.9) that if $\partial f(x_0)$ is empty, then the set at the right hand side of (1.8) is also empty. If $\partial f(x_0)$ is nonempty, i.e., f is subdifferentiable at x_0, then $N_{\mathrm{dom} f}(x_0)$ forms the recession cone of $\partial f(x_0)$. In any case it follows from (1.9) that

$$\int_\Omega \partial F(x_0, \omega)\, \mathrm{d}P(\omega) + N_{\mathrm{dom} f}(x_0) \subset \partial f(x_0). \tag{1.10}$$

Note that inclusion (1.10) holds irrespective of assumption (iv).

Proving the converse of inclusion (1.10) is a more delicate problem. Let us outline main steps of such a proof based on the interchangeability property of the directional derivative and integral operators. We can assume that both sets at the left and right hand sides of (1.9) are nonempty. Since the subdifferentials $\partial F(x_0, \omega)$ are convex, it is quite easy to show that the set $\int_\Omega \partial F(x_0, \omega)\, \mathrm{d}P(\omega)$ is convex. With some additional effort it is possible to show that this set is closed. Let us denote by $s_1(\cdot)$ and $s_2(\cdot)$ the support functions of the sets at the left and right hand sides of (1.10), respectively. By virtue of inclusion (1.9), $N_{\mathrm{dom}f}(x_0)$ forms the recession cone of the set at the left hand side of (1.10) as well. Since the tangent cone $T_{\mathrm{dom}f}(x_0)$ is the polar of $N_{\mathrm{dom}f}(x_0)$, it follows that $s_1(h) = s_2(h) = +\infty$ for any $h \notin T_{\mathrm{dom}f}(x_0)$. Suppose now that (1.8) does not hold, i.e., inclusion (1.10) is strict. Then $s_1(h) < s_2(h)$ for some $h \in T_{\mathrm{dom}f}(x_0)$. Moreover, by assumption (iv), the tangent cone $T_{\mathrm{dom}f}(x_0)$ has a nonempty interior and there exists $\overline{h}$ in the interior of $T_{\mathrm{dom}f}(x_0)$ such that $s_1(\overline{h}) < s_2(\overline{h})$. For such $\overline{h}$ the directional derivative $f'(x_0, h)$ is finite for all h in a neighborhood of $\overline{h}$, $f'(x_0, \overline{h}) = s_2(\overline{h})$ and (see Remark 5)

$$f'(x_0, \overline{h}) = \int_\Omega F'_\omega(x_0, \overline{h})\, \mathrm{d}P(\omega).$$

Also, $F'_\omega(x_0, h)$ is finite for a.e. ω and for all h in a neighborhood of $\overline{h}$, and hence $F'_\omega(x_0, \overline{h}) = \overline{h}^T G(\omega)$ for some $G(\omega) \in \partial F(x_0, \omega)$. Moreover, since the multifunction $\omega \mapsto \partial F(x_0, \omega)$ is measurable, we can choose a measurable $G(\omega)$ here. Consequently,

$$\int_\Omega F'_\omega(x_0, \overline{h})\, \mathrm{d}P(\omega) = \overline{h}^T \int_\Omega G(\omega)\, \mathrm{d}P(\omega).$$

Since $\int_\Omega G(\omega)\, \mathrm{d}P(\omega)$ is a point of the set at the left hand side of (1.9), we obtain that $s_1(\overline{h}) \geq f'(x_0, \overline{h}) = s_2(\overline{h})$, a contradiction. □

2 Two-stage stochastic programming problems

2.1 Linear two-stage problems. Discrete distributions

In this section we discuss two-stage stochastic linear programs of the form

$$\begin{aligned} &\mathrm{Min}_x\ c^T x + \mathbb{E}[Q(x, \xi)] \\ &\text{s.t.}\quad Ax = b,\ x \geq 0, \end{aligned} \tag{2.1}$$

where $Q(x, \xi)$ is the optimal value of the second stage problem

$$\operatorname*{Min}_{y} q^T y \quad \text{subject to} \quad Tx + Wy = h,\ y \geq 0. \tag{2.2}$$

Here $\xi = (q, h, T, W)$ is the data of the problem with some (all) elements of which can be random. The above two-stage problem is the same as problem (2.2)–(2.3) discussed in Section 2 of Chapter 1.

The second stage problem (2.2) is a particular case of the convex problem (2.51) discussed in the following Section 2.4. Therefore, in principle, one can apply results of Section 2.4 to problem (2.2). Note, however, that problem (2.2) has a specific linear structure which allows us to give a more detailed description of its properties. In particular, some regularity conditions can be relaxed.

The second stage problem (2.2) is a linear programming problem. Its dual can be written in the form

$$\operatorname*{Max}_{\pi} \pi^T(h - Tx) \quad \text{subject to} \quad W^T\pi \leq q. \tag{2.3}$$

By the theory of linear programming, the optimal values of problems (2.2) and (2.3) are equal unless both problems are infeasible. Moreover, if their common optimal value is finite, then each problem has a nonempty set of optimal solutions.

Consider the function

$$s_q(\chi) := \inf\{q^T y\colon Wy = \chi,\ y \geq 0\}. \tag{2.4}$$

Clearly $Q(x, \xi) = s_q(h - Tx)$. By the duality theory of linear programming we have that if the set

$$\Pi(q) := \{\pi\colon W^T\pi \leq q\} \tag{2.5}$$

is nonempty, then

$$s_q(\chi) = \sup_{\pi \in \Pi(q)} \pi^T\chi, \tag{2.6}$$

i.e., $s_q(\cdot)$ is the support function of the set $\Pi(q)$. The set $\Pi(q)$ is convex closed and polyhedral, and hence has a finite number of extreme points. It follows that if $\Pi(q)$ is nonempty, then $s_q(\cdot)$ is a positively homogeneous *polyhedral* function (see the following definition of polyhedral functions). If the set $\Pi(q)$ is empty, then the infimum at the right hand side of (2.4) may take only two

values: $+\infty$ or $-\infty$. In any case it is not difficult to verify directly that the function $s_q(\cdot)$ is convex.

Definition 10. An extended real valued function $g: \mathbb{R}^m \to \overline{\mathbb{R}}$ is called *polyhedral* if its epigraph is a convex closed polyhedron, and $g(x)$ is finite for at least one x (which implies that the function g is proper). In other words, $g(\cdot)$ is polyhedral if it is proper convex and lower semicontinuous, its domain is a convex closed polyhedron and $g(\cdot)$ is piecewise linear on its domain.

Proposition 11. *For any given ξ, the function $Q(\cdot, \xi)$ is convex. Moreover, if the set $\{\pi: W^T\pi \leq q\}$ is nonempty and problem* (2.2) *is feasible for at least one x, then $Q(\cdot, \xi)$ is polyhedral.*

Proof. Since $Q(x, \xi) = s_q(h - Tx)$, the above properties of $Q(\cdot, \xi)$ follow from the corresponding properties of $s_q(\cdot)$. □

Proposition 12. *Suppose that for a given $x = x_0$, the value $Q(x_0, \xi)$ is finite. Then $Q(\cdot, \xi)$ is subdifferentiable at x_0 and*

$$\partial Q(x_0, \xi) = -T^T \mathcal{D}(x_0, \xi), \tag{2.7}$$

where $\mathcal{D}(x, \xi) := \arg\max_{\pi \in \Pi(q)} \pi^T(h - Tx)$ is the set of optimal solutions of the dual problem (2.3).

Proof. Since $Q(x_0, \xi)$ is finite, the set $\Pi(q)$ (defined in (2.5)) is nonempty, and hence the function $s_q(\chi)$ is the support function of the set $\Pi(q)$. It is straightforward to see from the definitions that the support function $s_q(\cdot)$ is the conjugate function of the indicator function[1]

$$i_q(\pi) := \begin{cases} 0, & \text{if } \pi \in \Pi(q), \\ +\infty & \text{otherwise.} \end{cases}$$

Since the set $\Pi(q)$ is closed and convex, the function $i_q(\cdot)$ is convex and lower semicontinuous. It follows then by the Fenchel–Moreau Theorem that the conjugate of $s_q(\cdot)$ is $i_q(\cdot)$. Therefore, for $\chi_0 := h - Tx_0$, we have (see (9.15))

$$\partial s_q(\chi_0) = \arg\max_{\pi} \{\pi^T \chi_0 - i_q(\pi)\} = \arg\max_{\pi \in \Pi(q)} \pi^T \chi_0. \tag{2.8}$$

Since the set $\Pi(q)$ is polyhedral and $s_q(\chi_0)$ is finite, it follows that $\partial s_q(\chi_0)$ is nonempty. Moreover, the function $s_0(\cdot)$ is piecewise linear, and hence formula (2.7) follows from (2.8) by the chain rule of subdifferentiation. □

[1] For a set A its indicator function $i_A(\cdot)$ takes zero value on A and $+\infty$ otherwise.

Problem (2.2) is a particular case of problems of form (9.26). Therefore formula (2.7) is a particular case of the class of such formulas discussed in Section 8.2 (and also in the following Section 1.4, like formula (2.56)). In the present case there is no need for an additional regularity condition (constraint qualification) because the function $Q(\cdot, \xi)$ is polyhedral.

It follows from the above that if the function $Q(\cdot, \xi)$ has a finite value in at least one point, then it is subdifferentiable at that point, and hence it is proper. Its domain can be described in a more explicit way.

The *positive hull* of a matrix W is defined as

$$\operatorname{pos} W := \{\chi : \chi = Wy,\ y \geq 0\}. \tag{2.9}$$

It is a convex polyhedral cone generated by the columns of W. Directly from the definition (2.4) we see that $\operatorname{dom} s_q = \operatorname{pos} W$. Therefore,

$$\operatorname{dom} Q(\cdot, \xi) = \{x : h - Tx \in \operatorname{pos} W\}. \tag{2.10}$$

Suppose that x is such that $\chi = h - Tx \in \operatorname{pos} W$ and let us analyze formula (2.7). The recession cone of $\Pi(q)$ is equal to

$$\Pi_0 := \Pi(0) = \{\pi : W^T \pi \leq 0\}. \tag{2.11}$$

Then it follows from (2.6) that $s_q(\chi)$ is finite iff $\pi^T \chi \leq 0$ for every $\pi \in \Pi_0$, that is, iff χ is an element of the polar cone to Π_0. This polar cone is nothing else but $\operatorname{pos} W$.

If $\chi_0 \in \operatorname{int}(\operatorname{pos} W)$, then the set of maximizers in (2.6) must be bounded. Indeed, if it was unbounded, there would exist an element $\pi_0 \in \Pi_0$ such that $\pi_0^T \chi = 0$. By perturbing χ_0 a little to some χ, we would be able to keep χ within $\operatorname{pos} W$ and get $\pi_0^T \chi > 0$, which is a contradiction, because $\operatorname{pos} W$ is the polar of Π_0. Therefore, the set of maximizers in (2.6) is the convex hull of the vertices v of $\Pi(q)$ for which $v^T \chi = s_q(\chi)$. Note that $\Pi(q)$ must have vertices in this case, because otherwise the polar to Π_0 would have no interior.

If χ_0 is a boundary point of $\operatorname{pos} W$, then the set of maximizers in (2.6) is unbounded. Its recession cone is the intersection of the recession cone Π_0 of $\Pi(q)$ and the set $\{\pi : \pi^T \chi_0 = 0\}$. This intersection is nonempty for boundary points χ_0 and is equal to the normal cone to $\operatorname{pos} W$ at χ_0. Indeed, let π_0 be normal to $\operatorname{pos} W$ at χ_0. Since both χ_0 and $-\chi_0$ are feasible directions at χ_0, we must have $\pi_0^T \chi_0 = 0$. Next, for every $\chi \in \operatorname{pos} W$ we have $\pi_0^T \chi = \pi_0^T (\chi - \chi_0) \leq 0$, so $\pi_0 \in \Pi_0$. The converse argument is similar.

Let us consider now the expected value function

$$\phi(x) := \mathbb{E}[Q(x, \xi)]. \tag{2.12}$$

The expectation here is taken with respect to random components of the data ξ. Suppose that the distribution of ξ has a finite support, i.e., ξ has a finite number of realizations (scenarios) $\xi_k = (q_k, h_k, T_k, W_k)$ with respective (positive) probabilities p_k, $k = 1, \ldots, K$. Then

$$\mathbb{E}[Q(x, \xi)] = \sum_{k=1}^{K} p_k Q(x, \xi_k). \tag{2.13}$$

For a given x, the expectation $\mathbb{E}[Q(x, \xi)]$ is equal to the optimal value of the linear program

$$\begin{aligned} &\operatorname*{Min}_{y_1, \ldots, y_k} \sum_{k=1}^{K} p_k q_k^T y_k \\ &\text{s.t.} \quad T_k x + W_k y_k = h_k, \\ &\qquad y_k \geq 0, \quad k = 1, \ldots, K. \end{aligned} \tag{2.14}$$

Now if for at least one $k \in \{1, \ldots, K\}$ the system $T_k x + W_k y_k = h_k$, $y_k \geq 0$, has no solution, i.e., the corresponding second stage problem is infeasible, then the above problem (2.14) is infeasible, and hence its optimal value is $+\infty$. From that point of view the sum in the right hand side of (2.13) equals $+\infty$ if at least one of $Q(x, \xi_k) = +\infty$. That is, we assume here that $+\infty + (-\infty) = +\infty$.

Proposition 13. *Suppose that the distribution of ξ has a finite support and that the expectation function $\phi(\cdot) := \mathbb{E}[Q(\cdot, \xi)]$ has a finite value in at least one point $\overline{x} \in \mathbb{R}^n$. Then the function $\phi(\cdot)$ is polyhedral, and for any $x_0 \in \operatorname{dom} \phi$,*

$$\partial \phi(x_0) = \sum_{k=1}^{K} p_k \partial Q(x_0, \xi_k). \tag{2.15}$$

Proof. Since $\phi(\overline{x})$ is finite, it follows that all values $Q(\overline{x}, \xi_k)$, $k = 1, \ldots, K$, are finite. Consequently, by Proposition 12, every function $Q(\cdot, \xi_k)$ is polyhedral. It is not difficult to see that a positive linear combination of polyhedral functions is also polyhedral. Therefore, it follows that $\phi(\cdot)$ is polyhedral. We also have that $\operatorname{dom} \phi = \cap_{k=1}^{K} \operatorname{dom} Q_k$, where $Q_k(\cdot) := Q(\cdot, \xi_k)$, and for any $h \in \mathbb{R}^n$, the directional derivatives $Q_k'(x_0, h) > -\infty$ and

$$\phi'(x_0, h) = \sum_{k=1}^{K} p_k Q_k'(x_0, h). \tag{2.16}$$

Formula (2.15) then follows from (2.16) by duality arguments. Note that equation (2.15) is a particular case of the Moreau–Rockafellar Theorem (Theorem 50). Since the functions Q_k are polyhedral, there is no need here for an additional regularity condition for (2.15) to hold. □

The subdifferential $\partial Q(x_0, \xi_k)$ of the second stage optimal value function is described in Proposition 12. That is, if $Q(x_0, \xi_k)$ is finite, then

$$\partial Q(x_0, \xi_k) = -T_k^T \arg\max\{\pi^T(h_k - T_k x_0) \colon W_k^T \pi \leq q_k\}. \tag{2.17}$$

It follows that the expectation function ϕ is differentiable at x_0 iff for every $\xi = \xi_k$, $k = 1, \ldots, K$, the maximum in the right hand side of (2.17) is attained at a unique point, i.e., the corresponding second stage dual problem has a unique optimal solution.

Example 14 (Capacity Expansion). Let us consider a simplified, single-commodity version of the capacity expansion example discussed in Chapter 1 (Example 4). We have a directed graph with node set $\mathcal{N}$ and arc set $\mathcal{A}$. With each arc $a \in \mathcal{A}$ we associate a decision variable x_a and call it the *capacity* of a. There is a cost c_a for each unit of capacity of arc a. The vector x constitutes the vector of first stage variables. They are restricted to satisfy the inequalities $x \geq x^{\min}$, where $x^{\min}$ are the existing capacities.

At each node n of the graph we have a random demand ξ_n for shipments to n (if ξ_n is negative, its absolute value represents shipments from n and we have $\sum_{n \in \mathcal{N}} \xi_n = 0$). These shipments have to be sent through the network and they can be arbitrarily split into pieces taking different paths. We denote by y_a the amount of the shipment sent through arc a. There is a unit cost q_a for shipments on each arc a.

Our objective is to assign the arc capacities and to organize the shipments in such a way that the expected total cost, comprising the capacity cost and the shipping cost, is minimized. The condition is that the capacities have to be assigned *before* the actual demands ξ_n become known, while the shipments can be arranged *after* that.

Let us define the second stage problem. For each node n denote by $\mathcal{A}_+(n)$ and $\mathcal{A}_-(n)$ the sets of arcs entering and leaving node i. The second stage problem is the network flow problem

$$\text{Min} \sum_{a \in \mathcal{A}} q_a y_a \tag{2.18}$$

$$\text{s.t.} \quad \sum_{a \in \mathcal{A}_+(n)} y_a - \sum_{a \in \mathcal{A}_-(n)} y_a = \xi_n, \quad n \in \mathcal{N}, \tag{2.19}$$

$$0 \le y_a \le x_a, \quad a \in \mathcal{A}. \tag{2.20}$$

This problem depends on the random demand vector ξ and on the arc capacities, x. Its optimal value will be denoted $Q(x, \xi)$.

Suppose that for a given $x = x_0$ the second stage problem (2.18)–(2.20) is feasible. Let λ_n, $n \in \mathcal{N}$, be the optimal Lagrange multipliers (node potentials) associated with the node balance equations (2.19). With no loss of generality we can assume that $\lambda_{n_0} = 0$ for some fixed node n_0. Let us denote by $\Lambda(x_0, \xi)$ the set of optimal multiplier vectors (satisfying the additional condition $\lambda_{n_0} = 0$) for a given demand vector ξ.

For each arc $a = (i, j)$ the multiplier π_{ij} associated with the constraint (2.20) has the form

$$\pi_{ij} = \max\{0, \lambda_i - \lambda_j - q_{ij}\}.$$

Roughly, if the difference of node potentials $\lambda_i - \lambda_j$ is greater than q_{ij}, the arc is saturated and the capacity constraint $y_{ij} \le x_{ij}$ is relevant. Since $T = -I$ in this case, formula (2.17) provides the description of the subdifferential of $Q(\cdot, D)$ at x_0:

$$\partial Q(x_0, \xi) = \left\{ \left(\max\{0, \lambda_i - \lambda_j - q_{ij}\} \right)_{(i,j)\in\mathcal{A}} : \lambda \in \Lambda(x_0, \xi) \right\}.$$

The first stage problem has the form

$$\operatorname*{Min}_{x \ge x^{\min}} \sum_{a \in \mathcal{A}} c_a x_a + \mathbb{E}[Q(x, \xi)].$$

If ξ has finitely many realizations ξ_k attained with probabilities p_k, $k = 1, \ldots, K$, the subdifferential of the overall objective can be calculated by (2.15):

$$\partial f(x) = c + \sum_{k=1}^{K} p_k \partial Q(x_0, \xi_k).$$

2.2 *Linear two-stage problems, general distributions*

Let us discuss now the case of a general distribution of the random vector $\xi \in \mathbb{R}^d$. We have here that $Q(\cdot, \cdot)$ is the minimum value of the integrand which is a random lower semicontinuous function. Therefore, it follows by Theorem 19 from the Appendix of Chapter 1 that $Q(\cdot, \cdot)$ is measurable with respect to the Borel sigma algebra of $\mathbb{R}^n \times \mathbb{R}^d$. Also for every ξ the function $Q(\cdot, \xi)$ is lower semicontinuous. It follows that $Q(x, \xi)$ is a random lower

semicontinuous function. Recall that in order to ensure that the expectation $\phi(x)$ is well defined we have to verify two types of conditions, namely that: (i) $Q(x,\cdot)$ is measurable (with respect to the Borel sigma algebra of $\mathbb{R}^d$), (ii) either $\mathbb{E}[Q(x,\xi)_+]$ or $\mathbb{E}[(-Q(x,\xi))_+]$ is finite. Since $Q(x,\cdot)$ is measurable, we only have to verify condition (ii). We describe below some important particular situations where this condition holds.

Two-stage linear program (2.1)–(2.2) is said to have *fixed* recourse if the matrix W is fixed (not random). Moreover, we say that the recourse is *complete* if the system $Wy = \chi$ and $y \geq 0$ has a solution for any χ, or in other words the positive hull of W is equal to the corresponding vector space. By duality arguments the fixed recourse is complete iff the feasible set $\Pi(q)$ of the dual problem (2.3) is bounded (in particular it may be empty) for every q. Then its recession cone, $\Pi_0 = \Pi(0)$, must contain only the point $\{0\}$, provided that $\Pi(q)$ is nonempty. Therefore, another equivalent condition for complete recourse is that $\pi = 0$ is the only solution of the system $W^T\pi \leq 0$.

It is said that the recourse is *relatively complete* if for every x in the set $\{x\colon Ax = b,\ x \geq 0\}$, the feasible set of the second stage problem (2.2) is nonempty for a.e. $\omega \in \Omega$. That is, the recourse is relatively complete if for every feasible first stage point x the inequality $Q(x,\xi(\omega)) < +\infty$ holds w.p.1. This definition is in accordance with the general principle that an event which happens with zero probability is irrelevant for the calculation of the corresponding expected value.

Let Ξ be the support of the random vector $\xi \in \mathbb{R}^d$, i.e., Ξ is the smallest closed subset of $\mathbb{R}^d$ such that the probability of the event $\{\xi \notin \Xi\}$ is zero. A sufficient condition for relatively complete recourse is the following.

- For every $x \geq 0$ such that $Ax = b$, the inequality $Q(x,\xi) < +\infty$ holds for all $\xi \in \Xi$.

Unfortunately, in general, this condition is not necessary. This condition is necessary and sufficient in the following two cases: (i) the random vector $\xi(\omega)$ has a finite support, (ii) the recourse is fixed. Indeed, sufficiency is clear. If $\xi(\omega)$ has a finite support, i.e., the set Ξ is finite, then the necessity is also clear. In order to show the necessity in the case of fixed recourse, we argue as follows. Suppose that the recourse is relatively complete. This means that if x is a feasible point of the first stage problem, then $Q(x,\xi) < +\infty$ for all ξ in Ξ accept possibly for a subset of Ξ of measure zero. We have that $Q(x,\xi) < +\infty$ iff $h - Tx \in \operatorname{pos} W$. Note that this condition does not depend on q and that W is fixed here. Therefore, the condition $h - Tx \in \operatorname{pos} W$ should hold for almost every (h,T). This implies that the set $\{(h,T)\colon h - Tx \in \operatorname{pos} W\}$ should form a dense subset of the support of the probability distribution of $(h(\omega), T(\omega))$. Now if (h_n,T_n) is a sequence converging to a point (h,T) and $h_n - T_n x \in \operatorname{pos} W$, then $h_n - T_n x \to h - Tx$ and hence $h - Tx \in \operatorname{pos} W$ since the set $\operatorname{pos} W$ is closed. This completes the argument. Let us also note that, of course, if a two-stage linear program with fixed recourse is complete, then it is relatively complete.

Example 15. Consider

$$Q(x,\xi) := \inf\{y\colon \xi y = x,\ y \geq 0\},$$

with $x \in [0,1]$ and ξ being a random variable whose probability density function is $p(z) := 2z$, $0 \leq z \leq 1$. We have that for any $\xi > 0$ and $x \in [0,1]$, $Q(x,\xi) = x/\xi$, and hence

$$\mathbb{E}[Q(x,\xi)] = \int_0^1 \left(\frac{x}{z}\right) 2z \, \mathrm{d}z = 2x.$$

That is, the recourse here is relatively complete and the expectation of $Q(x,\xi)$ is finite. On the other hand, the support of $\xi(\omega)$ is the interval $[0,1]$, and for $\xi = 0$ and $x > 0$ the value of $Q(x,\xi)$ is $+\infty$ since the corresponding problem is infeasible. Of course, probability of the event "$\xi = 0$" is zero, and from the mathematical point of view the expected value function $\mathbb{E}[Q(x,\xi)]$ is well defined and finite for all $x \in [0,1]$. Note, however, that arbitrary small perturbation of the probability distribution of ξ may change that. Take, for example, a discretization of the distribution of ξ with the first discretization point $t = 0$. Then, does not matter how small is the assigned (positive) probability at $t = 0$, we obtain that $Q(x,\xi) = +\infty$ with positive probability, and hence, $\mathbb{E}[Q(x,\xi)] = +\infty$, for any $x > 0$. That is, from a numerical point of view the above problem is extremely unstable and is not well posed. As it was discussed above, such behavior cannot happen if the recourse is fixed.

Let us consider the support function $s_q(\cdot)$ of the set $\Pi(q)$. We want to find sufficient conditions for the existence of the expectation $\mathbb{E}[s_q(h - Tx)]$. By Hoffman's lemma (Theorem 53) there exists a constant κ, depending on W, such that if for some q_0 the set $\Pi(q_0)$ is nonempty, then for any q the following inclusion holds

$$\Pi(q) \subset \Pi(q_0) + \kappa \|q - q_0\| B, \tag{2.21}$$

where $B := \{\pi\colon \|\pi\| \leq 1\}$ and $\|\cdot\|$ denotes the Euclidean norm. Consider the support function $s_q(\cdot)$ of the set $\Pi(q)$. Since the support function of the unit ball B is $\|\cdot\|$, it follows from (2.21) that if the set $\Pi(q_0)$ is nonempty, then

$$s_q(\cdot) \leq s_{q_0}(\cdot) + \kappa \|q - q_0\| \, \|\cdot\|. \tag{2.22}$$

In particular, for the support function $s_0(\cdot)$ of the cone Π_0 we have that $s_0(\chi) = 0$ if $\chi \in \operatorname{pos} W$ and $s_0(\chi) = +\infty$ otherwise. Therefore, by taking $q_0 = 0$ in (2.22) we obtain that if $\Pi(q)$ is nonempty, then $s_q(\chi) \leq \kappa \|q\| \, \|\chi\|$ for $\chi \in \operatorname{pos} W$, and $s_q(\chi) = +\infty$ for $\chi \notin \operatorname{pos} W$. Since $\Pi(q)$ is polyhedral, if $\Pi(q)$ is

nonempty then $s_q(\cdot)$ is piecewise linear on its domain, which coincides with pos W, and

$$|s_q(\chi_1) - s_q(\chi_2)| \leq \kappa \|q\| \, \|\chi_1 - \chi_2\|, \quad \forall \, \chi_1, \chi_2 \in \text{pos } W. \tag{2.23}$$

We say that a random variable $Z(\omega)$ has a finite r-th moment if $\mathbb{E}|Z(\omega)|^r < +\infty$. It is said that $\xi(\omega)$ has finite r-th moments if each component of $\xi(\omega)$ has a finite r-th moment.

Proposition 16. *Suppose that the recourse is fixed and*

$$\mathbb{E}[\|q\| \, \|h\|] < +\infty \quad \textit{and} \quad \mathbb{E}[\|q\| \, \|T\|] < +\infty. \tag{2.24}$$

Consider a point $x \in \mathbb{R}^n$. Then $\mathbb{E}[Q(x,\xi)_+]$ is finite if and only if the following condition holds w.p.1:

$$h - Tx \in \text{pos } W. \tag{2.25}$$

Proof. We have that $Q(x,\xi) < +\infty$ iff condition (2.25) holds. Therefore, if condition (2.25) does not hold w.p.1, then $Q(x,\xi) = +\infty$ with positive probability, and hence $\mathbb{E}[Q(x,\xi)_+] = +\infty$.

Conversely, suppose that condition (2.25) holds w.p.1. Then $Q(x,\xi) = s_q(h - Tx)$ with $s_q(\cdot)$ being the support function of the set $\Pi(q)$. By (2.22) there exists a constant κ such that for any χ,

$$s_q(\chi) \leq s_0(\chi) + \kappa \|q\| \, \|\chi\|. \tag{2.26}$$

Also for any $\chi \in \text{pos } W$ we have that $s_0(\chi) = 0$, and hence w.p.1,

$$s_q(h - Tx) \leq \kappa \|q\| \, \|h - Tx\| \leq \kappa \|q\| (\|h\| + \|T\| \, \|x\|). \tag{2.27}$$

It follows then by (2.24) that $\mathbb{E}\big[s_q(h - Tx)_+\big] < +\infty$. $\square$

Remark 17. If q and (h,T) are independent and have finite first moments, then

$$\mathbb{E}\|q\| \, \|h\|] = \mathbb{E}[\|q\|] \, \mathbb{E}[\|h\|] \quad \text{and} \quad \mathbb{E}[\|q\| \, \|T\|] = \mathbb{E}[\|q\|] \, \mathbb{E}[\|T\|],$$

and hence condition (2.24) follows. Also condition (2.24) holds if (h, T, q) has finite second moments.

We obtain that, under the assumptions of Proposition 16, the expectation $\phi(x)$ is well defined and $\phi(x) < +\infty$ iff condition (2.25) holds w.p.1. If, moreover, the recourse is complete, then (2.25) holds for any x and ξ, and

hence $\phi(\cdot)$ is well defined and is less than $+\infty$. Since the function $\phi(\cdot)$ is convex, we have that if $\phi(\cdot)$ is less than $+\infty$ on $\mathbb{R}^n$ and is finite valued in at least one point, then $\phi(\cdot)$ is finite valued on the entire space $\mathbb{R}^n$.

Proposition 18. *Suppose that:* (i) *the recourse is fixed,* (ii) *for a.e. q the set $\Pi(q)$ is nonempty,* (iii) *condition (2.24) holds. Then the expectation function $\phi(x)$ is well defined and $\phi(x) > -\infty$ for all $x \in \mathbb{R}^n$. Moreover, ϕ is convex, lower semicontinuous and Lipschitz continuous on* $\operatorname{dom}\phi$*, its domain* $\operatorname{dom}\phi$ *is a convex closed subset of $\mathbb{R}^n$ and*

$$\operatorname{dom}\phi = \{x \in \mathbb{R}^n : h - Tx \in \operatorname{pos} W \ w.p.1\}. \tag{2.28}$$

Proof. Since by assumption (ii) the feasible set $\Pi(q)$ of the dual problem is nonempty w.p.1, we have that $Q(x,\xi)$ is equal to $s_q(h - Tx)$ w.p.1 for any x, where $s_q(\cdot)$ is the support function of the set $\Pi(q)$. Let $\pi(q)$ be the element of the set $\Pi(q)$ that is closest to 0. Since $\Pi(q)$ is closed, such an element exists. By Hoffman's lemma (see (2.21)) there is a constant κ such that $\|\pi(q)\| \le \kappa\|q\|$. Then for any x the following holds w.p.1:

$$s_q(h - Tx) \ge \pi(q)^T(h - Tx) \ge -\kappa\|q\|(\|h\| + \|T\|\,\|x\|). \tag{2.29}$$

By condition (2.24) it follows from (2.29) that $\phi(\cdot)$ is well defined and $\phi(x) > -\infty$ for all $x \in \mathbb{R}^n$. Moreover, since $s_q(\cdot)$ is lower semicontinuous, the lower semicontinuity of $\phi(\cdot)$ follows by Fatou's lemma. The convexity and the closedness of $\operatorname{dom}\phi$ follow from the convexity and the lower semicontinuity of ϕ. We have by Proposition 16 that $\phi(x) < +\infty$ iff condition (2.25) holds w.p.1. This implies (2.28).

Consider two points $x, x' \in \operatorname{dom}\phi$. Then by (2.28) the following holds w.p.1:

$$h - Tx \in \operatorname{pos} W \quad \text{and} \quad h - Tx' \in \operatorname{pos} W. \tag{2.30}$$

By (2.23) we have that if the set $\Pi(q)$ is nonempty and (2.30) holds, then

$$|s_q(h - Tx) - s_q(h - Tx')| \le \kappa\|q\|\,\|T\|\,\|x - x'\|. \tag{2.31}$$

It follows that

$$|\phi(x) - \phi(x')| \le \kappa\,\mathbb{E}[\|q\|\,\|T\|]\,\|x - x'\|.$$

Together with condition (2.24) this implies the Lipschitz continuity of ϕ on its domain.

Denote by Σ the support[2] of the probability distribution (measure) of (h, T). Formula (2.28) means that a point x belongs to $\operatorname{dom}\phi$ iff the probability of the event $\{h - Tx \in \operatorname{pos} W\}$ is one. Note that the set $\{(h, T): h - Tx \in \operatorname{pos} W\}$ is convex and polyhedral, and hence is closed. Consequently x belongs to $\operatorname{dom}\phi$ iff for every $(h, T) \in \Sigma$ it follows that $h - Tx \in \operatorname{pos} W$. Therefore, we can write formula (2.28) in the form

$$\operatorname{dom}\phi = \bigcap_{(h,T)\in\Sigma} \{x: h - Tx \in \operatorname{pos} W\}. \tag{2.32}$$

It should be noted that the above holds since we assume here that the recourse is fixed.

Let us observe that for any set $\mathcal{H}$ of vectors h, the set $\cap_{h\in\mathcal{H}}(-h + \operatorname{pos} W)$ is convex and polyhedral. Indeed, we have that $\operatorname{pos} W$ is a convex polyhedral cone, and hence can be represented as the intersection of a finite number of half spaces $A_i = \{\chi: a_i^T\chi \le 0\}$, $i = 1, \ldots, \ell$. Since the intersection of any number of half spaces of the form $b + A_i$, with $b \in B$, is still a half space of the same form (provided that this intersection is nonempty), we have the set $\cap_{h\in\mathcal{H}}(-h + \operatorname{pos} W)$ can be represented as the intersection of half spaces of the form $b_i + A_i$, $i = 1, \ldots, \ell$, and hence is polyhedral. It follows that if T and W are fixed, then the set at the right hand side of (2.32) is convex and polyhedral.

Let us discuss now the differentiability properties of the expectation function $\phi(x)$. By Theorem 9 and formula (2.7) of Proposition 12 we have the following result.

Proposition 19. *Suppose that the expectation function $\phi(\cdot)$ is proper and its domain has a nonempty interior. Then for any $x_0 \in \operatorname{dom}\phi$,*

$$\partial\phi(x_0) = -T^T\mathbb{E}[\mathcal{D}(x_0, \xi)] + N_{\operatorname{dom}\phi}(x_0), \tag{2.33}$$

where $\mathcal{D}(x, \xi) := \arg\max_{\pi\in\Pi(q)}\pi^T(h - Tx)$. Moreover, ϕ is differentiable at x_0 if and only if x_0 belongs to the interior of $\operatorname{dom}\phi$ and the set $\mathcal{D}(x_0, \xi)$ is a singleton w.p.1.

As we discussed earlier, when the distribution of ξ has a finite support (i.e., there is a finite number of scenarios) the expectation function ϕ is piecewise linear on its domain and is differentiable everywhere only in the trivial case if it is linear.[3] In the case of a continuous distribution of ξ the expectation operator smoothes the piecewise linear function $Q(\cdot, \xi)$ out. For example, the following result holds.

[2] Recall that the support of a probability distribution (measure) is the smallest closed set such that the probability (measure) of its complement is zero.

[3] By linear we mean here that it is of the form $a^Tx + b$. It is more accurate to call such a function affine.

Proposition 20. *Suppose that the assumptions of Proposition 18 hold, and the conditional distribution of h, given (T, q), is absolutely continuous for almost all (T, q). Then ϕ is continuously differentiable on the interior of its domain.*

Proof. By Proposition 18 we have here that the expectation function $\phi(\cdot)$ is well defined and is greater than $-\infty$. Let x be a point in the interior of $\operatorname{dom}\phi$. For fixed T and q, consider the multifunction

$$\mathcal{Z}(h) := \arg\max_{\pi \in \Pi(q)} \pi^T (h - Tx).$$

Clearly, conditional on (T, q), the set $\mathcal{D}(x, \xi)$ coincides with $\mathcal{Z}(h)$. Since $x \in \operatorname{dom}\phi$, we have by (2.28) that $h - Tx \in \operatorname{pos} W$ w.p.1. Also we have that if $h - Tx \in \operatorname{pos} W$, then $\mathcal{Z}(h)$ is nonempty and forms a face of the polyhedral set $\Pi(q)$. Moreover, there is a set A given by the union of a finite number of linear subspaces of $\mathbb{R}^m$ (where m is the dimension of h) perpendicular to the faces of sets $\Pi(q)$, such that if $h - Tx \in (\operatorname{pos} W) \backslash A$, then $\mathcal{Z}(h)$ is a singleton. Since an affine subspace of $\mathbb{R}^m$ has Lebesgue measure zero, it follows that the Lebesgue measure of A is zero. Therefore, since the conditional distribution of h given (T, q) is absolutely continuous, the probability of the event that $\mathcal{Z}(h)$ is not a singleton is zero. By integrating this probability over the distribution of (T, q), we obtain that the probability of the event that $\mathcal{D}(x, \xi)$ is not a singleton is zero. By Proposition 19 this implies the differentiability of $\phi(\cdot)$. Since $\phi(\cdot)$ is convex, it follows that for any $x \in \operatorname{int}(\operatorname{dom}\phi)$ the gradient $\nabla\phi(x)$ coincides with the (unique) subgradient of ϕ at x, and that $\nabla\phi(x)$ is continuous at x. □

Of course, if h and (T, q) are independent, then the conditional distribution of h given (T, q) is the same as the unconditional (marginal) distribution of h. Therefore, if h and (T, q) are independent, then it suffices to assume in the above proposition that the (marginal) distribution of h is absolutely continuous.

2.3 Polyhedral two-stage problems

Let us consider a slightly more general formulation of a two-stage stochastic programming problem:

$$\operatorname*{Min}_x \; f_1(x) + \mathbb{E}[Q(x, \omega)], \tag{2.34}$$

where $Q(x, \omega)$ is the optimal value of the second stage problem

$$\operatorname*{Min}_y \; f_2(y, \omega) \quad \text{subject to} \quad T(\omega)x + W(\omega)y = h(\omega). \tag{2.35}$$

We assume in this section that the above two-stage problem is *polyhedral*. That is, the following holds.

- The function $f_1(\cdot)$ is *polyhedral* (see Definition 10). This means that there exist vectors c_j and scalars α_j, $j = 1, \ldots, J_1$, vectors a_k and scalars b_k, $k = 1, \ldots, K_1$, such that $f_1(x)$ can be represented as follows:

$$f_1(x) = \begin{cases} \max_{1 \le j \le J_1} \alpha_j + c_j^T x, & \text{if } a_k^T x \le b_k, \quad k = 1, \ldots, K_1, \\ +\infty, & \text{otherwise,} \end{cases} \tag{2.36}$$

  and its domain $\operatorname{dom} f_1 = \{x : a_k^T x \le b_k,\ k = 1, \ldots, K_1\}$ is nonempty.
- The function f_2 is *random polyhedral*. That is, there exist random vectors $q_j = q_j(\omega)$ and random scalars $\gamma_j = \gamma_j(\omega)$, $j = 1, \ldots, J_2$, random vectors $d_k = d_k(\omega)$ and random scalars $r_k = r_k(\omega)$, $k = 1, \ldots, K_2$, such that $f_2(y, \omega)$ can be represented as follows:

$$f_2(y, \omega) = \begin{cases} \max_{1 \le j \le J_2} \gamma_j(\omega) + q_j(\omega)^T y, & \text{if } d_k(\omega)^T y \le r_k(\omega),\ k = 1, \ldots, K_2, \\ +\infty, & \text{otherwise,} \end{cases} \tag{2.37}$$

  and for a.e. ω the domain of $f_2(\cdot, \omega)$ is nonempty.

Clearly, the linear two-stage model (2.1)–(2.2) is a special case of a polyhedral two-stage problem. The converse is also true, i.e., every polyhedral two-stage model can be reformulated as a linear two-stage model. For example, the second stage problem (2.35) can be written as follows:

$$\begin{aligned} \operatorname*{Min}_{y, v} \ & v \\ \text{s.t.} \ & T(\omega)x + W(\omega)y = h(\omega), \\ & \gamma_j(\omega) + q_j(\omega)^T y \le v, \quad j = 1, \ldots, J_2, \\ & d_k(\omega)^T y \le r_k(\omega), \quad k = 1, \ldots, K_2. \end{aligned}$$

Here both v and y play the role of the second stage variables, and the data (q, T, W, h) in (2.2) have to be re-defined in an appropriate way. In order to avoid all these manipulations and the unnecessary notational complications that come together with such a conversion, we shall address polyhedral problems in a more abstract way. This will also help us to deal with the general convex case.

Consider the Lagrangian

$$L(y, \pi;\ x, \omega) := f_2(y, \omega) + \pi^T (h(\omega) - T(\omega)x - W(\omega)y)$$

of the second stage problem (2.35). We have

$$\inf_y L(y,\pi;\ x,\omega) = \pi^T(h(\omega) - T(\omega)x) + \inf_y \big[f_2(y,\omega) - \pi^T W(\omega)y\big]$$
$$= \pi^T(h(\omega) - T(\omega)x) - f_2^*(W(\omega)^T\pi,\omega),$$

where $f_2^*(\cdot,\omega)$ is the conjugate[4] of $f_2(\cdot,\omega)$. We obtain that the dual of problem (2.35) can be written as follows

$$\operatorname*{Max}_\pi \big[\pi^T(h(\omega) - T(\omega)x) - f_2^*(W(\omega)^T\pi,\omega)\big]. \tag{2.38}$$

By the duality theory of linear programming we have that if, for some (x,ω), the optimal value $Q(x,\omega)$ of problem (2.35) is less than $+\infty$ (i.e., problem (2.35) is feasible), then it is equal to the optimal value of the dual problem (2.38).

Let us denote, as before, by $\mathcal{D}(x,\omega)$ the set of optimal solutions of the dual problem (2.38). We then have an analogue of Proposition 12.

Proposition 21. *Let $\omega \in \Omega$ be given and suppose that $Q(\cdot,\omega)$ is finite in at least one point x. Then the function $Q(\cdot,\omega)$ is convex polyhedral. Moreover, $Q(\cdot,\omega)$ is subdifferentiable at every $x = x_0$, at which the value $Q(x_0,\omega)$ is finite, and*

$$\partial Q(x_0,\omega) = -T(\omega)^T\mathcal{D}(x_0,\omega). \tag{2.39}$$

Proof. Let us define the function $\psi(\pi) := f_2^*(W^T\pi)$ (for simplicity we suppress the argument ω). We have that if $Q(x,\omega)$ is finite, then it is equal to the optimal value of problem (2.38), and hence $Q(x,\omega) = \psi^*(h - Tx)$. Therefore, $Q(\cdot,\omega)$ is a polyhedral function. Moreover, it follows by the Fenchel–Moreau Theorem that

$$\partial\psi^*(h - Tx_0) = \mathcal{D}(x_0,\omega),$$

and the chain rule for subdifferentiation yields formula (2.39). Note that we do not need here additional regularity conditions because of the polyhedricity of the considered case. □

If $Q(x,\omega)$ is finite, then the set $\mathcal{D}(x,\omega)$ of optimal solutions of problem (2.38) is a nonempty convex closed polyhedron. If, moreover, $\mathcal{D}(x,\omega)$ is bounded, then it is the convex hull of its finitely many vertices (extreme points), and $Q(\cdot,\omega)$ is finite in a neighborhood of x. If $\mathcal{D}(x,\omega)$ is unbounded,

[4] Note that since $f_2(\cdot,\omega)$ is polyhedral, so is $f_2^*(\cdot,\omega)$.

then its recession cone (which is polyhedral) is the normal cone to the domain of $Q(\cdot, \omega)$ at the point x.

Let us consider now the expected value function $\phi(x) := \mathbb{E}[Q(x, \omega)]$. Suppose that the probability measure P has a finite support, i.e., there exists a finite number of scenarios ω_k with respective (positive) probabilities p_k, $k = 1, \ldots, K$. Then

$$\mathbb{E}[Q(x, \omega)] = \sum_{k=1}^{K} p_k Q(x, \omega_k). \tag{2.40}$$

For a given x, the expectation $\mathbb{E}[Q(x, \omega)]$ is equal to the optimal value of the program

$$\begin{aligned} &\operatorname*{Min}_{y_1, \ldots, y_K} \sum_{k=1}^{K} p_k f_2(y_k, \omega_k) \\ &\text{s.t.} \quad T_k x + W_k y_k = h_k, \quad k = 1, \ldots, K, \end{aligned} \tag{2.41}$$

where $(h_k, T_k, W_k) := (h(\omega_k), T(\omega_k), W(\omega_k))$. Similarly to the linear case, if for at least one $k \in \{1, \ldots, K\}$ the set

$$\operatorname{dom} f_2(\cdot, \omega_k) \cap \{y : T_k x + W_k y = h_k\}$$

is empty, i.e., the corresponding second stage problem is infeasible, then problem (2.41) is infeasible, and hence its optimal value is $+\infty$.

Proposition 22. *Suppose that the probability measure P has a finite support and that the expectation function $\phi(\cdot) := \mathbb{E}[Q(\cdot, \omega)]$ has a finite value in at least one point $x \in \mathbb{R}^n$. Then the function $\phi(\cdot)$ is polyhedral, and for any $x_0 \in \operatorname{dom} \phi$,*

$$\partial \phi(x_0) = \sum_{k=1}^{K} p_k \partial Q(x_0, \omega_k). \tag{2.42}$$

The proof is identical to the proof of Proposition 13. Since the functions $Q(\cdot, \omega_k)$ are polyhedral, formula (2.42) follows by the Moreau–Rockafellar Theorem.

The subdifferential $\partial Q(x_0, \omega_k)$ of the second stage optimal value function is described in Proposition 21. That is, if $Q(x_0, \omega_k)$ is finite, then

$$\partial Q(x_0, \omega_k) = -T_k^T \arg\max \{\pi^T (h_k - T_k x_0) - f_2^*(W_k^T \pi, \omega_k)\}. \tag{2.43}$$

It follows that the expectation function ϕ is differentiable at x_0 iff for every ω_k, $k = 1, \ldots, K$, the maximum at the right hand side of (2.43) is attained at a unique point, i.e., the corresponding second stage dual problem has a unique optimal solution.

Let us now consider the case of a general probability distribution P. We need to ensure that the expectation function $\phi(x) := \mathbb{E}[Q(x, \omega)]$ is well defined. General conditions are messy, so we resort again to the case of fixed recourse.

We say that the two-stage polyhedral problem has *fixed recourse* if the matrix W and the set[5] $Y := \operatorname{dom} f_2(\cdot, \omega)$ are fixed, i.e., do not depend on ω. In that case,

$$f_2(y, \omega) = \begin{cases} \max_{1 \le j \le J_2} \gamma_j(\omega) + q_j(\omega)^T y, & \text{if } y \in Y, \\ +\infty, & \text{otherwise.} \end{cases}$$

Denote $W(Y) := \{Wy : y \in Y\}$. Let x be such that

$$h(\omega) - T(\omega)x \in W(Y) \text{ w.p.1.} \tag{2.44}$$

This means that for a.e. ω the system

$$y \in Y, \; y = h(\omega) - T(\omega)x \tag{2.45}$$

has a solution. Let for some $\omega_0 \in \Omega$, y_0 be a solution of the above system, i.e., $y_0 \in Y$ and $h(\omega_0) - T(\omega_0)x = Wy_0$. Since system (2.45) is defined by linear constraints, we have by Hoffman's lemma that there exists a constant κ such that for almost all ω we can find a solution $\overline{y}(\omega)$ of the system (2.45) with

$$\|\overline{y}(\omega) - y_0\| \le \kappa \|(h(\omega) - T(\omega)x) - (h(\omega_0) - T(\omega_0)x)\|.$$

Therefore the optimal value of the second stage problem can be bounded from above as follows:

$$\begin{aligned} Q(x, \omega) &\le \max_{1 \le j \le J_2} \left\{ \gamma_j(\omega) + q_j(\omega)^T \overline{y}(\omega) \right\} \\ &\le Q(x, \omega_0) + \sum_{j=1}^{J_2} |\gamma_j(\omega) - \gamma_j(\omega_0)| \\ &\quad + \kappa \sum_{j=1}^{J_2} \|q_j(\omega)\| (\|h(\omega) - h(\omega_0)\| + \|x\| \, \|T(\omega) - T(\omega_0)\|). \end{aligned} \tag{2.46}$$

[5] Note that since it is assumed that $f_2(\cdot, \omega)$ is polyhedral, it follows that the set Y is nonempty and polyhedral.

Proposition 23. *Suppose that the recourse is fixed and*

$$\mathbb{E}|\gamma_j| < +\infty,\ \mathbb{E}\big[\|q_j\|\ \|h\|\big] < +\infty \quad \text{and} \quad \mathbb{E}\big[\|q_j\|\ \|T\|\big] < +\infty,\quad j = 1, \ldots, J_2. \tag{2.47}$$

Consider a point $x \in \mathbb{R}^n$. Then $\mathbb{E}[Q(x, \omega)_+]$ is finite if and only if condition (2.44) holds.

The proof uses (2.46), similarly to the proof of Proposition 16.

Let us now formulate conditions under which the expected recourse cost is bounded from below. Let C be the recession cone of Y, and C^* be its polar. Consider the conjugate function $f_2^*(\cdot, \omega)$. It can be verified that

$$\operatorname{dom} f_2^*(\cdot, \omega) = \operatorname{conv}\{q_j(\omega), j = 1, \ldots, J_2\} + C^*. \tag{2.48}$$

Indeed, by the definition of the function $f_2(\cdot, \omega)$ and its conjugate, we have that $f_2^*(z, \omega)$ is equal to the optimal value of the problem

$$\begin{aligned} &\operatorname{Max}_{y,\, v}\ v \\ &\text{s.t.}\quad z^T y - \gamma_j(\omega) - q_j(\omega)^T y \geq v, \quad j = 1, \ldots, J_2,\ y \in Y. \end{aligned}$$

Since it is assumed that the set Y is nonempty, the above problem is feasible, and since Y is polyhedral, it is linear. Therefore its optimal value is equal to the optimal value of its dual. In particular, its optimal value is less than $+\infty$ iff the dual problem is feasible. Now the dual problem is feasible iff there exist $\pi_j \geq 0$, $j = 1, \ldots, J_2$, such that $\sum_{j=1}^{J_2} \pi_j = 1$ and

$$\sup_{y \in Y} y^T \Big(z - \textstyle\sum_{j=1}^{J_2} \pi_j q_j(\omega) \Big) < +\infty.$$

The last condition holds iff $z - \sum_{j=1}^{J_2} \pi_j q_j(\omega) \in C^*$, which completes the argument.

Let us define the set

$$\Pi(\omega) := \{\pi \colon W^T \pi \in \operatorname{conv}\big\{q_j(\omega), j = 1, \ldots, J_2\big\} + C^*\}.$$

We may remark that in the case of a linear two stage program the above set coincides with the one defined in (2.5).

Proposition 24. *Suppose that:* (i) *the recourse is fixed,* (ii) *the set $\Pi(\omega)$ is nonempty w.p.1,* (iii) *condition (2.47) holds. Then the expectation function $\phi(x)$ is well defined and $\phi(x) > -\infty$ for all $x \in \mathbb{R}^n$. Moreover, ϕ is convex, lower*

semicontinuous and Lipschitz continuous on $\operatorname{dom}\phi$*, its domain* $\operatorname{dom}\phi$ *is a convex closed subset of* $\mathbb{R}^n$ *and*

$$\operatorname{dom}\phi = \{x \in \mathbb{R}^n : h - Tx \in W(Y) \ w.p.1\}. \tag{2.49}$$

Note that the dual problem (2.38) is feasible iff $W^T\pi \in \operatorname{dom} f_2^*(\cdot,\omega)$. By formula (2.48) assumption (ii) means that problem (2.38) is feasible, and hence $Q(x,\omega)$ is equal to the optimal value of (2.38), for a.e. ω. The remainder of the proof is similar to the linear case (Proposition 18).

2.4 Convex two-stage problems

Let us consider the following two-stage problem

$$\operatorname*{Min}_{x\in X}\{f(x) := \mathbb{E}[F(x,\omega)]\}, \tag{2.50}$$

where $F(x,\omega)$ is the optimal value of the second stage problem

$$\operatorname*{Min}_{y\in Y} q(y,\omega) \text{ subject to } g_i(y,\omega) + \chi_i \le 0, \quad i = 1,\ldots,m, \tag{2.51}$$

and $\chi_i = t_i(x,\omega)$. Here X is a subset of $\mathbb{R}^n$, Y is a subset of $\mathbb{R}^s$, and $q(y,\omega)$, $g_i(y,\omega)$ and $t_i(x,\omega)$ are real valued functions. The above problem is a particular case of the two-stage problem (2.19)–(2.20) discussed in Chapter 1. We assume throughout this section that for a.e. $\omega \in \Omega$ the problem (2.51) is *convex*, that is, the set Y is convex, and the functions $q(\cdot,\omega)$ and $g_i(\cdot,\omega)$, $t_i(\cdot,\omega)$, $i = 1,\ldots,m$, are convex. Recall that real valued convex functions are continuous, in fact they are even locally Lipschitz continuous.

The second stage constraints can be absorbed into the objective function by defining $\overline{q}(y,\chi,\omega) := q(y,\omega)$ if (y,χ) satisfies the constraints of (2.51), and $\overline{q}(y,\chi,\omega) := +\infty$ otherwise. Consequently, problem (2.51) can be written as

$$\operatorname*{Min}_{y\in\mathbb{R}^s} \overline{q}(y,\chi,\omega). \tag{2.52}$$

Our convexity assumptions imply that for a.e. $\omega \in \Omega$ the function $\overline{q}(\cdot,\cdot,\omega)$ is convex. Therefore we can study this problem in the framework of conjugate duality discussed in Section 7.2 of the Appendix.[6]

Let us denote by $\vartheta(\chi,\omega)$ the optimal value of problem (2.51), or equivalently of problem (2.52). Note that $F(x,\omega) = \vartheta(T(x,\omega),\omega)$, where

[6] Note that in order to be consistent with the notation of two-stage programming, in the present case the optimization in (2.51) is performed with respect to y while in Section 7.2 the corresponding optimization is performed with respect to x.

$T(x,\omega) := (t_1(x,\omega),\ldots,t_m(x,\omega))$. The dual of problem (2.51) can be written in the form

$$\operatorname*{Max}_{\pi\geq 0}\left\{\pi^T\chi + \inf_{y\in Y} L(y,\pi,\omega)\right\}, \tag{2.53}$$

where

$$L(y,\pi,\omega) := q(y,\omega) + \sum_{i=1}^{m} \pi_i g_i(y,\omega)$$

is the Lagrangian of problem (2.51). By the theory of conjugate duality we have the following results (see Proposition 55).

Proposition 25. *Let χ and $\omega\in\Omega$ be given. Suppose that problem (2.51) is convex. Then the following holds.* (i) *The functions $\vartheta(\cdot,\omega)$ and $F(\cdot,\omega)$ are convex.* (ii) *Suppose that problem (2.51) is subconsistent. Then there is no duality gap between problem (2.51) and its dual (2.53) if and only if the optimal value function $\vartheta(\cdot,\omega)$ is lower semicontinuous at χ.* (iii) *There is no duality gap between problems (2.51) and (2.53) and the dual problem (2.53) has a nonempty set of optimal solutions if and only if the optimal value function $\vartheta(\cdot,\omega)$ is subdifferentiable at χ.* (iv) *Suppose that the optimal value of (2.51) is finite. Then there is no duality gap between problems (2.51) and (2.53) and the dual problem (2.53) has a nonempty and bounded set of optimal solutions if and only if $\chi\in\operatorname{int}(\operatorname{dom}\vartheta(\cdot,\omega))$.*

The regularity condition $\chi\in\operatorname{int}(\operatorname{dom}\vartheta(\cdot,\omega))$ means that for all small perturbations of χ the corresponding problem (2.51) remains feasible. We can also characterize the differentiability properties of the optimal value functions in terms of the dual problem (2.53). Let us denote by $\mathcal{D}(\chi,\omega)$ the (possibly empty) set of optimal solutions of the dual problem (2.53).

Proposition 26. *Let $\omega\in\Omega$ and $\chi = T(x,\omega)$ be given. Suppose that problem (2.51) is convex, and that problems (2.51) and (2.53) have finite and equal optimal values. Then*

$$\partial\vartheta(\chi,\omega) = \mathcal{D}(\chi,\omega). \tag{2.54}$$

Suppose, further, that the functions $t_i(\cdot,\omega)$, $i=1,\ldots,m$, are differentiable, and that the condition

$$0\in\operatorname{int}\{T(x,\omega) + \nabla T(x,\omega)\mathbb{R}^s - \operatorname{dom}\vartheta(\cdot,\omega)\} \tag{2.55}$$

holds. Then

$$\partial F(x,\omega) = \nabla T(x,\omega)^T \mathcal{D}(\chi,\omega). \tag{2.56}$$

As before, all subdifferentials and derivatives in the above formulas are taken with respect to x and χ.

Corollary 27. *Let $\omega \in \Omega$ and $\chi = T(x,\omega)$ be given. Suppose that problem (2.51) is convex. Then $\vartheta(\cdot,\omega)$ is differentiable at χ if and only if $\mathcal{D}(\chi,\omega)$ is a singleton. Suppose, further, that the functions $t_i(\cdot,\omega)$, $i = 1,\ldots,m$, are differentiable. Then $F(\cdot,\omega)$ is differentiable at x if $\mathcal{D}(\chi,\omega)$ is a singleton.*

Proof. If $\mathcal{D}(\chi,\omega)$ is a singleton, then the set of optimal solutions of the dual problem (2.53) is nonempty and bounded, and hence there is no duality gap between problems (2.51) and (2.53). Thus formula (2.54) holds. Conversely, if $\partial\vartheta(\chi,\omega)$ is a singleton and hence is nonempty, then again there is no duality gap between problems (2.51) and (2.53), and hence formula (2.54) holds.

Now if $\mathcal{D}(\chi,\omega)$ is a singleton, then $\vartheta(\cdot,\omega)$ is continuous at χ and hence the regularity condition (2.55) holds. It follows then by formula (2.56) that $F(\cdot,\omega)$ is differentiable at x and formula

$$\nabla F(x,\omega) = \nabla T(x,\omega)^T \mathcal{D}(\chi,\omega) \tag{2.57}$$

holds. □

Let us discuss now properties of the expectation function $f(x) := \mathbb{E}[F(x,\omega)]$. If the set Ω is finite, say $\Omega = \{\omega_1,\ldots,\omega_K\}$ with corresponding probabilities p_k, $k = 1,\ldots,K$, then $f(x) = \sum_{k=1}^K p_k F(x,\omega_k)$ and subdifferentiability of $f(x)$ is described in Theorem 6 together with formula (2.56). In particular, we obtain that $f(\cdot)$ is differentiable at a point x if the functions $t_i(\cdot,\omega)$, $i = 1,\ldots,m$, are differentiable at x and for every $\omega \in \Omega$ the corresponding dual problem (2.53) has a unique optimal solution.

Let us discuss now the general case where Ω does not need to be finite. Assume that the functions $q(y,\omega)$ and $g_i(y,\omega)$, $t_i(x,\omega)$, $i = 1,\ldots,m$, are random lower semicontinuous. Then it follows that the function $\overline{q}(y,T(x,\omega),\omega)$ is also random lower semicontinuous. Consequently, we obtain by Theorem 19 from the Appendix of Chapter 1 that for any x the optimal (minimal) value $F(x,\cdot)$ is measurable. If, moreover, for a.e. $\omega \in \Omega$ the function $F(\cdot,\omega)$ is lower semicontinuous, then the integrand $F(x,\omega)$ is random lower semicontinuous. Since, by convexity, the functions $t_i(\cdot,\omega)$ are continuous, we have that $F(\cdot,\omega)$ is lower semicontinuous if $\vartheta(\cdot,\omega)$ is lower semicontinuous. Also since $F(x,\cdot)$ is measurable, in order to verify that $f(x)$ is well defined we only need to check that either $E[F(x,\omega)_+]$ or $E[(-F(x,\omega))_+]$ is finite.

By Theorem 9 and Proposition 26 we obtain the following result.

Theorem 28. *Suppose that:* (i) *the functions* $q(y,\omega)$ *and* $g_i(y,\omega)$, $t_i(x,\omega)$, $i=1,\ldots,m$, *are random lower semicontinuous,* (ii) *for a.e.* $\omega\in\Omega$ *the problem* (2.51) *is convex,* (iii) *for a.e.* $\omega\in\Omega$ *the optimal value function* $\vartheta(\cdot,\omega)$ *is lower semicontinuous,* (iv) *for a.e.* $\omega\in\Omega$ *the functions* $t_i(\cdot,\omega)$, $i=1,\ldots,m$, *are differentiable and the regularity condition* (2.55) *holds,* (v) *the expectation function* $f(x)$ *is proper and its domain has a nonempty interior. Then for any* $x\in\operatorname{dom} f$,

$$\partial f(x)=\int_\Omega \nabla T(x,\omega)^T\mathcal{D}(T(x,\omega),\omega)\,\mathrm{d}P(\omega)+N_{\operatorname{dom} f}(x). \tag{2.58}$$

Proof. It follows from assumptions (i)–(iii) that $F(x,\omega)$ is random lower semicontinuous. If $\vartheta(\chi,\omega)<+\infty$, then $\vartheta(\cdot,\omega)$ is lower semicontinuous at χ iff there is no duality gap between problems (2.51) and (2.53). Formula (2.58) follows then from the corresponding formula of Theorem 9 and formula (2.56) of Proposition 26. □

Under the assumptions of the above theorem we have that $f(\cdot)$ is differentiable at a point x iff $x\in\operatorname{int}(\operatorname{dom} f)$ and $\mathcal{D}(T(x,\omega),\omega)$ is a singleton w.p.1.

The above analysis can be applied to the second stage problem of form (2.35) with the function $f_2(\cdot,\omega)$ being convex (not necessarily polyhedral) for a.e. $\omega\in\Omega$. The dual of (2.35) can be still written in the form (2.38) However, in the non-polyhedral case one needs some additional conditions in order to ensure that there is no duality gap between the (primal) problem (2.35) and its dual (2.38). For example, we have that if, for a given (x,ω), the optimal value of (2.35) is finite, then there is no duality gap between (2.35) and (2.38) and the dual problem (2.38) has a nonempty and bounded set of optimal solutions iff the following condition holds

$$h(\omega)-T(\omega)x\in\operatorname{int}\{W(\omega)[\operatorname{dom} f_2(\cdot,w)]\}. \tag{2.59}$$

The above condition means that for small perturbations of $\chi=h(\omega)-T(\omega)x$ the corresponding (primal) problem remains feasible.

3 Multistage models

Consider the linear multistage problem

$$\begin{array}{lccccccccc}
\operatorname{Min}\mathbb{E}[& c_1^Tx_1 & + & c_2^Tx_2 & + & c_3^Tx_3 & + & \ldots & + & c_T^Tx_T] \\
\text{s.t.} & A_{11}x_1 & & & & & & & & = b_1, \\
& A_{21}x_1 & + & A_{22}x_2 & & & & & & = b_2, \\
& & & A_{32}x_2 & + & A_{33}x_3 & & & & = b_3, \\
& \ldots\ldots & & & & & & & & \ldots\ldots \\
& & & & & & & A_{T,T-1}x_{T-1} & + & A_{TT}x_T = b_T, \\
& x_1\geq 0, & & x_2\geq 0, & & x_3\geq 0, & & \ldots & & x_T\geq 0.
\end{array} \tag{3.1}$$

In this problem $x_1 \in \mathbb{R}^{n_1}, \ldots, x_T \in \mathbb{R}^{n_T}$ are the parts of the decision vector corresponding to stages $1, \ldots, T$, and the random variables associated with period t are $\xi_t := (c_t, A_{t,t-1}, A_{tt}, b_t)$. Each x_t is allowed to depend on $\xi_{[1,t]} := (\xi_1, \ldots, \xi_t)$, but not on future observations $\xi_{t+1}, \ldots, \xi_T$. That is, $x_t = x_t(\xi_{[1,t]})$ is viewed as a function of $(\xi_1, \ldots, \xi_t)$, and the minimization is performed over appropriate functional spaces. In particular, x_1 depends only on ξ_1 which is deterministic, and hence x_1 is deterministic. The constraints of (3.1) are assumed to hold for almost every realization of the random data $\xi = (\xi_1, \ldots, \xi_T)$. If the number of scenarios is finite, i.e., the distribution of ξ has a finite support, then problem (3.1) can be written as a large linear programming problem. See Chapter 1 for additional motivation and discussion of multistage models.

We discuss in this section a slightly more general structure, a *polyhedral* multistage model, which is formulated as follows:

$$
\begin{aligned}
\operatorname{Min}\ & \mathbb{E}\,[f_1(x_1,\xi_1) + f_2(x_2,\xi_2) + f_3(x_3,\xi_3) + \ldots + f_T(x_T,\xi_T)] \\
\text{s.t.}\ & A_{11}(\xi_1)x_1 = b_1(\xi_1), \\
& A_{21}(\xi_2)x_1 + A_{22}(\xi_2)x_2 = b_2(\xi_2), \\
& A_{32}(\xi_3)x_2 + A_{33}(\xi_3)x_3 = b_3(\xi_3), \\
& \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
& A_{T,T-1}(\xi_T)x_{T-1} + A_{TT}(\xi_T)x_T = b_T(\xi_T).
\end{aligned}
\tag{3.2}
$$

Here $\xi_1, \ldots, \xi_T$ is a vector valued random process associated with stages $1, \ldots, T$, and the objective parts $f_t(x_t, \xi_t)$, $t = 1, \ldots, T$, associated with the successive stages, are assumed to be *random polyhedral* functions. Random polyhedral functions were introduced in Section 2.3, see equation (2.37) in particular. Note a slight difference in notation here, it is explicitly assumed in (3.2) that all random data are a function of the process $\xi_1, \ldots, \xi_T$, which include random variables defining the (polyhedral) objective functions and the constraints. As in the linear multistage model (3.1), each x_t may only depend on $\xi_{[1,t]}$ but not on future observations, i.e., $x_t = x_t(\xi_{[1,t]})$ is a function of $\xi_{[1,t]}$, and the minimization is performed over appropriate functional spaces. Since ξ_1 becomes known before the first decision is made, we may assume that f_1 depends on x_1 only, but for the uniformity of notation we keep ξ_1 in the formulation of the problem.

Similarly to the two-stage case, every polyhedral multistage problem (with a finite number of scenarios) can be converted into a linear multistage problem by adding new variables and constraints. The form (3.2), though, is more convenient to analyze. So in the remainder of this section we deal with model (3.2). We denote $\xi := \xi_{[1,T]} = (\xi_1, \ldots, \xi_T)$ the random data of the considered problem.

Definition 29. A sequence of mappings $x_t(\xi)$, $t = 1, \ldots, T$, is called an *implementable policy* if each $x_t(\cdot)$ is a function of the history $\xi_{[1,t]}$ of the process. An implementable policy $x_t(\xi_{[1,t]})$, $t = 1, \ldots, T$, is said to be a *feasible policy* if it satisfies all constraints of (3.2) and $x_t(\xi_{[1,t]}) \in \mathrm{dom} f_t(\cdot, \xi_t)$, $t = 1, \ldots, T$, for a.e. ξ.

Let us denote, as in Chapter 1, by $Q_t(x_{t-1}, \xi_{[1,t]})$ the optimal value of the following problem (the *cost-to-go*)

$$\begin{array}{llllll}
\text{Min}\,\mathbb{E}[\, f_t(x_t,\xi_t) + f_{t+1}(x_{t+1},\xi_{t+1}) & + & \cdots & & + f_T(x_T,\xi_T) | \xi_{[1,t]}] & \\
\text{s.t.}\quad A_{t,t-1}(\xi_t)x_{t-1} + A_{tt}(\xi_t)x_t & & & & & = b_t(\xi_t), \\
\qquad\qquad A_{t+1,t}(\xi_{t+1})x_t & + & A_{t+1,t+1}(\xi_{t+1})x_{t+1} & & & = b_{t+1}(\xi_{t+1}), \\
\cdots\cdots\cdots\cdots & & & & & \\
 & & & A_{T,T-1}(\xi_T)x_{T-1} & + A_{TT}(\xi_T)x_T & = b_T(\xi_T).
\end{array} \tag{3.3}$$

In the above problem values of x_{t-1} and $\xi_1, \ldots, \xi_t$ are assumed to be known, and hence the optimal value of (3.3) is a function of these values. Problem (3.3) is viewed as a multistage problem with the first period starting at time t, and depending on $x_{t-1} \in \mathbb{R}^{n_{t-1}}$ through the first equation constraint.

As outlined in Chapter 1, functions $Q_t(x_{t-1}, \xi_{[1,t]})$ satisfy, for a.e. ξ, the following *dynamic programming* equation:

$$Q_t(x_{t-1}, \xi_{[1,t]}) = \inf_{x_t \in \mathbb{R}^{n_t}} \left\{ \varphi_t(x_t, \xi_{[1,t]}) \colon A_{t,t-1}(\xi_t)x_{t-1} + A_{tt}(\xi_t)x_t = b_t \right\}, \tag{3.4}$$

where

$$\varphi_t(x_t, \xi_{[1,t]}) := f_t(x_t, \xi_t) + \mathbb{E}\left[Q_{t+1}(x_t, \xi_{[1,t+1]}) | \, \xi_{[1,t]} \right]. \tag{3.5}$$

In the remainder of this section we focus our attention on distributions with a *finite* support of the random data vector ξ. Note that since it is assumed that the problem is polyhedral and the distribution of $\xi(\omega)$ has a finite support, the functions $\varphi_t(\cdot, \cdot)$ are random polyhedral. Let us analyze the Lagrangian

$$L_t(x_t, \pi_t) := \varphi_t(x_t, \xi_{[1,t]}) + \pi_t^T (b_t(\xi_t) - A_{t,t-1}(\xi_t)x_{t-1} - A_{tt}(\xi_t)x_t)$$

of problem (3.4) and the dual functional

$$D_t(\pi_t) := \inf_{x_t \in \mathbb{R}^{n_t}} \left\{ L_t(x_t, \pi_t) = \varphi_t(x_t) - \pi_t^T A_{tt} x_t + \pi_t^T (b_t - A_{t,t-1} x_{t-1}) \right\}.$$

We omit for brevity the arguments $\xi_{[1,t]}$ in these expressions. It follows that we can write the dual of the problem (3.4) as follows

$$\operatorname*{Max}_{\pi_t} \left\{ D_t(\pi_t) = -\varphi_t^*\left(A_{tt}^T \pi_t\right) + \pi_t^T (b_t - A_{t,t-1} x_{t-1}) \right\}, \tag{3.6}$$

where φ_t^* is the conjugate of φ_t. We deal here with polyhedral problems, so by the duality theory of linear programming, if $Q_t(x_{t-1}, \xi_{[1,t]}) < +\infty$, then there is no duality gap between problems (3.4) and (3.6), and hence

$$Q_t(x_{t-1}, \xi_{[1,t]}) = \sup_{\pi_t} \left[-\varphi_t^*\left(A_{tt}^T \pi_t\right) + \pi_t^T (b_t - A_{t,t-1} x_{t-1}) \right]. \tag{3.7}$$

Moreover, if $Q_t(x_{t-1}, \xi_{[1,t]})$ is finite, then both problems (3.4) and (3.6) have nonempty sets of optimal solutions. Let us denote, as before, by $\mathcal{D}_t(x_{t-1}, \xi_{[1,t]})$ the set of optimal solutions of the dual problem (3.6). We then have an analogue of Proposition 21.

Proposition 30. *For every $t = 2, \ldots, T$ the function $Q_t(\cdot, \xi_{[1,t]})$ is a convex polyhedral function. Moreover, $Q_t(\cdot, \xi_{[1,t]})$ is subdifferentiable at every x_{t-1}, at which $Q_t(x_{t-1}, \xi_{[1,t]})$ is finite, and*

$$\partial Q_t(x_{t-1}, \xi_{[1,t]}) = -A_{t,t-1}^T(\xi_t) \mathcal{D}_t(x_{t-1}, \xi_{[1,t]}). \tag{3.8}$$

Proof. The assertion is true for $t = T$ by Proposition 21. Suppose now that $t < T$ and the assertion holds for $t + 1$. Since the distribution of ξ_{t+1} is discrete, the function (3.5) is a convex polyhedral function, as a sum of finitely many convex polyhedral functions. Consequently, Proposition 21 applies to problem (3.4) and our assertion is true for all t. □

Identically to the two stage case, the set of maximizers in (3.6), denoted $\mathcal{D}_t(x_{t-1}, \xi_{[1,t]})$, is a convex closed polyhedron. Two cases are possible. If it is bounded, then it is the convex hull of its finitely many vertices, and $Q_t(\cdot, \xi_{[1,t]})$ is finite around x_{t-1}. If $\mathcal{D}_t(x_{t-1}, \xi_{[1,t]})$ is unbounded, then its recession cone (which is polyhedral) is the normal cone to the domain of $Q_t(\cdot, \xi_{[1,t]})$ at x_{t-1}.

Example 31 (Trucking). Let us return to the trucking example (Example 9) from Chapter 1. Let us develop the dynamic programming equations for

this model. We have here that $Q_t(r_{t-1}, D_{[1,t]})$ is equal to the optimal value of the following problem

$$\begin{aligned} \operatorname*{Max}_{y_t, z_t, r_t} & \left\{ \sum_{i,j=1}^{n} (q_{ij} z_{ijt} - c_{ij} y_{ijt}) + \mathbb{E}\big[Q_{t+1}\big(r_t, D_{[1,t+1]}\big) \mid D_{[1,t]}\big] \right\} \\ \text{s.t.} \quad & z_{ijt} \le D_{ijt}, \quad i,j = 1, \ldots, n, \\ & z_{ijt} \le y_{ijt}, \quad i,j = 1, \ldots, n, \\ & r_{i,t-1} + \sum_{k=1}^{n} y_{kit} - \sum_{j=1}^{n} y_{ijt} = r_{it}, \quad i = 1, \ldots, n, \\ & r_t \ge 0, \ y_t \ge 0, \ z_t \ge 0. \end{aligned} \tag{3.9}$$

We used here the fact that r_t is a sufficient state vector for this problem.

Let π_{it} denote the Lagrange multipliers associated with the state constraints in (3.9) and let $\Pi_t(r_{t-1}, D_{[1,t]})$ be the set of the optimal values of these multipliers. Then we know that the function $Q_t(\cdot, D_{[1,t]})$ is concave (we have a maximization problem here) and its superdifferential[7] is equal to

$$\partial Q_t(r_{t-1}, D_{[1,t]}) = \Pi_t(r_{t-1}, D_{[1,t]}). \tag{3.10}$$

4 Optimality conditions, basic case

In this section we discuss optimization problems of the form

$$\operatorname*{Min}_{x \in X} \{ f(x) := \mathbb{E}[F(x, \omega)] \}, \tag{4.1}$$

where $F: \mathbb{R}^n \times \Omega \to \overline{\mathbb{R}}$ is an integrand and X is a nonempty subset of $\mathbb{R}^n$. We assume that the expectation function $f(\cdot)$ is well defined on $\mathbb{R}^n$. Let us recall that from the theoretical point of view the feasibility constraint $x \in X$ can be absorbed into the objective function by defining[8] $\overline{F}(x, \omega) := F(x, \omega) + i_X(x)$, i.e.,

$$\overline{F}(x, \omega) = \begin{cases} F(x, \omega), & \text{if } x \in X, \\ +\infty, & \text{if } x \notin X. \end{cases}$$

[7] Since we deal here with a concave rather than convex function, we call $\partial Q_t(r_{t-1}, D_{[1,t]})$ the superdifferential rather than subdifferential.

[8] Recall that $i_X(\cdot)$ denotes the indicator function of the set X.

The optimization problem (4.1) can be formulated then in the form

$$\operatorname*{Min}_{x\in\mathbb{R}^n} \{\overline{f}(x) := \mathbb{E}[\overline{F}(x,\omega)]\}. \tag{4.2}$$

Clearly it follows that $\overline{f}(x) = f(x) + i_X(x)$, and hence $\operatorname{dom}\overline{f} = X \cap \operatorname{dom} f$. Note that the feasible set of the problem (4.1) is given by the intersection of the set X and $\operatorname{dom} f$. This set coincides, of course, with the feasible set of the problem (4.2) given by $\operatorname{dom}\overline{f}$. In this section we discuss the case where the set X is convex and the integrand $F(\cdot,\omega)$ is convex for a.e. $\omega \in \Omega$, and hence the expectation function $f(\cdot)$ is also convex.

In the following proposition we present necessary and sufficient conditions for a feasible point x_0 to be an optimal solution of the problem (4.1). In order to ensure necessity of these conditions we need a regularity assumption (constraint qualification). A simple constraint qualification of such type is the following:

$$\operatorname{ri}(X) \cap \operatorname{ri}(\operatorname{dom} f) \neq \emptyset, \tag{4.3}$$

i.e., there exists a point $\overline{x} \in \operatorname{ri}(X)$ belonging to the relative interior of the domain of f. Note that if a point $x \in X$, then any neighborhood of x has a nonempty intersection with $\operatorname{ri}(X)$. Therefore, if the domain of f has a nonempty interior, and hence $\operatorname{ri}(\operatorname{dom} f) = \operatorname{int}(\operatorname{dom} f)$, then the above constraint qualification (4.3) is equivalent to the following

$$X \cap \operatorname{int}(\operatorname{dom} f) \neq \emptyset. \tag{4.4}$$

Note also that if f is finite in at least one point of the interior of its domain, then f is continuous and subdifferentiable at that point, and hence is proper.

Proposition 32. *Suppose that the set X and the function f are convex. Consider a point $x_0 \in X$ such that $f(x_0)$ is finite. Then x_0 is an optimal solution of problem (4.1) if the following condition holds:*

$$0 \in \partial f(x_0) + N_X(x_0). \tag{4.5}$$

Moreover, if f is proper and the constraint qualification (4.3) is satisfied, then condition (4.5) is also necessary for x_0 to be an optimal solution of the problem (4.1).

Proof. We have here that $x_0 \in \operatorname{dom}\overline{f}$, where $\overline{f}(x) := f(x) + i_X(x)$. It follows immediately from the definition of subdifferentials that x_0 is an optimal solution of the problem (4.2) iff $0 \in \partial\overline{f}(x_0)$. Since the set X is nonempty, the indicator function i_X is proper. If condition (4.5) holds, then $\partial f(x_0)$ is nonempty,

and hence f is a proper function. It follows then by the Moreau–Rockafellar Theorem (Theorem 50) that $\partial \overline{f}(x_0)$ includes the set $\partial f(x_0) + \partial i_X(x_0)$. Also we have that $\partial i_X(x_0)$ coincides with $N_X(x_0)$. Therefore, we obtain that $\partial f(x_0) + N_X(x_0) \subset \partial \overline{f}(x_0)$. Consequently, if condition (4.5) holds, then $0 \in \partial \overline{f}(x_0)$, and hence x_0 is an optimal solution of (4.2). Moreover, under the constraint qualification (4.3) and the assumption that f is proper, $\partial f(x_0) + \partial i_X(x_0)$ is equal to $\partial \overline{f}(x_0)$ by the Moreau–Rockafellar Theorem, and hence the necessity of (4.5) follows. □

The above optimality conditions can be combined, of course, with the formula for the subdifferential of the expectation function given in Theorem 9. Note that the constraint qualification condition (4.4) implies that the domain of f has a nonempty interior.

Theorem 33. *Suppose that:* (i) *the function* $F(x, \omega)$ *is random lower semicontinuous,* (ii) *for a.e.* $\omega \in \Omega$ *the function* $F(\cdot, \omega)$ *is convex,* (iii) *the expectation function* f *is proper,* (iv) *the set* X *is convex,* (v) *the constraint qualification* (4.4) *is satisfied. Then a point* $x_0 \in X \cap \operatorname{dom} f$ *is an optimal solution of the problem* (4.1) *if and only if the following condition holds:*

$$0 \in \int_\Omega \partial F(x_0, \omega)\, \mathrm{d}P(\omega) + N_{\operatorname{dom} f}(x_0) + N_X(x_0). \tag{4.6}$$

In particular, if x_0 belongs to the interior of the domain of f, then $N_{\operatorname{dom} f}(x_0) = \{0\}$, and hence in that case the optimality condition (4.6) takes on the form

$$0 \in \int_\Omega \partial F(x_0, \omega)\, \mathrm{d}P(\omega) + N_X(x_0). \tag{4.7}$$

5 Optimality conditions for multistage models

Let us now turn to the polyhedral multistage model (3.2). We assume that the random vector $\xi = (\xi_1, \ldots, \xi_T)$ has a distribution with a *finite* support. Recall Definition 29 of a feasible policy $x_t(\xi_{[1,t]})$ of problem (3.2).

Since the distribution of ξ has a finite support, the value of the objective function of (3.2) is finite for every feasible policy. A question arises when a feasible policy is optimal. Note that since we deal with distributions with a finite support, a statement that a certain property holds for a.e. ξ is equivalent here to that this property holds for every realization of ξ. We write this simply as "for every realization $\xi \ldots$".

Theorem 34. *A feasible policy* $\hat{x}_t(\xi_{[1,t]})$, $t = 1, \ldots, T$, *is optimal for* (3.2) *if and only if, for all* $t = 1, \ldots, T$, *and every realization* ξ,

$$\hat{x}_t(\xi_{[1,t]}) \in \arg\min_{x_t \in \mathbb{R}^{n_t}} \big\{ \varphi_t(x_t, \xi_{[1,t]}) : A_{tt}(\xi_t) x_t = b_t(\xi_t) - A_{t,t-1}(\xi_t) \hat{x}_{t-1}(\xi_{[1,t-1]}) \big\}, \tag{5.1}$$

where $\varphi_t(x_t, \xi_{[1,t]})$ *is defined in* (3.5) *with the term* Q_{T+1} *omitted for* $t = T$.

Proof. The assertion is obvious for $T = 1$. Let us suppose that it is true for $T - 1$. Consider problem (3.3) for $t = 2$, where $\hat{x}_1$ is assumed to be fixed. By our assumption, the policy $\hat{x}_t(\xi_{[1,t]})$, $t = 2, \ldots, T$ is optimal for this problem if and only if relations (5.1) hold for $t = 2, \ldots, T$. On the other hand, since $Q_2(x_1, \xi_{[1,2]})$ is the optimal value of (3.3) for $t = 2$ for any $(x_1, \xi_{[1,2]})$, the first stage decision $\hat{x}_1$ is optimal for (3.2) if and only if (5.1) is true for $t = 1$. □

It follows from the above result that the multistage problem can be viewed as a nested family of stochastic optimization problems of form (4.1), and we can apply all the results derived in the preceding section.

Theorem 35. (i) *A feasible policy* $\hat{x}_t(\xi_{[1,t]})$, $t = 1, \ldots, T$, *is optimal for* (3.2) *if and only if, for all* $t = 1, \ldots, T$, *and every realization* $\xi_{[1,t]}$, *there exist multipliers* $\hat{\pi}_t(\xi_{[1,t]})$ *such that*[9]

$$0 \in \partial f_t(\hat{x}_t(\xi_{[1,t]}), \xi_t) - A_{tt}(\xi_t)^T \hat{\pi}_t(\xi_{[1,t]}) + \mathbb{E}\big[\partial Q_{t+1}(\hat{x}_t(\xi_{[1,t]}), \xi_{[1,t+1]}) \mid \xi_{[1,t]}\big]. \tag{5.2}$$

(ii) *Multipliers* $\hat{\pi}_t(\xi_{[1,t]})$ *satisfy* (5.2) *for a feasible policy* $\hat{x}_t(\xi_{[1,t]})$, $t = 1, \ldots, T$, *if and only if for every realization* $\xi_{[1,t]}$,

$$\hat{\pi}_t(\xi_{[1,t]}) \in \mathcal{D}_t(\hat{x}_{t-1}(\xi_{[1,t-1]}), \xi[1,t]), \tag{5.3}$$

where $\mathcal{D}_t(\hat{x}_{t-1}(\xi_{[1,t-1]}), \xi_{[1,t]})$ *is the set of optimal solutions of the dual problem* (3.6).

Proof. By Proposition 30 the functions in problem (5.1) are polyhedral, so the optimality conditions of Theorem 33 hold without any additional constraint qualification assumptions. Relation (5.2) follows then from Theorem 34. Relation (5.3) is the consequence of the duality theorem in convex programming applied to problem (5.1). □

[9] For $t = T$ we omit the term with Q_{T+1} in (1.2).

Proposition 30 provides us with the explicit form of the subdifferentials involved in (5.2):

$$\partial Q_{t+1}(x_t, \xi_{[1,t+1]}) = -A_{t+1,t}(\xi_{t+1})^T \mathcal{D}_{t+1}(x_t, \xi_{[1,t+1]}). \tag{5.4}$$

This allows us to reformulate part (i) of Theorem 35 as follows.

Corollary 36. *A feasible policy $\hat{x}_t(\xi_{[1,t]})$, $t = 1, \ldots, T$, is optimal for (3.2) if and only if, for all $t = 1, \ldots, T$, and every realization $\xi_{[1,t]}$, there exist multipliers $\hat{\pi}_t(\xi_{[1,t]})$ such that*[10]

$$0 \in \partial f_t(\hat{x}_t(\xi_{[1,t]}), \xi_t) - A_{tt}(\xi_t)^T \hat{\pi}_t(\xi_{[1,t]}) - \mathbb{E}\big[A_{t+1,t}(\xi_{t+1})^T \hat{\pi}_{t+1}(\xi_{[1,t+1]}) \mid \xi_{[1,t]}\big], \quad t = 1, \ldots, T. \tag{5.5}$$

Proof. Suppose that a policy $\hat{x}_t(\xi_{[1,t]})$, $t = 1, \ldots, T$, is optimal for (3.2). Consider $t = 1$. By Theorem 35, we can choose multipliers $\hat{\pi}_1 \in \mathcal{D}_1$ such that (5.2) holds for $t = 1$. Note that there is no preceding stage, so the set $\mathcal{D}_1$ is fixed. It follows from (5.2) and (5.4) that we can choose a measurable selection

$$\hat{\pi}_2(\xi_{[1,2]}) \in \mathcal{D}_2(\hat{x}_1, \xi_{[1,2]})$$

such that[11]

$$0 \in \partial f_1(\hat{x}_1(\xi_1), \xi_1) - A_{11}(\xi_1)^T \hat{\pi}_1(\xi_1) - \mathbb{E}\big[A_{2,1}(\xi_2)^T \hat{\pi}_2(\xi_{[1,2]}) \mid \xi_1\big], \tag{5.6}$$

so formula (5.5) is true for $t = 1$. By Theorem 35 (ii), the same selection $\hat{\pi}_2(\xi_{[1,2]})$ can be used in (5.2) for $t = 2$. Then there exists a measurable selection

$$\hat{\pi}_3(\xi_{[1,3]}) \in \mathcal{D}_3(\hat{x}_1, \xi_{[1,3]})$$

such that (5.5) is true for $t = 2$. Proceeding in this way we find selections (5.3) such that (5.5) holds for all $t = 1, \ldots, T$. □

It should be stressed that both the polyhedrality of the objective and the finite number of realizations of ξ are essential for Theorems 34 and 35, and for Corollary 36, because we could avoid the verification of constraint qualification conditions for problems appearing in (5.1). For the nested formulation (3.4) these conditions are difficult to ensure, in general. However, when the distribution of ξ remains finite, we can formulate the necessary and sufficient conditions for problems of form (3.2) with general convex functions.

[10] Again, for $t = T$ we omit the term with $T + 1$ in (5.5).

[11] Since ξ_1 is not random, the conditional expectation in (5.6) does not depend on ξ_1, we write it for uniformity of the notation.

Let us rewrite (3.2) in a compact form

$$\text{Min } \mathbb{E}[f(x,\xi)] \tag{5.7}$$

$$\text{s.t.} \quad Ax = b, \tag{5.8}$$

where

$$f(x,\xi) := \sum_{t=1}^{T} f_t(x_t(\xi_{[1,t]}),\xi_t),$$

A is the block matrix defining the constraints of (3.2), and b is the corresponding vector of the right hand sides. We should keep in mind that the decision vector $x = x(\cdot)$ is an *implementable policy*, that is,

$$x(\cdot) = (x_t(\xi_{[1,t]}))_{t=1,\ldots,T},$$

and each constraint of (3.2), associated with stage t, has as many realizations as there are different values of $\xi_{[1,t]}$ possible. All these numbers are finite, so (5.7)–(5.8) is a finite dimensional convex optimization problem.

Denote by X the linear manifold (affine space) defined by (5.8). It follows from Proposition 32 that a policy $\hat{x} \in X \cap \operatorname{dom} \mathbb{E}f(\cdot,\xi)$ is optimal if the following condition holds:

$$0 \in \partial\, \mathbb{E}[f(\hat{x},\xi)] + N_X(\hat{x}). \tag{5.9}$$

Moreover, if $\mathbb{E}[f(\cdot,\xi)]$ is proper and the constraint qualification[12] (4.3) is satisfied, then condition (5.9) is also necessary for $\hat{x}$ to be an optimal solution of the problem (3.2).

Since X is a linear manifold, the normal cone $N_X(x)$ is constant (i.e., does not depend on $x \in X$) and coincides with the linear space orthogonal to X. Consequently, $N_X(x)$ is equal to the set of vectors of form $A^T\lambda$, where

$$\lambda = \left(\lambda_t(\xi_{[1,t]})\right)_{t=1,\ldots,T}.$$

Let us introduce an equivalent representation of the normal cone. For each possible realization $\xi^k_{[1,t]}$ of $\xi_{[1,t]}$ we define[13]

$$\hat{\pi}_t\left(\xi^k_{[1,t]}\right) := -\left(\frac{1}{\mathbb{P}\{\xi_{[1,t]} = \xi^k_{[1,t]}\}}\right)\lambda_t\left(\xi^k_{[1,t]}\right). \tag{5.10}$$

[12] Since X is an affine space, its relative interior coincides with X.

[13] To avoid collisions of subscripts we slightly change our notation and use superscripts to denote realizations (scenarios).

It is legitimate, because $\mathbb{P}\{\xi_{[1,t]} = \xi^k_{[1,t]}\} > 0$. Then $N_X(x)$ is the set of vectors of the form

$$\big(A_{tt}(\xi_t)^T \hat{\pi}_t(\xi_{[1,t]}) + \mathbb{E}\big[A_{t+1,t}(\xi_{t+1})^T \hat{\pi}_{t+1}(\xi_{[1,t+1]})|\ \xi_{[1,t]}\big]\big)_{t=1,\ldots,T},$$

where, for uniformity, we take the convention that all the terms involving $T+1$ are 0.

Then simple manipulations show that the relation (5.9) is identical with (5.5).

Corollary 37. *Suppose that the distribution of ξ has a finite support and the functions $f_t(\cdot, \xi_t)$, $t = 1, \ldots, T$, are convex for every realization of ξ. Then for a feasible policy $\hat{x}_t(\xi_{[1,t]})$, $t = 1, \ldots, T$, to be optimal it is sufficient that for all $t = 1, \ldots, T$, and every realization $\xi_{[1,t]}$, there exist multipliers $\hat{\pi}_t(\xi_{[1,t]})$ such that the relations (5.5) are satisfied. If, in addition, the constraint qualification (4.3) holds, then conditions (5.5) are also necessary for $\hat{x}_t(\xi_{[1,t]})$ to be optimal for the problem (3.2).*

6 Duality, basic case

Let us discuss first the case where the set Ω is finite, say $\Omega = \{\omega_1, \ldots, \omega_K\}$ with corresponding probabilities $p_k > 0$, $k = 1, \ldots, K$. As it was mentioned in Chapter 1, in that case we can formulate problem (4.2) in the following equivalent form

$$\begin{aligned} &\mathrm{Min}_{x_1,\ldots,x_K,z} \sum_{k=1}^{K} p_k \overline{F}(x_k, \omega_k), \\ &\text{subject to} \quad x_k = z, \quad k = 1, \ldots, K, \end{aligned} \tag{6.1}$$

where $x_1, \ldots, x_K$ and z are n-dimensional vectors. Of course, we can eliminate z from the above problem by expressing it in terms of $x_1, \ldots, x_K$. However, it will be convenient to view z as an additional variable.

Since the probabilities p_k are assumed to be positive, the constraints of problem (6.1) can be also written as the equations $p_k(x_k - z) = 0$. This suggests the following Lagrangian for problem (6.1),

$$L(x_1, \ldots, x_K, z, \lambda_1, \ldots, \lambda_K) := \sum_{k=1}^{K} p_k \overline{F}(x_k, \omega_k) + \sum_{k=1}^{K} p_k \lambda_k^T (x_k - z), \tag{6.2}$$

where $\lambda_k \in \mathbb{R}^n$, $k = 1, \ldots, K$, are called *Lagrange multipliers*. The problem (6.1) can be represented as the min–max problem

$$\operatorname*{Min}_{x_1,\ldots,x_K,\,z} \left\{ \sup_{\lambda_1,\ldots,\lambda_K} L(x_1, \ldots, x_K, z, \lambda_1, \ldots, \lambda_K) \right\}. \tag{6.3}$$

Its dual problem is obtained by interchanging the order of the Min and Max operators. Since the infimum of the Lagrangian over z is $-\infty$ unless $\sum_{k=1}^K p_k \lambda_k = 0$, this leads to the following dual of the problem (6.1):

$$\begin{aligned} &\operatorname{Max}_{\lambda_1,\ldots,\lambda_K} \left\{ \inf_{x_1,\ldots,x_K} \sum_{k=1}^K p_k \big(\overline{F}(x_k, \omega_k) + \lambda_k^T x_k \big) \right\} \\ &\text{subject to} \quad \sum_{k=1}^K p_k \lambda_k = 0 \end{aligned} \tag{6.4}$$

Note the separable structure of the above problem,[14] that is

$$\inf_{x_1,\ldots,x_K} \sum_{k=1}^K p_k \big(\overline{F}(x_k, \omega_k) + \lambda_k^T x_k \big) = \sum_{k=1}^K p_k \left[\inf_{x_k} \big(\overline{F}(x_k, \omega_k) + \lambda_k^T x_k \big) \right]. \tag{6.5}$$

We also have that

$$\inf_{x_k} \big(\overline{F}_k(x_k) + \lambda_k^T x_k \big) = -\sup_{x_k} \big((-\lambda_k)^T x_k - \overline{F}_k(x_k) \big) = -\overline{F}_k^*(-\lambda_k),$$

where $\overline{F}_k(\cdot) := \overline{F}(\cdot, \omega_k)$ and $\overline{F}_k^*$ is the conjugate of $\overline{F}_k$. Therefore we can write the dual problem (6.4) in the form

$$\begin{aligned} &\operatorname{Max}_{\lambda_1,\ldots,\lambda_K} \left\{ -\sum_{k=1}^K p_k \overline{F}_k^*(-\lambda_k) \right\}, \\ &\text{subject to} \quad \sum_{k=1}^K p_k \lambda_k = 0. \end{aligned} \tag{6.6}$$

Problems (6.1) and (6.4) can be represented by employing the min–max and max–min operators, respectively, applied to the Lagrangian. It follows that

[14] One should be careful in writing equation (6.5) since some of the optimal values there can be $+\infty$ or $-\infty$, and adding $+\infty$ and $-\infty$ should be avoided.

the optimal value of the problem (6.1) is always greater than or equal to the the optimal value of its dual problem (6.4). It is also not difficult to see this directly. Indeed, for any $\lambda_1, \ldots, \lambda_K$ such that $\sum_{k=1}^K p_k \lambda_k = 0$ the following inequality holds

$$\inf_{x_1,\ldots,x_K} \left\{ \sum_{k=1}^K p_k \overline{F}(x_k, \omega_k) + \sum_{k=1}^K p_k \lambda_k^T x_k \right\} \leq \inf_x \left\{ \overline{f}(x) = \sum_{k=1}^K p_k \overline{F}(x, \omega_k) \right\}. \tag{6.7}$$

The above inequality is obtained by restricting the search at the left hand side to $x_k = x$, $k = 1, \ldots, K$. Since (6.7) holds for any multipliers λ_k satisfying constraint $\sum_{k=1}^K p_k \lambda_k = 0$, it follows that the optimal value of (6.4) is less than or equal to the optimal value of (4.2).

In order to ensure that problems (6.1) and (6.4) have equal optimal values, i.e., that there is no duality gap between (6.1) and (6.4), one needs a constraint qualification condition. It is possible to deal with the duality gap problem by employing various techniques of convex analysis. We outline below a particular approach which is relatively elementary and easy to extend to infinite dimensional cases.

By the min–max representation, there is no duality gap between (6.1) and (6.4) and both problems have optimal solutions iff the Lagrangian has a *saddle point*. That is, there is a feasible point $(\overline{x}_1, \ldots, \overline{z})$ of problem (6.1) and a feasible point $(\overline{\lambda}_1, \ldots, \overline{\lambda}_K)$ of (6.4) such that $L(\cdot, \ldots, \cdot, \overline{\lambda}_1, \ldots, \overline{\lambda}_K)$ attains its (unconstrained) minimum at $(\overline{x}_1, \ldots, \overline{z})$ and $L(\overline{x}_1, \ldots, \overline{z}, \cdot, \ldots, \cdot)$ attains its maximum at $(\overline{\lambda}_1, \ldots, \overline{\lambda}_K)$. Equivalently, $(\overline{x}_1, \ldots, \overline{z}, \overline{\lambda}_1, \ldots, \overline{\lambda}_K)$ is a saddle point iff

$$\overline{x}_k = \overline{z},\ k = 1, \ldots, K; \quad \sum_{k=1}^K p_k \overline{\lambda}_k = 0, \tag{6.8}$$

and $(\overline{x}_1, \ldots, \overline{x}_K)$ is an optimal solution of the (unconstrained) problem

$$\operatorname*{Min}_{x_1,\ldots,x_K} \left\{ \sum_{k=1}^K p_k \overline{F}(x_k, \omega_k) + \sum_{k=1}^K p_k \overline{\lambda}_k^T x_k \right\}. \tag{6.9}$$

In that case $\overline{z}$ is an optimal solution of (4.2) (or, equivalently, of (4.1)) and $(\overline{\lambda}_1, \ldots, \overline{\lambda}_K)$ is an optimal solution of (6.4). Since problem (6.9) is separable, $(\overline{x}_1, \ldots, \overline{x}_K)$ is an optimal solution of problem (6.9) iff

$$\overline{x}_k \in \arg\min_{x_k} \left\{ \overline{F}(x_k, \omega_k) + \overline{\lambda}_k^T x_k \right\}, \quad k = 1, \ldots, K. \tag{6.10}$$

The functions $\overline{F}(\cdot, \omega_k)$ are assumed to be convex, so condition (6.10) holds iff $\overline{\lambda}_k \in -\partial \overline{F}(\overline{x}_k, \omega_k)$, $k = 1, \ldots, K$. Therefore $(x_0, \ldots, x_0, \overline{\lambda}_1, \ldots, \overline{\lambda}_K)$ is a saddle point of the Lagrangian iff

$$\overline{\lambda}_k \in -\partial \overline{F}(x_0, \omega_k), \ k = 1, \ldots, K, \quad \text{and} \quad \sum_{k=1}^{K} p_k \overline{\lambda}_k = 0. \tag{6.11}$$

Theorem 38. *Suppose that the set X and the functions $F(\cdot, \omega_k)$, $k = 1, \ldots, K$, are convex. Then the following holds.* (i) *Points x_0 and $(\overline{\lambda}_1, \ldots, \overline{\lambda}_K)$ are optimal solutions of problems (4.1) and (6.4), respectively, and there is no duality gap between these problems if and only if condition (6.11) is satisfied.* (ii) *Problems (4.1) and (6.4) have optimal solutions and there is no duality gap between these problems if and only if there exists a point x_0 such that*

$$0 \in \sum_{k=1}^{K} p_k \partial \overline{F}(x_0, \omega_k). \tag{6.12}$$

(iii) *Problems (4.1) and (6.4) have optimal solutions and there is no duality gap between these problems if problem* (4.1) *has an optimal solution x_0, the function f is proper and the regularity condition (4.3) is satisfied.*

Proof. By the discussion preceding the theorem, $(x_0, \ldots, x_0, \overline{\lambda}_1, \ldots, \overline{\lambda}_K)$ is a saddle point of the Lagrangian iff condition (6.11) is satisfied. This proves assertion (i). Since it is assumed that all p_k are positive, condition (6.12) is equivalent to existence of $\overline{\lambda}_k$ satisfying (6.11). Note that (6.12) implies that all functions $\overline{F}(\cdot, \omega_k)$ are subdifferentiable at x_0 and hence $\overline{F}(x_0, \omega_k)$ is finite, and consequently $\overline{f}(x_0)$ is finite. Assertion (ii) then follows. By the Moreau–Rockafellar Theorem, if the regularity condition (4.3) is satisfied and $x_0 \in \operatorname{dom} \overline{f}$, then

$$\partial \overline{f}(x_0) = \sum_{k=1}^{K} p_k \partial \overline{F}(x_0, \omega_k). \tag{6.13}$$

Moreover, if x_0 is an optimal solution of (4.1), then $0 \in \partial \overline{f}(x_0)$, and hence assertion (iii) follows from (ii). □

The above results can be formulated in the following form. Consider the problem

$$\operatorname*{Min}_{x(\omega) \in X} \mathbb{E}\big[F(x(\omega), \omega) + \lambda(\omega)^T x(\omega)\big]. \tag{6.14}$$

Here $x(\cdot)\colon \Omega \to \mathbb{R}^n$ is a mapping, the constraint $x(\omega) \in X$ is assumed to hold for every $\omega \in \Omega$, and $\lambda(\cdot)\colon \Omega \to \mathbb{R}^n$ is such that $\mathbb{E}[\lambda(\omega)] = 0$. Since it is assumed that Ω is finite, mappings $x(\omega)$ and $\lambda(\omega)$ can be identified with vectors $(x_1, \ldots, x_K)$ and $(\lambda_1, \ldots, \lambda_K)$, respectively. Therefore, problem (6.14) is the same as problem (6.9) (for $\lambda_k = \overline{\lambda}_k$). By Theorem 38, if the set X and the functions $F(\cdot, \omega_k)$, $k = 1, \ldots, K$, are convex, x_0 is an optimal solution of the problem (4.1), the function f is proper and the regularity condition (4.3) is satisfied, then there exists $\lambda(\omega)$ such that $\mathbb{E}[\lambda(\omega)] = 0$ and $\overline{x}(\omega) \equiv x_0$ is an optimal solution of (6.14).

We can also investigate dual problems (6.1) and (6.4) in the framework of conjugate duality. Let $v(\eta)$ be the optimal value of the problem

$$\begin{aligned} &\mathrm{Min}_{x_1,\ldots,x_K,z} \sum_{k=1}^{K} p_k \overline{F}(x_k, \omega_k), \\ &\text{subject to} \quad p_k(x_k - z) + \eta_k = 0, \quad k = 1, \ldots, K, \end{aligned} \tag{6.15}$$

where $\eta = (\eta_1, \ldots, \eta_K)$ and $\eta_k \in \mathbb{R}^n$, $k = 1, \ldots, K$, are viewed as parameters giving perturbations of problem (6.1). Clearly, for $\eta = 0$ problem (6.15) coincides with problem (6.1), and $v(0)$ is the optimal value of (6.1). It is straightforward to verify that the function $v(\eta)$ is convex and its conjugate is

$$\begin{aligned} v^*(\lambda) &= \sup\left\{ \sum_{k=1}^{K} \lambda_k^T y_k - \sum_{k=1}^{K} p_k \overline{F}(x_k, \omega_k) : p_k(x_k - z) + \eta_k = 0, k = 1, \ldots, K \right\} \\ &= \sup\left\{ -\sum_{k=1}^{K} p_k \lambda_k^T (x_k - z) - \sum_{k=1}^{K} p_k \overline{F}(x_k, \omega_k) \right\} \\ &= -\inf\left\{ \sum_{k=1}^{K} p_k \overline{F}(x_k, \omega_k) + \sum_{k=1}^{K} p_k \lambda_k^T (x_k - z) \right\}. \end{aligned}$$

Consequently the dual problem (6.4) coincides with the problem of maximization of $-v^*(\lambda)$. That is, the optimal value of the dual problem is equal to $v^{**}(0)$. By the theory of conjugate duality (see Section 9.2 in the Appendix) we have the following results.

Proposition 39. *Suppose that the set X and the functions $F(\cdot, \omega_k)$, $k = 1, \ldots, K$, are convex. Then the following holds.* (i) *Suppose that the problem (6.1) is subconsistent. Then there is no duality gap between problems (4.1) and (6.4) if and only if the optimal value function $v(\eta)$ is lower semicontinuous at $\eta = 0$.*

(ii) *There is no duality gap between problems* (4.1) and (6.4) *and the dual problem* (6.4) *has a nonempty set of optimal solutions if and only if* $v(\eta)$ *is subdifferentiable at* $\eta = 0$. (iii) *There is no duality gap between problems* (4.1) and (6.4) *and the dual problem* (6.4) *has a nonempty and bounded set of optimal solutions if and only if* $v(0)$ *is finite and* $0 \in \mathrm{int}(\mathrm{dom}\, v)$. (iv) *If the dual problem* (6.4) *has a nonempty and bounded set of optimal solutions, then* $v(\eta)$ *is continuous at* $\eta = 0$ *and there is no duality gap between problems* (4.1) and (6.4).

Remark 40. In the case of polyhedral problem (2.34)–(2.35) we have that $\overline{F}(x, \omega) = f_1(x) + Q(x, \omega)$, where $Q(x, \omega)$ is the optimal value of the second stage problem (2.35). In that case problems (4.1) and (6.4) form a pair of dual linear programming problems, and hence there is no duality gap between these problems unless both of them are infeasible. Moreover, if the (common) optimal value of the primal and dual problems is finite, then both problems have nonempty sets of optimal solutions. That is, in the polyhedral case with a finite number of scenarios, there is no need for additional regularity conditions for the strong duality relation to hold.

Example 41 (Betting on Horses). There are n horses in a race. For every horse i we know the probability p_i that it wins and the amount s_i that the rest of the public is betting on it. The track keeps a certain proportion $C \in (0, 1)$ of the total amount bet and distributes the rest among the public in proportion to the amounts bet on the winning horse. We want to place bets totaling b dollars to maximize the expected net return.

Let us denote by x_i the amount bet on horse i. There are n scenarios $\omega_1, \ldots, \omega_n$ in this problem, with scenario ω_k representing the event that horse k wins the race. Then that the amount $F_k(x) = F(x, \omega_k)$ gained in scenario k is

$$F_k(x) = Ax_k/(x_k + s_k),$$

where $A := (1 - C)(b + \sum_{i=1}^n s_i)$ is the total sum to be split. We can now write the corresponding optimization problem as follows:

$$\operatorname*{Max}_x \; A \sum_{k=1}^n \frac{p_k x_k}{x_k + s_k} \tag{6.16}$$

$$\text{s.t.} \quad \sum_{i=1}^n x_i = b, \tag{6.17}$$

$$x \geq 0. \tag{6.18}$$

In the extended formulation (6.1) we have separated decision vectors $x^k = (x_1^k, \ldots, x_n^k)$ for each scenario $k = 1, \ldots, n$. The problem takes on the form

$$\text{Max } A \sum_{k=1}^{n} \frac{p_k x_k^k}{x_k^k + s_k} \tag{6.19}$$

$$\text{s.t.} \quad \sum_{i=1}^{n} x_i^k = b, \quad k = 1, \ldots, n, \tag{6.20}$$

$$x^k \geq 0, \quad k = 1, \ldots, n, \tag{6.21}$$

$$x^k = z, \quad k = 1, \ldots, n. \tag{6.22}$$

Without the nonaticipativity constraints (6.22) this would mean the comfortable situation of knowing the winning horse before placing the bet. The optimal solution would be then, of course, $x_k^k = b$, and $x_i^k = 0$ for $i \neq k$.

Note that $F_k(x) = A(1 - s_k/(x_k + s_k))$, and hence functions F_k are concave on $\mathbb{R}^n_+$, and therefore, since we deal here with a maximization problem, (6.19)–(6.22) is a convex problem. Clearly its feasible set is nonempty and bounded, and hence it has an optimal solution. Since, in fact, functions F_k are strictly concave, problem (6.19)–(6.22) possesses a unique optimal solution. Of course, similar remarks apply to problem (6.16)–(6.18) as well. Moreover, under small perturbations of the nonanticipativity constraints (6.22) it remains feasible, and hence by Proposition 39 (iii) we have that the dual of (6.19)–(6.22) has a nonempty and bounded set of optimal solutions and there is no duality gap between these problems. In the present case we will be able to write these optimal solutions explicitly.

Suppose that a 'friend' with inside information offers us 'protection' against the uncertainty inherent in betting on horses. He offers to provide us with a table of payments λ_i^k, $k, i = 1, \ldots, n$, such that in the event that horse k wins we shall pay him the amount $\lambda_i^k x_i$ for each horse i, proportionally to the amount x_i bet on this horse. The payments λ_i^k can be negative (in which case he pays us), and in fact

$$\sum_{k=1}^{N} p_k \lambda_i^k = 0, \quad i = 1, \ldots, n, \tag{6.23}$$

so that the expected cost of the deal is zero. If we shall enter the deal, he will tell us which horse is going to win.

It is intuitively clear that getting such information at the average cost zero should give us a certain advantage. And, indeed, in that case we could optimize our behavior by solving for $k = 1, \ldots, K$ the problems

$$\operatorname*{Max}_{x_k} \; A\left(\frac{x_k^k}{x_k^k + s_k}\right) - (\lambda^k)^T x^k \tag{6.24}$$

$$\text{s.t.} \quad \sum_{i=1}^{n} x_i^k = b, \tag{6.25}$$

$$x^k \geq 0. \tag{6.26}$$

The expectation of the optimal values of the above problems (6.24)–(6.26) represents the value of the dual problem for the agreed choice of the costs (i.e., multipliers) λ_i^k. Consequently, by the weak duality we have that this expectation is always greater than or equal to the optimal value of the problem (6.16)–(6.18). That is, in such a deal we expect on average to gain an additional nonnegative amount of money. In particular, if all λ_i^k are zeros, i.e., the information is given to us free of charge, then the expected value of our additional gain is equal to the *expected value of perfect information* (see Chapter 1 for a discussion of EVPI).

Suppose now that our 'friend' makes his own optimization by choosing λ_i^k to minimize our expected gain. Thus he minimizes the dual value subject to the constraints (6.23), i.e., the multipliers λ_i^k form an optimal solution of the dual problem. It turns out that in this case the deal will not help us at all. Since in the present example there is no duality gap between the optimal values of the primal and dual problems, for the optimal choice of λ_i^k the expected value of the additional gain is zero, and the optimal solution of the original problem (6.16)–(6.18) provides the optimal solution of every problem (6.24)–(6.26), $k = 1, \ldots, n$, of the well-informed. By the strict concavity, the opposite is true, too: the optimal solutions of (6.24)–(6.26), $k = 1, \ldots, n$, form the optimal solution of (6.16)–(6.18).

In other words, knowing the winning horse is not harmful, so the expected optimal value of problems (6.24)–(6.26), in view of (6.23), is at least as good as the optimal value of our original problem. If the multipliers λ_i^k are chosen in the optimal way, these values become equal, and this is the essence of duality in this case.

We can find the optimal solution $\hat{x}$ and the payment table λ by the following argument. Denoting by μ the multiplier associated with the budget constraint (6.17), we see that the optimal solution has the form

$$\hat{x}_i = \max\left(0, \sqrt{Ap_i s_i/\mu} - s_i\right), \quad i = 1, \ldots, n.$$

Ordering the horses (and scenarios) in such a way that $p_1/s_1 \geq p_2/s_2 \geq \ldots \geq p_n/s_n$ we see that there must exist l such that

$$\hat{x}_i = \begin{cases} \sqrt{Ap_i s_i/\mu} - s_i, & i = 1, \ldots, l, \\ 0, & \text{otherwise.} \end{cases}$$

Since the budget b must be used, we can find l as the smallest integer for which

$$\sqrt{\frac{p_l}{s_l}} > \frac{\sum_{i=1}^{l} \sqrt{p_i s_i}}{b + \sum_{i=1}^{l} s_i} \geq \sqrt{\frac{p_{l+1}}{s_{l+1}}}.$$

Note that the left inequality holds for $l = 1$. If such an integer does not exist, we set $l = n$. In any case

$$\mu = \frac{A\left(\sum_{i=1}^{l} \sqrt{p_i s_i}\right)^2}{\left(b + \sum_{i=1}^{l} s_i\right)^2}$$

We leave to the reader the elementary manipulations that support these results. Then we get

$$\lambda_i^k = \begin{cases} -\mu & \text{if } i \neq k, \\ -\mu + \mu/p_k & \text{if } i = k. \end{cases}$$

It turns out that the (optimal) payment table λ_i^k has a very special form. Our friend pays us up front the amount μb. We have to bet the amount b the way we wish. In return we promise to pay him $\mu x_{k^*}/p_{k^*}$, where k^* is the winning horse. If we enter this deal, he will tell us what k^* will be. It will not help us at all. Our bets will be the same as if we did not know it.

Let us consider now the general case where the probability space Ω is not necessarily finite. Recall that the constraint $x \in X$ can be absorbed into the objective function, and hence problem (4.1) can be written in form (4.2). Problem (4.2), in turn, can be formulated in the following equivalent form

$$\begin{aligned} &\operatorname{Min}_{x(\cdot)\in\mathcal{M},\, z\in\mathbb{R}^n} \mathbb{E}\big[\overline{F}(x(\omega), \omega)\big], \\ &\text{subject to} \qquad x(\omega) = z, \text{ a.e. } \omega \in \Omega. \end{aligned} \tag{6.27}$$

Since, in fact, optimization in the above problem is performed over constant mappings $x(\omega)$, the set $\mathcal{M}$ can be any space of measurable mappings $x\colon \Omega \to \mathbb{R}^n$ such that the expectation in (6.27) is well defined. The choice of

the space $\mathcal{M}$ affects, however, the corresponding duality relations. It appears natural to take $\mathcal{M}$ to be the space of all measurable mappings $x(\omega)$. Unfortunately, this may create problems with the definition of the expectation of the functions $F(x(\omega),\omega)$ and $\lambda(\omega)^T x(\omega)$. Therefore, it is convenient to restrict the space $\mathcal{M}$ to essentially bounded[15] mappings. That is, we assume that $\mathcal{M} := \mathcal{L}_\infty^n(\Omega,\mathcal{F},P)$, where $\mathcal{L}_\infty^n(\Omega,\mathcal{F},P)$ is the linear space of essentially bounded measurable mappings $x\colon \Omega \to \mathbb{R}^n$. We assume in the subsequent analysis that the expectation $\mathbb{E}[\overline{F}(x(\omega),\omega)]$ is well defined for all $x(\cdot) \in \mathcal{L}_\infty^n(\Omega,\mathcal{F},P)$. Note that if $x(\cdot) \in \mathcal{L}_\infty^n(\Omega,\mathcal{F},P)$ and[16] $\lambda(\cdot) \in \mathcal{L}_1^n(\Omega,\mathcal{F},P)$, then $\mathbb{E}[\lambda(\omega)^T x(\omega)]$ is well defined and finite. When it will not cause a confusion we will use subsequently the shortened notation $\mathcal{L}_\infty^n = \mathcal{L}_\infty^n(\Omega,\mathcal{F},P)$ and $\mathcal{L}_1^n = \mathcal{L}_1^n(\Omega,\mathcal{F},P)$.

With problem (6.27) is associated the Lagrangian

$$L(x,z,\lambda) := \mathbb{E}\big[\overline{F}(x(\omega),\omega) + \lambda(\omega)^T(x(\omega) - z)\big], \tag{6.28}$$

where $(x,z) \in \mathcal{L}_\infty^n \times \mathbb{R}^n$ and $\lambda \in \mathcal{L}_1^n$. By minimizing this Lagrangian with respect to $x(\cdot)$ and z and maximizing with respect to $\lambda(\cdot)$ we obtain the following dual of the problem (6.27):

$$\begin{aligned} &\mathrm{Max}_{\lambda(\cdot)\in\mathcal{L}_1^n}\ \big\{\inf_{x\in\mathcal{L}_\infty^n} \mathbb{E}[\overline{F}(x(\omega),\omega) + \lambda(\omega)^T x(\omega)]\big\}, \\ &\text{subject to} \quad \mathbb{E}[\lambda(\omega)] = 0. \end{aligned} \tag{6.29}$$

We have that a point $(\overline{x},\overline{z},\overline{\lambda}) \in \mathcal{L}_\infty^n \times \mathbb{R}^n \times \mathcal{L}_1^n$ is a saddle point of the Lagrangian iff

$$\overline{x}(\omega) = \overline{z}, \text{ a.e. } \omega \in \Omega, \quad \text{and} \quad \mathbb{E}\big[\overline{\lambda}(\omega)\big] = 0, \tag{6.30}$$

and

$$\overline{x} \in \arg\min\big\{\mathbb{E}[\overline{F}(x(\omega),\omega) + \overline{\lambda}(\omega)^T x(\omega)]\colon x \in \mathcal{L}_\infty^n\big\}. \tag{6.31}$$

Let us observe that condition (6.31) can be equivalently expressed as

$$\overline{z} \in \arg\min\big\{\overline{F}(z,\omega) + \overline{\lambda}(\omega)^T z\colon z \in \mathbb{R}^n\big\}, \quad \text{a.e. } \omega \in \Omega. \tag{6.32}$$

[15] A function $x(\omega)$ is called essentially bounded if there exists a constant c such that $||x(\omega)|| \le c$ for a.e. $\omega \in \Omega$.

[16] $\mathcal{L}_1^n(\Omega,\mathcal{F},P)$ denotes the linear space of measurable mappings $y\colon \Omega \to \mathbb{R}^n$ such that $\int_\Omega ||y(\omega)||\,\mathrm{d}P(\omega) < +\infty$.

Indeed, for a constant w.p.1 mapping $\overline{x}(\omega) \equiv \overline{z}$, condition (6.31) is satisfied iff for any $x(\cdot) \in \mathcal{L}_\infty^n$ the inequality

$$\overline{F}(x(\omega), \omega) + \overline{\lambda}(\omega)^T x(\omega) \geq \overline{F}(\overline{z}, \omega) + \overline{\lambda}(\omega)^T \overline{z}$$

holds w.p.1. This, in turn, is equivalent to (6.32).

Since $\overline{F}(\cdot, \omega)$ is convex w.p.1, condition (6.32) is equivalent to

$$\overline{\lambda}(\omega) \in -\partial \overline{F}(\overline{z}, \omega), \quad \text{a.e. } \omega \in \Omega. \tag{6.33}$$

Therefore, we obtain that a point $(\overline{x}, \overline{z}, \overline{\lambda}) \in \mathcal{L}_\infty^n \times \mathbb{R}^n \times \mathcal{L}_1^n$ is a saddle point of the Lagrangian iff conditions (6.30) and (6.33) hold. Suppose, further, that the function $F(x, \omega)$ is random lower semicontinuous, and hence the expectation

$$\mathbb{E}[\partial \overline{F}(\overline{z}, \omega)] = \int_\Omega \partial \overline{F}(\overline{z}, \omega)\, \mathrm{d}P(\omega)$$

is well defined. Then, by the definition of the integral of a multifunction, (6.33) and the second equation of (6.30) imply that

$$0 \in \mathbb{E}[\partial \overline{F}(\overline{z}, \omega)]. \tag{6.34}$$

Conversely, if (6.34) holds, then there exists $\overline{\lambda} \in \mathcal{L}_1^n$ satisfying (6.33) and (6.30). Therefore, the Lagrangian, given in (6.28), has a saddle point iff there exists $\overline{z} \in \mathbb{R}^n$ satisfying condition (6.34). We can formulate this results in the following form.

Proposition 42. *Suppose that the set X is convex, for a.e. $\omega \in \Omega$ the function $F(\cdot, \omega)$ is convex, and the function $F(x, \omega)$ is random lower semicontinuous. Then there is no duality gap between problems (4.1) and (6.29) and both problems have optimal solutions if and only if there exists $\overline{z} \in \mathbb{R}^n$ satisfying condition (6.34).*

Recall that the inclusion $\mathbb{E}[\partial \overline{F}(\overline{z}, \omega)] \subset \partial \overline{f}(\overline{z})$ always holds. Therefore, condition (6.34) implies that $0 \in \partial \overline{f}(\overline{z})$, and hence $\overline{z}$ is an optimal solution of (4.1). Conversely, if $\overline{z}$ is an optimal solution of (4.1), then $0 \in \partial \overline{f}(\overline{z})$, and hence if in addition $\mathbb{E}[\partial \overline{F}(\overline{z}, \omega)] = \partial \overline{f}(\overline{z})$, then (6.34) follows. Therefore, Theorem 9 and Proposition 42 imply the following result.

Theorem 43. *Suppose that:* (i) *the function $F(x, \omega)$ is random lower semicontinuous,* (ii) *for a.e. $\omega \in \Omega$ the function $F(\cdot, \omega)$ is convex,* (iii) *the set X is convex,* (iv) *problem (4.1) possesses an optimal solution x_0 such that $x_0 \in \operatorname{int}(\operatorname{dom} f)$. Then there is no duality gap between problems (4.1) and (6.29),*

the dual problem (6.29) has an optimal solution $\overline{\lambda}$*, and the constant mapping* $\overline{x}(\omega) \equiv x_0$ *is an optimal solution of the problem*

$$\operatorname*{Min}_{x(\cdot)\in\mathcal{L}^n_\infty} \mathbb{E}\big[\,\overline{F}(x(\omega),\omega) + \overline{\lambda}(\omega)^T x(\omega)\big]. \tag{6.35}$$

Proof. Since x_0 is an optimal solution of (4.1) we have that $x_0 \in X$ and $f(x_0)$ is finite. Moreover, since $x_0 \in \operatorname{int}(\operatorname{dom} f)$ and f is convex, it follows that f is proper and $N_{\operatorname{dom} f}(x_0) = \{0\}$. Therefore, it follows by Theorem 9 that $\partial f(x_0) = \mathbb{E}[\partial F(x_0,\omega)]$. Furthermore, since $x_0 \in \operatorname{int}(\operatorname{dom} f)$ we have that $\partial \overline{f}(x_0) = \partial f(x_0) + N_X(x_0)$. Consequently, $0 \in \mathbb{E}\big[\partial \overline{F}(x_0,\omega)\big]$, and hence the assertions follow by Proposition 42. □

It is also possible to investigate dual problems (4.1) and (6.29) in the framework of conjugate duality. However, since we deal here with infinite dimensional spaces like $\mathcal{L}^n_\infty(\Omega,\mathcal{F},P)$, this would require an application of functional analysis which will go beyond the scope of this book.

One can note again the separable structure of the problem (6.35). For each $\omega \in \Omega$,

$$\inf_{x\in\mathbb{R}^n}\big[\,\overline{F}(x,\omega) + \lambda^T x\big] = -\overline{F}^*(-\lambda,\omega),$$

where $\overline{F}^*(\cdot,\omega)$ is the conjugate of the function $\overline{F}(\cdot,\omega)$. For a given $\lambda \in \mathcal{L}^n_1(\Omega,\mathcal{F},P)$, denote by $\mathcal{M}_\lambda = \mathcal{M}_\lambda(\Omega,\mathcal{F},\mathbb{R}^n)$ the space of all measurable mappings $x\colon \Omega \to \mathbb{R}^n$ such that the expectation $\mathbb{E}\,[\overline{F}(x(\omega),\omega) + \lambda(\omega)^T x(\omega)]$ is well defined. By Proposition 5 of Chapter 1, the infimum with respect to $x \in \mathcal{M}_\lambda(\Omega,\mathcal{F},\mathbb{R}^n)$ can be taken inside the expected value, that is

$$\inf_{x(\cdot)\in\mathcal{M}_\lambda} \mathbb{E}\big[\,\overline{F}(x(\omega),\omega) + \lambda(\omega)^T x(\omega)\big] = \mathbb{E}\big[-\overline{F}^*(-\lambda(\omega),\omega)\big], \tag{6.36}$$

provided that $F(x,\omega)$ (and hence $\overline{F}(x,\omega)$) is random lower semicontinuous and the expectation of the minimal value is well defined. The space $\mathcal{M}_\lambda(\Omega,\mathcal{F},\mathbb{R}^n)$ in the optimization problem at the left hand side of (6.36) can be replaced by, possibly smaller, space $\mathcal{L}^n_\infty(\Omega,\mathcal{F},\mathbb{R}^n)$ if this optimization problem has an optimal (nearly optimal) solution which is essentially bounded. This happens, for example, if the set X is bounded. In that case the dual problem (6.29) can be written in the form

$$\begin{aligned} &\operatorname{Max}_{\lambda(\cdot)\in\mathcal{L}^n_1} \mathbb{E}\big[-\overline{F}^*(-\lambda(\omega),\omega)\big], \\ &\text{subject to} \quad \mathbb{E}[\lambda(\omega)] = 0. \end{aligned} \tag{6.37}$$

7 Duality for multistage stochastic programs

Let us consider again the multistage stochastic programming problem (3.2). Unless stated otherwise we assume in this section that the distribution of ξ has a finite support and the functions $f_t(\cdot,\cdot)$ are random polyhedral, i.e., the problem (3.2) is *polyhedral*.

The first approach introduces Lagrange multipliers π_t, $t=1,\ldots,T$, associated with the constraints of (3.2) and the Lagrangian

$$\begin{aligned} L(x,\pi) &:= \mathbb{E}\left\{\sum_{t=1}^{T}\Big[f_t(x_t,\xi_t)+\pi_t^T(b_t(\xi_t)-A_{tt}(\xi_t)x_t-A_{t,t-1}(\xi_t)x_{t-1})\Big]\right\} \\ &= \mathbb{E}\left\{\sum_{t=1}^{T}\Big[f_t(x_t,\xi_t)-\pi_t^T A_{tt}(\xi_t)x_t-\pi_{t+1}^T A_{t+1,t}(\xi_{t+1})x_t+\pi_t^T b_t(\xi_t)\Big]\right\}, \end{aligned} \tag{7.1}$$

with the convention that $x_0=0$ and the terms involving $T+1$ are zeros. The multipliers π_t, similarly to the decisions x_t, may depend on $\xi_{[1,t]}$, but not on $\xi_{t+1},\ldots,\xi_T$. That is, $x_t=x_t(\xi_{[1,t]})$ and $\pi_t=\pi_t(\xi_{[1,t]})$ are viewed as functions of $\xi_{[1,t]}$.

The dual functional is defined as

$$D(\pi) := \inf_{x\in\mathcal{M}} \mathbb{E}[L(x,\pi)], \tag{7.2}$$

where $\mathcal{M}$ is an appropriate functional space of allowable mappings $x(\cdot)=(x_1(\cdot),\ldots,x_T(\cdot))$. Since, for given π, the Lagrangian $L(\cdot,\pi)$ is separable in $x_t(\cdot)$, we can move the operation of minimization with respect to x_t under the conditional expectation $\mathbb{E}(\cdot|\xi_{[1,t]})$ (see Proposition 5 of Chapter 1). Therefore, we obtain

$$\begin{aligned} D(\pi) = \mathbb{E}\left\{\sum_{t=1}^{T}\inf_{x_t\in\mathbb{R}^{n_t}}\left\{f_t(x_t,\xi_t)-\pi_t^T A_{tt}(\xi_t)x_t-\mathbb{E}\big[\pi_{t+1}^T A_{t+1,t}(\xi_{t+1})|\xi_{[1,t]}\big]x_t\right\}\right\} \\ +\mathbb{E}\left\{\sum_{t=1}^{T}\pi_t^T b_t(\xi_t)\right\}. \end{aligned}$$

It follows that

$$D(\pi)=\mathbb{E}\left\{\sum_{t=1}^{T} D_t(\pi_t,\pi_{t+1},\xi_{[1,t]})\right\}+\mathbb{E}\left\{\sum_{t=1}^{T}\pi_t^T b_t(\xi_t)\right\},$$

where

$$D_t(\pi_t,\ \pi_{t+1},\ \xi_{[1,t]}) := \inf_{x_t \in \mathbb{R}^{n_t}} \big[f_t(x_t,\ \xi_t) - \big(\pi_t^T A_{tt}(\xi_t) + \mathbb{E}\big[\pi_{t+1}^T A_{t+1,t}(\xi_{t+1}) | \xi_{[1,t]}\big]\big) x_t \big]. \tag{7.3}$$

The dual problem has the form

$$\operatorname*{Max}_{\pi}\ D(\pi), \tag{7.4}$$

where the maximization is performed over such $\pi_t(\cdot)$ which depend only on $\xi_{[1,t]}$, $t = 1, \ldots, T$.

Since we assume here that there is a finite number of scenarios, the allowable mappings $x_t(\cdot)$ and $\pi_t(\cdot)$ can be identified with finite dimensional vectors. Moreover, since we deal with the polyhedral case, both primal and dual problems can be written as large linear programming problems. Therefore, the following duality result is a consequence of the general theory of linear programming.

Theorem 44. *The optimal values of problems* (3.2) and (7.4) *are equal unless both problems are infeasible. If the* (*common*) *optimal value of these problems is finite, then both problems have optimal solutions.*

If the functions $f_t(\cdot, \xi_t)$ are convex (not necessarily polyhedral), a constraint qualification like (4.3) is needed to ensure that there is no duality gap between problems (3.2) and (7.4).

The form of the dual problem is particularly simple in the case of the linear multistage problem (3.1). Indeed, let

$$f_t(x_t, \xi_t) := \begin{cases} c_t^T(\xi_t) x_t, & \text{if } x_t \geq 0, \\ +\infty, & \text{otherwise.} \end{cases}$$

Then the infimum in (7.3) is $-\infty$, unless

$$A_{tt}(\xi_t)^T \pi_t + \mathbb{E}\big[A_{t+1,t}(\xi_{t+1})^T \pi_{t+1} | \xi_{[1,t]}\big] \leq c_t(\xi_t),$$

in which case the infimum is zero. Thus the dual problem (7.4) takes on the form

$$\begin{aligned} &\operatorname*{Max}_{\pi} \ \mathbb{E}\left[\sum_{t=1}^{T} b_t(\xi_t)^T \pi_t\right] \\ &\text{s.t.} \quad A_{tt}(\xi_t)^T \pi_t + \mathbb{E}\big[A_{t+1,t}(\xi_{t+1})^T \pi_{t+1} | \xi_{[1,t]}\big] \le c_t(\xi_t), \quad t = 1, \ldots, T, \end{aligned} \tag{7.5}$$

where for the uniformity of notation we set all '$T+1$ terms' equal to 0. The multipliers π_t in problem (7.5) are restricted to depend only on $\xi_{[1,t]}$, that is, they have to form a dual implementable policy.

For the dual problem (7.5) we can develop dynamic programming equations, similarly to the primal problem (3.2). Let us consider the problem

$$\begin{aligned} &\operatorname*{Max}_{\pi_t, \ldots, \pi_T} \ \mathbb{E}\left[\sum_{\tau=t}^{T} b_\tau^T(\xi_\tau)\pi_\tau | \xi_{[1,t]}\right] \\ &\text{s.t. } A_{\tau\tau}(\xi_\tau)^T \pi_\tau + \mathbb{E}\big[A_{\tau+1,\tau}(\xi_{\tau+1})^T \pi_{\tau+1} | \xi_{[1,\tau]}\big] \\ &\le c_t(\xi_\tau), \ \tau = t-1, \ldots, T, \end{aligned} \tag{7.6}$$

In this problem, the values of π_{t-1} and of $\xi_{[1,t]}$ are assumed to be known. We denote the optimal value of this problem by $S_t(\pi_{t-1}, \xi_{[1,t]})$. These values are related for $t = 1, \ldots, T$ through the dual dynamic programming equation: $S_t(\pi_{t-1}, \xi_{[1,t]})$ is equal to the optimal value of the following problem

$$\begin{aligned} &\operatorname{Max}_{\pi_t} \ b_t(\xi_t)^T \pi_t + \mathbb{E}\big[S_{t+1}(\pi_t, \xi_{[1,t+1]}) | \xi_{[1,t]}\big] \\ &\text{s.t.} \quad A_{t-1,t-1}(\xi_{t-1})^T \pi_{t-1} + \mathbb{E}\big[A_{t,t-1}(\xi_t)^T \pi_t | \xi_{[1,t-1]}\big] \le c_{t-1}(\xi_{t-1}), \end{aligned} \tag{7.7}$$

where, for the uniformity of the notation, we assume that all terms involving $t = 0$ are zero.

There is a duality relation between the primal cost-to-go functions $Q_t(x_{t-1}, \xi_{[1,t]})$, defined in (3.3), and their dual counterparts $S_t(\pi_{t-1}, \xi_{[1,t]})$.

Theorem 45. *A feasible policy $\hat{x}_t(\xi_{[1,t]})$, $t = 1, \ldots, T$, is optimal for (3.1) and a dual feasible policy $\hat{\pi}_t(\xi_{[1,t]})$, $t = 1, \ldots, T$, is optimal for (7.5) if and only if for every realization of ξ the following holds*

$$Q_t\big(\hat{x}_{t-1}(\xi_{[1,t-1]}), \xi_{[1,t]}\big) = S_t\big(\hat{\pi}_{t-1}(\xi_{[1,t-1]}), \xi_{[1,t]}\big), \quad t = 1, \ldots, T.$$

There is another group of duality relations for multistage stochastic programs, associated with the nonaticipativity constraints.

Let us consider problem (3.2) again, but let us assume now that each decision x_t may depend on *all* random data, ξ. Since ξ has finitely many

realizations, ξ^k, $k = 1, \ldots, K$ (attained with probabilities $p_1, \ldots, p_K$), we may model our assumption by assigning a decision sequence,

$$x^k = (x_1^k, \ldots, x_T^k),$$

to the k-th realization of ξ.[17] The problem takes on the form

$$\begin{array}{lll}
\operatorname{Min} \sum_{k=1}^{K} p_k \Big[f_1(x_1^k, \xi_1^k) + f_2(x_2^k, \xi_2^k) + f_3(x_3^k, \xi_3^k) + \ldots + f_T(x_T^k, \xi_T^k) \Big] & & \\
\text{s.t.} \quad A_{11}(\xi_1^k) x_1^k & = b_1(\xi_1^k), & \\
\quad A_{21}(\xi_2^k) x_1^k + A_{22}(\xi_2^k) x_2^k & = b_2(\xi_2^k), & \\
\quad\quad A_{32}(\xi_3^k) x_2^k + A_{33}(\xi_3^k) x_3^k & = b_3(\xi_3^k), & \\
\quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & & \\
\quad\quad A_{T,T-1}(\xi_T^k) x_{T-1}^k + A_{TT}(\xi_T^k) x_T^k & = b_T(\xi_T^k), & \\
\quad\quad k = 1, \ldots, K. & &
\end{array} \tag{7.8}$$

Although similar in appearance, this formulation is not equivalent to the original problem (3.2), unless we introduce additional constraints that limit the dependence of x_t on ξ to the information that is available up to time t. As discussed in Chapter 1, these conditions take the form of *nonanticipativity constraints*,

$$x_t^k = x_t^j \text{ for all } k, j \text{ for which } \xi_{[1,t]}^k = \xi_{[1,t]}^j, \quad t = 1, \ldots, T. \tag{7.9}$$

This allows us to write problem (7.8)–(7.9) in a more abstract way. Define

$$f^k(x^k) = \begin{cases} \sum_{t=1}^{T} f_t(x_t^k, \xi_t^k), & \text{if the constraints of (7.8) are satisfied} \\ & \text{for scenario } k, \\ +\infty & \text{otherwise.} \end{cases}$$

Also, let $\mathcal{W}$ be the set of policies satisfying the nonanticipativity constraints (7.9). We see that $\mathcal{W}$ is a linear subspace of the set of all policies. The problem can be now written in a lucid form

$$\operatorname{Min} \left\{ f(x) := \sum_{k=1}^{K} p_k f^k(x^k) \right\} \quad \text{s.t.} \quad x \in \mathcal{W}. \tag{7.10}$$

[17] To avoid collisions of subscripts we slightly change our notation and use superscripts to denote realizations (scenarios).

Clearly, f is a polyhedral function, so if this problem has a solution, the optimality conditions and duality relations hold. Let us introduce the Lagrangian associated with (7.10):

$$L(x, \lambda) := f(x) + \langle \lambda, x \rangle. \tag{7.11}$$

The scalar product $\langle \lambda, x \rangle$ is understood in the usual way, as

$$\langle \lambda, x \rangle := \sum_{k=1}^{k} \sum_{t=1}^{T} \langle \lambda_t^k, x_t^k \rangle.$$

Theorem 46. *A policy $\hat{x} \in \mathcal{W}$ is an optimal solution of (7.10) if and only if there exist multipliers $\hat{\lambda} \in \mathcal{W}^{\perp}$ such that*

$$\hat{x} \in \arg \min_x \; L(x, \hat{\lambda}). \tag{7.12}$$

Proof. The result follows from Proposition 32. Indeed, $N_W(x) = \mathcal{W}^{\perp}$ for all $x \in \mathcal{W}$. Denoting by $\hat{\lambda}$ the element of $N_{\mathcal{W}}(\hat{x})$ that appears in the optimality conditions, we get

$$0 \in \partial L(\hat{x}, \hat{\lambda}). \tag{7.13}$$

Since $\mathcal{W}$ is a linear space, this is necessary and sufficient for (7.12). □

Also, we can define the dual function

$$D(\lambda) := \min_x L(x, \lambda),$$

and the dual problem

$$\operatorname*{Max}_{\lambda \in \mathcal{W}^{\perp}} \; D(\lambda). \tag{7.14}$$

Theorem 47. *The optimal values of problems (7.10) and (7.14) are equal unless both problems are infeasible. If their (common) optimal value is finite, then both problems have optimal solutions.*

The crucial role in our approach is played by the requirement that $\lambda \in \mathcal{W}^{\perp}$. Let us decipher this condition. For

$$\lambda = \left(\lambda_t^k\right)_{t=1,\ldots,T,\, k=1,\ldots,K}$$

the condition $\lambda \in \mathcal{W}^{\perp}$ is equivalent to

$$\sum_{k=1}^{k} \sum_{t=1}^{T} \left\langle \lambda_t^k, x_t^k \right\rangle = 0 \quad \text{for all } x \in \mathcal{W}.$$

Substituting

$$\pi^k = \lambda^k / p_k, \quad k = 1, \ldots, K,$$

we can write the last relation in a more abstract form as

$$\mathbb{E}\left[\sum_{t=1}^{T} \langle \pi_t, x_t \rangle \right] = 0, \quad \text{for all } x \in \mathcal{W}. \tag{7.15}$$

Since[18] $\mathbb{E}_t x_t = x_t$ for all $x \in \mathcal{W}$, we obtain from (7.15) that

$$\mathbb{E}\left[\sum_{t=1}^{T} \langle E_t \pi_t, x_t \rangle \right] = 0, \quad \text{for all } x \in \mathcal{W},$$

which is equivalent to

$$\mathbb{E}_t \pi_t = 0, \quad t = 1, \ldots, T. \tag{7.16}$$

We can now rewrite our necessary conditions of optimality and duality relations in a more explicit form. Let us re-define the Lagrangian (with a slight abuse of notation)

$$L(x, \pi) = f(x) + \mathbb{E}\left[\sum_{t=1}^{T} \langle \pi_t, x_t \rangle \right],$$

the dual functional

$$D(\pi) = \min_{x} \; L(x, \pi),$$

[18] In order to simplify notation we denote in the remainder of this section by $\mathbb{E}_t$ the conditional expectation conditional on $\xi_{[1,t]}$.

and the dual problem

$$\text{Max } D(\pi) \quad \text{s.t.} \quad E_t\pi_t = 0, \quad t = 1, \ldots, T. \tag{7.17}$$

Corollary 48. *A policy $\hat{x} \in \mathcal{W}$ is an optimal solution of (7.10) if and only if there exist multipliers $\hat{\pi}$ satisfying (7.16) such that*

$$\hat{x} \in \arg\min_x L(x, \hat{\pi}). \tag{7.18}$$

Moreover, problem (7.10) has an optimal solution if and only if problem (7.17) has an optimal solution. The optimal values of these problems are equal unless both are infeasible.

An equivalent approach to formulating the dual problem is to use algebraic expressions for the nonanticipativity constraints (7.9) and incorporate them (with the corresponding multipliers) into the Lagrangian. For example, if we write (7.9) as

$$\mathbb{E}_t x_t = x_t, \quad t = 1, \ldots, T,$$

we may formulate the Lagrangian

$$L(x, \pi) = f(x) + \mathbb{E}\left[\sum_{t=1}^{T} \langle \pi_t, x_t - \mathbb{E}_t x_t \rangle\right],$$

the dual functional

$$D(\pi) = \min_x L(x, \pi),$$

and the dual problem

$$\text{Max } D(\pi). \tag{7.19}$$

There are no constraints on π in this dual problem. Since f is polyhedral and the nonanticipativity conditions linear, Kuhn–Tucker optimality conditions and duality relations hold for this formulation, similarly to Corollary 48, but without additional constraints on π of form (7.16). However, these constraints may be included into the dual problem without any loss of optimality. To prove that, let us consider any dual solution π and define

$$\overline{\pi}_t = \pi_t - \mathbb{E}_t\pi_t, \quad t = 1, \ldots, T.$$

Clearly, it satisfies (7.16). Now, for any x we have

$$L(x,\pi) - L(x,\overline{\pi}) = \mathbb{E}\left[\sum_{t=1}^{T} \langle \mathbb{E}_t \pi_t, x_t - \mathbb{E}_t x_t \rangle\right]$$
$$= \sum_{t=1}^{T} (\mathbb{E}\langle \mathbb{E}_t \pi_t, x_t \rangle - \mathbb{E}\langle \mathbb{E}_t \pi_t, \mathbb{E}_t x_t \rangle) = 0.$$

Consequently, $\partial L(x,\pi) = \partial L(x,\overline{\pi})$, $D(\pi) = D(\overline{\pi})$, so $\overline{\pi}$ can be substituted for π in the optimality conditions and duality relations.

There are many different ways to express the nonanticipativity constraints (7.9), and thus there are many equivalent ways to formulate the Lagrangian and the dual problem. Some of them may be more convenient for some computational methods, other may be more suitable for other methods. We shall return to these issues in the sections devoted to numerical methods for solving stochastic programming problems.

8 Min–max stochastic optimization

In practical applications the required probability distributions are either estimated from available historical data or assigned by a subjective judgment. Consequently, these distributions are never known exactly and to some extent are also uncertain. We already briefly discussed that problem in Section 4 of Chapter 1. In order to deal with the distribution uncertainty one can formulate the following min–max analogue of problem (4.1):

$$\operatorname*{Min}_{x\in X} \operatorname*{Max}_{\mu\in\mathbb{S}} \mathbb{E}_\mu[F(x,\omega)]. \tag{8.1}$$

Here $\mathbb{S}$ is a given set of probability measures (distributions), defined on a sample space $(\Omega,\mathcal{F})$, and the notation $\mathbb{E}_\mu$ means that the expectation is taken with respect to measure $\mu \in \mathbb{S}$. Of course, if $\mathbb{S} := \{P\}$ is a singleton, then the above problem (8.1) coincides with problem (4.1). In this section we discuss some basic properties of the min–max problem (8.1).

We assume that the sets X and $\mathbb{S}$ are nonempty, and that for every $x \in X$ and $\mu \in \mathbb{S}$, the expectation $\phi(x,\mu) := \mathbb{E}_\mu[F(x,\omega)]$ is well defined. Interchanging the order of Min and Max operators, we obtain the following dual of the problem (8.1):

$$\operatorname*{Max}_{\mu\in\mathbb{S}} \operatorname*{Min}_{x\in X} \mathbb{E}_\mu[F(x,\omega)]. \tag{8.2}$$

By the general theory of min–max duality we have that the optimal value of (8.1) is always greater than or equal to the optimal value of (8.2).

The function $\phi(x, \mu)$ is linear in μ. Therefore, the max-value of problem (8.1) is not changed if the set $\mathcal{S}$ is substituted by its convex hull $\mathcal{S}^* := \text{conv}(\mathcal{S})$. This substitution may effect, however, the optimal value of problem (8.2). We assume throughout this section that the set X is convex and for every $\omega \in \Omega$, the function $F(\cdot, \omega)$ is convex. This implies, of course, that the expectation function $\phi(\cdot, \mu)$ is also convex for any $\mu \in \mathcal{S}$. In order to get a better insight into the problem let us discuss the following particular cases.

Suppose that the set $\mathcal{S}$ is finite, say $\mathcal{S} := \{P_1, \ldots, P_\ell\}$. We already briefly discussed that case in Section 4 of Chapter 1. Suppose also, for the sake of simplicity, that the functions $f_i(x) := \mathbb{E}_{P_i}[F(x, \omega)]$, $i = 1, \ldots, \ell$, are real valued for all $x \in \mathbb{R}^n$. By the convexity assumption we have that these functions are convex. Then the max-function

$$f_0(x) := \max_{i \in \{1, \ldots, \ell\}} f_i(x)$$

is also real valued and convex. Since functions f_i, $i = 0, \ldots, \ell$, are real valued and convex, these functions are continuous and subdifferentiable on $\mathbb{R}^n$. We have the following formula for the subdifferential of the max-function (see Theorem 51 in the Appendix)

$$\partial f_0(x) = \text{conv}\big\{\cup_{i \in I(x)} \partial f_i(x)\big\}, \tag{8.3}$$

where

$$I(x) := \big\{i : f_0(x) = f_i(x),\ i = 1, \ldots, \ell\big\}$$

is the set of active at x functions. By the optimality condition (4.5) we have[19] that a point $\overline{x} \in X$ is an optimal solution of the corresponding min–max problem (8.1) iff there exist nonnegative multipliers λ_i, $i \in I(\overline{x})$, such that $\sum_{i \in I(\overline{x})} \lambda_i = 1$ and

$$0 \in \sum_{i \in I(\overline{x})} \lambda_i \partial f_i(\overline{x}) + N_X(\overline{x}). \tag{8.4}$$

[19] Since it is assumed here that the function $f(x)$ is real valued, and hence its domain is $\mathbb{R}^n$, constraint qualification (4.4) holds automatically.

By the Moreau–Rockafellar Theorem we also have that

$$\sum_{i\in I(\bar{x})} \lambda_i \partial f_i(\bar{x}) = \partial\left(\sum_{i\in I(\bar{x})} \lambda_i f_i(\bar{x})\right).$$

Therefore, $\bar{x}$ is an optimal solution of the problem

$$\operatorname*{Min}_{x\in X} \mathbb{E}_{P^*}[F(x,\omega)], \tag{8.5}$$

where $P^* := \sum_{i\in I(\bar{x})} \lambda_i P_i$. It also follows that $(\bar{x}, P^*)$ is a saddle point of the corresponding min–max problem with the set $\mathbb{S}$ replaced by its convex hull $\mathbb{S}^* := \operatorname{conv}(\mathbb{S})$. We have here that there is no duality gap[20] between problem (8.1) and its dual

$$\operatorname*{Max}_{\mu\in\mathbb{S}^*} \operatorname*{Min}_{x\in X} \mathbb{E}_{\mu}[F(x,\omega)], \tag{8.6}$$

and the set of optimal solutions of the dual problem (8.6) is nonempty. Note that the optimal value of problem (8.2) can be smaller that the optimal value of (8.6), and therefore it is essential here that the (finite) set $\mathbb{S}$ is replaced in (8.6) by its convex hull $\mathbb{S}^*$.

Suppose now that the set $\Omega = \{\omega_1, \ldots, \omega_K\}$ is finite. Then a probability measure (distribution) on Ω is defined by a vector $P = (p_1, \ldots, p_K) \in \mathbb{R}^K_+$ such that $\sum_{k=1}^K p_k = 1$. Let $\mathbb{S} \subset \mathbb{R}^K_+$ be a set of such vectors. Then the corresponding min–max problem can be written as

$$\operatorname*{Min}_{x\in X}\left\{f_0(x) := \sup_{P\in\mathbb{S}} \sum_{k=1}^K p_k F(x,\omega_k)\right\}. \tag{8.7}$$

Suppose that the set $\mathbb{S}$ is convex and compact, the set X is convex and the functions $F(\cdot,\omega_k)$ are convex real valued. We have then that the max-function $f_0(x)$ is convex real valued, and by Theorem 51,

$$\partial f_0(x) = \operatorname{conv}\left\{\bigcup_{P\in\mathbb{S}_0(x)} \left(\sum_{k=1}^K p_k \partial F(x,\omega_k)\right)\right\}, \tag{8.8}$$

[20] The above derivations are based on existence of an optimal solution of (8.1). In the present case, however, this is not essential for the "no duality gap" property to hold, which can be proved directly, for example, by using the conjugate duality approach.

where

$$\mathbb{S}_0(x) := \arg\max_{P\in\mathbb{S}} \sum_{k=1}^{K} p_k F(x,\omega_k).$$

Note that the set $\mathbb{S}_0(x)$ is convex, nonempty and compact since the set $\mathbb{S}$ is convex, nonempty and compact.

By the optimality condition (4.5) together with formula (8.8) we obtain that $(\overline{x}, P^*)$ is a saddle point of the corresponding min–max problem iff $\overline{x} \in X$, $P^* \in \mathbb{S}_0(\overline{x})$ and

$$0 \in \partial(\mathbb{E}_{P^*}[F(\overline{x},\omega)]) + N_X(\overline{x}). \tag{8.9}$$

We also have that a point $\overline{x} \in X$ is an optimal solution of problem (8.7) iff there exists $P^* \in \mathbb{S}_0(\overline{x})$ such that (8.9) holds. Therefore, if problem (8.7) has an optimal solution, then its dual problem also has an optimal solution P^*, there is no duality gap between (8.7) and its dual, and the set of optimal solutions of (8.7) coincides with the set of optimal solutions of the corresponding problem (8.5).

The above analysis can be extended to the following general case of a metric space Ω and its Borel sigma algebra $\mathcal{B}$. Denote $\mathbb{S}_0(x) := \arg\max_{\mu\in\mathbb{S}} \mathbb{E}_\mu[F(x,\omega)]$.

Theorem 49. *Let Ω be a metric space equipped with its Borel sigma algebra $\mathcal{B}$, X be a nonempty convex subset of $\mathbb{R}^n$ and $\mathbb{S}$ be a nonempty convex set of probability measures on $(\Omega,\mathcal{B})$. Suppose that Ω is compact, for every $x \in \mathbb{R}^n$ the function $F(x,\cdot)$ is continuous on Ω, and for every $\omega \in \Omega$ the function $F(\cdot,\omega)$ is real valued and convex on $\mathbb{R}^n$. Then there is no duality gap between problem (8.1) and its dual (8.2). Suppose, further, that the set $\mathbb{S}$ is closed in the weak topology of the space of probability measures on $(\Omega,\mathcal{B})$, and the optimal value of problem (8.1) is finite. Then the dual problem (8.2) has a nonempty set of optimal solutions, and a point $\overline{x} \in X$ is an optimal solution of (8.1) if and only if there exists $P^* \in \mathbb{S}_0(\overline{x})$ such that $\overline{x}$ is an optimal solution of problem (8.5).*

Proof. We outline a proof in the case where problem (8.1) has an optimal solution. Since $F(x,\cdot)$ is continuous and Ω is compact, the expectation $\mathbb{E}_\mu[F(x,\omega)]$ is well defined for any $(x,\mu) \in \mathbb{R}^n \times \mathbb{S}$. Let us equip $\mathbb{S}$ with the weak topology and consider function $\phi(x,\mu) := \mathbb{E}_\mu[F(x,\omega)]$. We have that for any $\mu \in \mathbb{S}$ the function $\phi(\cdot,\mu)$ is convex and for any $x \in \mathbb{R}^n$ the function $\phi(x,\cdot)$ is continuous on $\mathbb{S}$. Since Ω is compact, we have by Prohorov's theorem that the topological closure of $\mathbb{S}$ is compact. We can assume that $\mathbb{S}$ is closed and therefore is compact. It follows that the max-function $f_0(x) := \sup_{\mu\in\mathbb{S}} \phi(x,\mu)$ is

convex real valued. Let $\bar{x}$ be an optimal solution of problem (8.1). By Theorem 51 (from the Appendix) we have that

$$\partial f_0(\bar{x}) = \operatorname{conv}\left(\bigcup_{\mu \in \mathbb{S}_0(\bar{x})} \partial \phi(\bar{x}, \mu)\right). \tag{8.10}$$

This implies existence of a saddle point $(\bar{x}, P^*)$, and hence the assertions of the theorem follow. □

Suppose now that the set $\mathbb{S}$ is defined as the set of probability measures μ on $(\Omega, \mathcal{B})$ satisfying the constraints

$$\mathbb{E}_\mu[\psi_j(\omega)] = b_j, \quad j = 1, \ldots, m. \tag{8.11}$$

Here $\psi_1(\omega), \ldots, \psi_m(\omega)$ are real valued measurable functions on $(\Omega, \mathcal{B})$ and $b_1, \ldots, b_m$ are given numbers. Then the problem

$$\operatorname*{Max}_{\mu \in \mathbb{S}} \ \mathbb{E}_\mu[F(x, \omega)] \tag{8.12}$$

is called the *problem of moments*. By Theorem 60 (from the Appendix) we have that it suffices to perform optimization in the above problem (8.12) over probability measures with a finite support of at most $s = m + 1$ points. That is, problem (8.12) is equivalent to the problem

$$\begin{aligned} \operatorname{Max} & \sum_{i=1}^{s} p_i F(x, \omega_i) \\ \text{s.t.} & \sum_{i=1}^{s} p_i \psi_j(\omega_i) = b_i, \quad j = 1, \ldots, m, \\ & \sum_{i=1}^{s} p_i = 1, \ p_i \geq 0, \quad i = 1, \ldots, s, \end{aligned} \tag{8.13}$$

where the maximum is taken with respect to $P = (p_1, \ldots, p_s)$ and $\omega_1, \ldots, \omega_s \in \Omega$.

9 Appendix

9.1 Differentiability and convex analysis

Consider a mapping $g: \mathbb{R}^n \to \mathbb{R}^m$. It is said that g is directionally differentiable at a point $x_0 \in \mathbb{R}^n$ in a direction $h \in \mathbb{R}^n$ if the limit

$$g'(x_0, h) := \lim_{t \downarrow 0} \frac{g(x_0 + th) - g(x_0)}{t} \tag{9.1}$$

exists, in which case $g'(x_0, h)$ is called the *directional derivative* of $g(x)$ at x_0 in the direction h. If g is directionally differentiable at x_0 in every direction $h \in \mathbb{R}^n$, then it is said that g is directionally differentiable at x_0. Note that whenever exists, $g'(x_0, h)$ is positively homogeneous in h, i.e., $g'(x_0, th) = tg'(x_0, h)$ for any $t \geq 0$. If $g(x)$ is directionally differentiable at x_0 and $g'(x_0, h)$ is *linear* in h, then it is said that $g(x)$ is Gâteux differentiable at x_0. (9.1) can be also written in the form

$$g(x_0 + h) = g(x_0) + g'(x_0, h) + r(h), \tag{9.2}$$

where the remainder term $r(h)$ is such that $r(th)/t \to 0$, as $t \downarrow 0$, for any fixed $h \in \mathbb{R}^n$. If, moreover, $g'(x_0, h)$ is linear in h and the remainder term $r(h)$ is 'uniformly small' in the sense that $r(h)/\|h\| \to 0$ as $h \to 0$, then it is said that $g(x)$ is differentiable at x_0 in the sense of Fréchet, or simply differentiable at x_0.

Clearly Fréchet differentiability implies Gâteux differentiability. The converse of that is not necessarily true. However, for locally Lipschitz continuous mappings both concepts do coincide. That is, if $g(x)$ is Lipschitz continuous in a neighborhood of x_0 and directionally differentiable at x_0, then $g(x)$ is directionally differentiable at x_0 in the sense of Fréchet and $g'(x_0, \cdot)$ is Lipschitz continuous on $\mathbb{R}^n$. Recall that $g(x)$ is said to be Lipschitz continuous on a set $X \subset \mathbb{R}^n$ if there is a positive constant c such that $||g(x_1) - g(x_2)|| \leq c||x_1 - x_2||$ for all $x_1, x_2 \in X$.

Let C be a subset of $\mathbb{R}^n$. It is said that $x \in \mathbb{R}^n$ is an *interior point* of C if there is a neighborhood N of x such that $N \subset C$. The set of interior points of C is denoted $\text{int}(C)$. The *convex hull* of C, denoted $\text{conv}(C)$, is the smallest convex set including C. It is said that C is a *cone* if for any $x \in C$ and $t \geq 0$ it follows that $tx \in C$. The *polar cone* of a cone $C \subset \mathbb{R}^n$ is defined as

$$C^* := \{z \in \mathbb{R}^n : z^T x \leq 0, \ \forall\, x \in C\}. \tag{9.3}$$

We have that the polar of the polar cone $C^{**} = (C^*)^*$ is equal to the topological closure of the convex hull of C, and that $C^{**} = C$ iff the cone C is convex and closed.

Let C be a *convex* subset of $\mathbb{R}^n$. The affine space generated by C is the space of points in $\mathbb{R}^n$ of the form $tx + (1 - t)y$, where $x, y \in C$ and $t \in \mathbb{R}$. It is said that a point $x \in \mathbb{R}^n$ belongs to the *relative interior* of the set C if x is an interior point of C relative to the affine space generated by C, i.e., there exists a neighborhood of x such that its intersection with the affine space generated by

C is included in C. The relative interior set of C is denoted $\mathrm{ri}(C)$. Note that if the interior of C is nonempty, then the affine space generated by C coincides with $\mathbb{R}^n$, and hence in that case $\mathrm{ri}(C) = \mathrm{int}(C)$. The *normal cone* to C at a point $x_0 \in C$ is defined as

$$N_C(x_0) := \{z : z^T(x - x_0) \leq 0, \ \forall\, x \in C\}. \tag{9.4}$$

The topological closure of the *radial cone* $R_C(x_0) := \cup_{t>0} \{t^{-1}(C - x_0)\}$ is called the *tangent cone* to C at x_0, and denoted $T_C(x_0)$. Both cones $T_C(x_0)$ and $N_C(x_0)$ are closed and convex, and each one is the polar cone of the other.

The *support function* of a set $C \subset \mathbb{R}^n$ is defined as

$$s(h) := \sup_{z \in C} z^T h. \tag{9.5}$$

The support function $s(\cdot)$ is convex, positively homogeneous and lower semicontinuous. If $s_1(\cdot)$ and $s_2(\cdot)$ are support functions of convex closed sets A and B, respectively, then $s_1(\cdot) \leq s_2(\cdot)$ iff $A \subset B$, and $s_1(\cdot) = s_2(\cdot)$ iff $A = B$.

Consider an extended real valued function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$. The domain of f is defined as $\mathrm{dom} f := \{x \in \mathbb{R}^n : f(x) < +\infty\}$. It is said that f is *proper* if its domain is nonempty and $f(x) > -\infty$ for all $x \in \mathbb{R}^n$. It is not difficult to show that f is convex iff its epigraph $\mathrm{epi} f := \{(x, \alpha) : f(x) \leq \alpha\}$ is a convex subset of $\mathbb{R}^{n+1}$.

Suppose now that $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is a *convex* function and $x_0 \in \mathbb{R}^n$ is a point such that $f(x_0)$ is *finite*. Then $f(x)$ is directionally differentiable at x_0, its directional derivative $f'(x_0, \cdot)$ is an extended real valued convex positively homogeneous function and can be written in the form

$$f'(x_0, h) = \inf_{t>0} \frac{f(x_0 + th) - f(x_0)}{t}. \tag{9.6}$$

Moreover, if x_0 is in the interior of the domain of $f(\cdot)$, then $f(x)$ is Lipschitz continuous in a neighborhood of x_0, the directional derivative $f'(x_0, h)$ is finite valued for any $h \in \mathbb{R}^n$, and $f(x)$ is differentiable at x_0 iff $f'(x_0, h)$ is linear in h.

It is said that a vector $z \in \mathbb{R}^n$ is a *subgradient* of $f(x)$ at x_0 if

$$f(x) - f(x_0) \geq z^T(x - x_0), \quad \forall\, x \in \mathbb{R}^n. \tag{9.7}$$

The set of all subgradients of $f(x)$, at x_0, is called the *subdifferential* and denoted $\partial f(x_0)$. The subdifferential $\partial f(x_0)$ is a closed convex subset of $\mathbb{R}^n$. It is said that f is *subdifferentiable* at x_0 if $\partial f(x_0)$ is nonempty. If f is subdifferentiable at x_0, then the normal cone $N_{\mathrm{dom} f}(x_0)$, to the domain of f at

x_0, forms the recession cone of the set $\partial f(x_0)$. It is also clear that if f is subdifferentiable at x_0, then $f(x) > -\infty$ for any x and hence f is proper.

By duality theory of convex analysis we have that if the directional derivative $f'(x_0, \cdot)$ is lower semicontinuous, then

$$f'(x_0, h) = \sup_{z \in \partial f(x_0)} z^T h, \tag{9.8}$$

i.e., $f'(x_0, \cdot)$ is the support function of the set $\partial f(x_0)$. In particular, if x_0 is an interior point of the domain of $f(x)$, then $f'(x_0, \cdot)$ is continuous, $\partial f(x_0)$ is nonempty and compact and (9.8) holds. Conversely, if $\partial f(x_0)$ is nonempty and compact, then x_0 is an interior point of the domain of $f(x)$. Also $f(x)$ is differentiable at x_0 iff $\partial f(x_0)$ is a singleton, i.e., contains only one element, which then coincides with the gradient $\nabla f(x_0)$.

Theorem 50 (Moreau–Rockafellar). *Let $f_i: \mathbb{R}^n \to \overline{\mathbb{R}}$, $i = 1, \ldots, m$, be proper convex functions, $f(\cdot) := f_1(\cdot) + \ldots + f_m(\cdot)$ and x_0 be a point such that $f_i(x_0)$ are finite, i.e., $x_0 \in \cap_{i=1}^m \mathrm{dom} f_i$ Then*

$$\partial f_1(x_0) + \cdots + \partial f_m(x_0) \subset \partial f(x_0). \tag{9.9}$$

Moreover,

$$\partial f_1(x_0) + \cdots + \partial f_m(x_0) = \partial f(x_0) \tag{9.10}$$

if any one of the following conditions holds: (i) *the set $\cap_{i=1}^m \mathrm{ri}(\mathrm{dom} f_i)$ is nonempty,* (ii) *the functions $f_1, \ldots, f_k$, $k \leq m$, are polyhedral and the intersection of the sets $\cap_{i=1}^k \mathrm{dom} f_i$ and $\cap_{i=k+1}^m \mathrm{ri}(\mathrm{dom} f_i)$ is nonempty,* (iii) *there exists a point $\overline{x} \in \mathrm{dom} f_m$ such that $\overline{x} \in \mathrm{int}(\mathrm{dom} f_i)$, $i = 1, \ldots, m-1$.*

In particular, if all functions $f_1, \ldots, f_m$ in the above theorem are polyhedral, then the equation (9.10) holds without an additional regularity condition.

The following result gives a description of subdifferentials of max-functions. By $\mathrm{cl}(A)$ we denote the topological closure of a set $A \subset \mathbb{R}^n$.

Theorem 51 (Levin-Valadier). *Let U be a compact topological space and $g: \mathbb{R}^n \times U \to \mathbb{R}$ be a real valued function. Suppose that:* (i) *for every $u \in U$ the function $g_u(\cdot) = g(\cdot, u)$ is convex on $\mathbb{R}^n$,* (ii) *for every $x \in \mathbb{R}^n$ the function $g(x, \cdot)$ is upper semicontinuous on U. Then the max-function $f(x) := \sup_{u \in U} g(x, u)$ is convex real valued and*

$$\partial f(x) = \mathrm{cl}\big\{\mathrm{conv}\big(\cup_{u \in U_0(x)} \partial g_u(x)\big)\big\}, \tag{9.11}$$

where $U_0(x) := \arg\max_{u\in U} g(x,u)$.

Let us make the following observations regarding the above theorem. Since U is compact and by the assumption (ii), we have that the set $U_0(x)$ is nonempty and compact. Since the function $f(\cdot)$ is convex real valued, it is subdifferentiable at every $x \in \mathbb{R}^n$ and its subdifferential $\partial f(x)$ is a convex, closed bounded subset of $\mathbb{R}^n$. It follows then from (9.11) that the set $A := \cup_{u\in U_0(x)} \partial g_u(x)$ is bounded. Suppose further that:

(iii) For every $x \in \mathbb{R}^n$ the function $g(x,\cdot)$ is continuous on U.

Then the set A is closed, and hence is compact. Indeed, consider a sequence $z_k \in A$. Then, by the definition of the set A, $z_k \in \partial g_{u_k}(x)$ for some sequence $u_k \in U_0(x)$. Since $U_0(x)$ is compact and A is bounded, by passing to a subsequence if necessary, we can assume that u_k converges to a point $\overline{u} \in U_0(x)$ and z_k converges to a point $\overline{z} \in \mathbb{R}^n$. By the definition of subgradients z_k we have that for any $x' \in \mathbb{R}^n$ the following inequality holds

$$g_{u_k}(x') - g_{u_k}(x) \geq z_k^T(x' - x).$$

By passing to the limit in the above inequality as $k \to \infty$, we obtain that $\overline{z} \in \partial g_{\overline{u}}(x)$. It follows that $\overline{z} \in A$, and hence A is closed. Now since convex hull of a compact subset of $\mathbb{R}^n$ is also compact, and hence is closed, we obtain that if the assumption (ii) in the above theorem is strengthened to the assumption (iii), then the set inside the parentheses in (9.11) is closed, and hence formula (9.11) takes the form

$$\partial f(x) = \operatorname{conv}\left(\bigcup_{u\in U_0(x)} \partial g_u(x)\right). \tag{9.12}$$

Let $f\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ be an extended real valued function. The *conjugate function* of f is

$$f^*(z) := \sup_{x\in\mathbb{R}^n} \{z^T x - f(x)\}. \tag{9.13}$$

The conjugate function $f^*\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ is always convex and lower semicontinuous.

Theorem 52 (Fenchel–Moreau). *Let* $f\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ *be a proper extended real valued function. Then*

$$f^{**} = \operatorname{lsc}(\operatorname{conv} f). \tag{9.14}$$

Here f^{**} denotes the conjugate of f^* and $\operatorname{conv} f$ denotes the convex hull of f,

$$(\operatorname{conv} f)(x) = \inf\{\alpha : (x, \alpha) \in \operatorname{conv}(\operatorname{epi} f)\},$$

i.e., $\operatorname{conv} f$ is the largest convex function majorized by f. Note that if $f(x) = -\infty$ at some $x \in \mathbb{R}^n$, then $f^*(\cdot) \equiv +\infty$ and $f^{**}(\cdot) \equiv -\infty$. It follows from (9.14) if f is proper, then $f^{**} = f$ iff f is convex and lower semicontinuous.

It immediately follows from the definitions that

$$z \in \partial f(x) \quad \text{iff} \quad f^*(z) = z^T x - f(x).$$

By applying that to the function f^{**}, instead of f, we obtain that $z \in \partial f^{**}(x)$ iff $f^{***}(z) + f^{**}(x) = z^T x$. Now by the Fenchel–Moreau theorem we have that $f^{***} = f^*$. Consequently, we obtain

$$\partial f^{**}(x) = \arg\max_{z \in \mathbb{R}^n} \{z^T x - f^*(z)\}. \tag{9.15}$$

The following result about Lipschitz continuity of linear systems is known as Hoffman's lemma.

Theorem 53 (Hoffman). *Consider the multifunction $\mathcal{M}(b) := \{x \in \mathbb{R}^n : Ax \leq b\}$, where A is a given $m \times n$ matrix. Then there exists a positive constant κ, depending on A, such that for any $x \in \mathbb{R}^n$ and any $b \in \operatorname{dom} \mathcal{M}$,*

$$\operatorname{dist}(x, \mathcal{M}(b)) \leq \kappa \|(Ax - b)_+\|. \tag{9.16}$$

The term[21] $\|(Ax - b)_+\|$, in the right hand side of (9.16), measures the infeasibility of the point x.

9.2 Duality of optimization problems

Consider a real valued function $L : X \times Y \to \mathbb{R}$, where X and Y are arbitrary sets. We can associate with the function $L(x, y)$ the following two optimization problems:

$$\operatorname*{Min}_{x \in X} \left\{ f(x) := \sup_{y \in Y} L(x, y) \right\}, \tag{9.17}$$

[21] The operator $(\cdot)_+$ applied to a vector is taken componentwise.

$$\operatorname*{Max}_{y\in Y}\left\{g(y) := \inf_{x\in X} L(x,y)\right\}, \tag{9.18}$$

viewed as dual to each other. We have that for any $x \in X$ and $y \in Y$,

$$g(y) = \inf_{x'\in X} L(x',y) \le L(x,y) \le \sup_{y'\in Y} L(x,y') = f(x),$$

and hence the optimal value of problem (9.17) is greater than or equal to the optimal value of problem (9.18). It is said that a point $(\overline{x},\overline{y}) \in X \times Y$ is a *saddle point* of $L(x,y)$ if

$$L(\overline{x},y) \le L(\overline{x},\overline{y}) \le L(x,\overline{y}), \quad \forall\, (x,y) \in X \times Y. \tag{9.19}$$

Proposition 54. *The following holds:* (i) *The optimal value of problem* (9.17) *is greater than or equal to the optimal value of problem* (9.18). (ii) *Problems* (9.17) and (9.18) *have the same optimal value and each has an optimal solution if and only if there exists a saddle point* $(\overline{x},\overline{y})$. *In that case* $\overline{x}$ *and* $\overline{y}$ *are optimal solutions of problems* (9.17) and (9.18), *respectively*. (iii) *If problems* (9.17) and (9.18) *have the same optimal value, then the set of saddle points coincides with the Cartesian product of the sets of optimal solutions of* (9.17) and (9.18).

In applications of the above results to optimization problems with constraints, the function $L(x,y)$ usually is the Lagrangian of the problem and y is a vector of Lagrange multipliers.

An alternative approach to duality, referred to as *conjugate duality*, is the following. Consider an extended real valued function $\psi\colon \mathbb{R}^n \times \mathbb{R}^m \to \overline{\mathbb{R}}$. Let $\vartheta(y)$ be the optimal value of the parameterized problem

$$\operatorname*{Min}_{x\in\mathbb{R}^n} \psi(x,y), \tag{9.20}$$

i.e., $\vartheta(y) := \inf_{x\in\mathbb{R}^n}\psi(x,y)$. Note that implicitly the optimization in the above problem is performed over the domain of the function $\psi(\cdot,y)$, i.e., $\operatorname{dom}\psi(\cdot,y)$ can be viewed as the feasible set of problem (9.20).

The conjugate of the function $\vartheta(y)$ can be expressed in terms of the conjugate of $\psi(x,y)$. That is, the conjugate of ψ is

$$\psi^*(x^*,y^*) := \sup_{(x,y)\in\mathbb{R}^n\times\mathbb{R}^m} \left\{(x^*)^T x + (y^*)^T y - \psi(x,y)\right\},$$

and hence the conjugate of ϑ can be written as

$$\begin{aligned}\vartheta^*(y^*) &:= \sup_{y\in\mathbb{R}^m}\{(y^*)^T y - \vartheta(y)\} = \sup_{y\in\mathbb{R}^m}\{(y^*)^T y - \inf_{x\in\mathbb{R}^n}\psi(x,y)\}\\ &= \sup_{(x,y)\in\mathbb{R}^n\times\mathbb{R}^m}\{(y^*)^T y - \psi(x,y)\} = \psi^*(0,y^*).\end{aligned}$$

Consequently, the conjugate of ϑ^* is

$$\vartheta^{**}(y) = \sup_{y^*\in\mathbb{R}^m}\{(y^*)^T y - \psi^*(0,y^*)\}. \tag{9.21}$$

This leads to the following dual of (9.20):

$$\operatorname*{Max}_{y^*\in\mathbb{R}^m}\{(y^*)^T y - \psi^*(0,y^*)\}. \tag{9.22}$$

In the above formulation of problem (9.20) and its (conjugate) dual (9.22) we have that $\vartheta(y)$ and $\vartheta^{**}(y)$ are optimal values of (9.20) and (9.22), respectively. We also have by the Fenchel–Moreau Theorem that either $\vartheta^{**}(\cdot)$ is identically $-\infty$, or

$$\vartheta^{**}(y) = \operatorname{lsc}(\operatorname{conv}\vartheta)(y), \quad \forall\, y\in\mathbb{R}^m. \tag{9.23}$$

It follows that $\vartheta^{**}(y) \le \vartheta(y)$ for any $y\in\mathbb{R}^m$. It is said that there is no duality gap between (9.20) and its dual (9.22) if $\vartheta^{**}(y) = \vartheta(y)$.

Suppose now that the function $\psi(x,y)$ is *convex* (as a function of $(x,y)\in\mathbb{R}^n\times\mathbb{R}^m$). It is straightforward then to verify that the optimal value function $\vartheta(y)$ is also convex, and hence $\operatorname{conv}\vartheta(\cdot)\equiv\vartheta(\cdot)$. It is said that the problem (9.20) is *subconsistent*, for a given value of y, if $\operatorname{lsc}\vartheta(y) < +\infty$. If problem (9.20) is feasible, i.e., $\operatorname{dom}\psi(\cdot,y)$ is nonempty, then $\vartheta(y) < +\infty$, and hence (9.20) is subconsistent.

Proposition 55. *Suppose that the function $\psi(\cdot,\cdot)$ is convex. Then the following holds:* (i) *The optimal value function $\vartheta(\cdot)$ is convex.* (ii) *If problem (9.20) is subconsistent, then $\vartheta^{**}(y) = \vartheta(y)$ if and only if the optimal value function $\vartheta(\cdot)$ is lower semicontinuous at y.* (iii) *If $\vartheta^{**}(y)$ is finite, then the set of optimal solutions of the dual problem (9.22) coincides with $\partial\vartheta^{**}(y)$.* (iv) *The set of optimal solutions of the dual problem (9.22) is nonempty and bounded if and only if $\vartheta(y)$ is finite and $\vartheta(\cdot)$ is continuous at y.*

A few words about the above statements are now in order. Assertion (ii) follows by the Fenchel–Moreau Theorem. Assertion (iii) follows from formula (9.15). If $\vartheta(\cdot)$ is continuous at y, then it is lower semicontinuous at y, and hence $\vartheta^{**}(y) = \vartheta(y)$. Moreover, in that case $\partial\vartheta^{**}(y) = \partial\vartheta(y)$ and is nonempty and bounded provided that $\vartheta(y)$ is finite. It follows then that the set of

optimal solutions of the dual problem (9.22) is nonempty and bounded. Conversely, if the set of optimal solutions of (9.22) is nonempty and bounded, then, by (iii), $\partial\vartheta^{**}(y)$ is nonempty and bounded, and hence by convex analysis $\vartheta(\cdot)$ is continuous at y. Note also that if $\partial\vartheta(y)$ is nonempty, then $\vartheta^{**}(y) = \vartheta(y)$ and $\partial\vartheta^{**}(y) = \partial\vartheta(y)$.

The above analysis can be also used in order to describe differentiability properties of the optimal value function $\vartheta(\cdot)$ in terms of its subdifferentials.

Proposition 56. *Suppose that the function $\psi(\cdot,\cdot)$ is convex and let $y \in \mathbb{R}^m$ be a given point. Then the following holds:* (i) *The optimal value function $\vartheta(\cdot)$ is subdifferentiable at y if and only if $\vartheta(\cdot)$ is lower semicontinuous at y and the dual problem (9.22) possesses an optimal solution.* (ii) *The subdifferential $\partial\vartheta(y)$ is nonempty and bounded if and only if $\vartheta(y)$ is finite and the set of optimal solutions of the dual problem (9.22) is nonempty and bounded.* (iii) *In both above cases $\partial\vartheta(y)$ coincides with the set of optimal solutions of the dual problem (9.22).*

Since $\vartheta(\cdot)$ is convex, we also have that $\partial\vartheta(y)$ is nonempty and bounded iff $\vartheta(y)$ is finite and $y \in \operatorname{int}(\operatorname{dom}\vartheta)$. The condition $y \in \operatorname{int}(\operatorname{dom}\vartheta)$ means the following: there exists a neighborhood N of y such that for any $y' \in N$ the domain of $\psi(\cdot, y')$ is nonempty.

As an example let us consider the following problem

$$\begin{aligned} &\operatorname{Min}_{x\in X} f(x) \\ &\text{subject to} \quad g_i(x) + y_i \le 0, \quad i = 1,\ldots,m, \end{aligned} \tag{9.24}$$

where X is a subset of $\mathbb{R}^n$, $f(x)$ and $g_i(x)$ are real valued functions, and $y = (y_1,\ldots,y_m)$ is a vector of parameters. We can formulate this problem in the form (9.20) by defining

$$\psi(x,y) := \overline{f}(x) + F(G(x) + y),$$

where $\overline{f}(x) := f(x) + i_X(x)$ and $F(\cdot)$ is the indicator function of the negative orthant, i.e., $F(z) := 0$ if $z_i \le 0$, $i = 1,\ldots,m$, and $F(z) := +\infty$ otherwise, and $G(x) := (g_1(x),\ldots,g_m(x))$.

Suppose that the problem (9.24) is convex, that is, the set X and the functions $f(x)$ and $g_i(x)$, $i = 1,\ldots,m$, are convex. Then it is straightforward to verify that the function $\psi(x,y)$ is also convex. Let us calculate the conjugate of the function $\psi(x,y)$,

$$\begin{aligned} \psi^*(x^*,y^*) &= \sup_{(x,y)\in\mathbb{R}^n\times\mathbb{R}^m}\left\{((x^*)^T x + (y^*)^T y - \overline{f}(x) - F(G(x)+y)\right\} \\ &= \sup_{x\in\mathbb{R}^n}\left\{((x^*)^T x - \overline{f}(x) - (y^*)^T G(x)\right. \\ &\quad \left. + \sup_{y\in\mathbb{R}^m}\left[(y^*)^T(G(x)+y) - F(G(x)+y)\right]\right\}. \end{aligned}$$

By change of variables $z = G(x) + y$ we obtain that

$$\sup_{y\in\mathbb{R}^m}\big[(y^*)^T(G(x)+y) - F(G(x)+y)\big] = \sup_{z\in\mathbb{R}^m}\big[(y^*)^T z - F(z)\big] = i_{\mathbb{R}^m_+}(y^*).$$

Therefore we obtain

$$\psi^*(x^*, y^*) = \sup_{x\in X}\big\{(x^*)^T x - L(x, y^*)\big\} + i_{\mathbb{R}^m_+}(y^*),$$

where $L(x, y^*) := f(x) + \sum_{i=1}^m y_i^* g_i(x)$, is the Lagrangian of the problem. Consequently, the dual of the problem (9.24) can be written in the form

$$\operatorname*{Max}_{\lambda\geq 0}\left\{\lambda^T y + \inf_{x\in X} L(x,\lambda)\right\}. \tag{9.25}$$

Note that we changed the notation from y^* to λ in order to emphasize that the above problem (9.25) is the standard Lagrangian dual of (9.24) with λ being vector of Lagrange multipliers. The results of Propositions 55 and 56 can be applied to problem (9.24) and its dual (9.25) in a straightforward way.

As another example consider the problem

$$\begin{aligned} &\operatorname{Min}_{x\in X} f(x) \\ &\text{subject to} \quad g_i(x) + h_i(z) \leq 0, \quad i = 1,\ldots,m, \end{aligned} \tag{9.26}$$

where X is a convex subset of $\mathbb{R}^n$, $f(x)$ and $g_i(x)$ are real valued convex functions, and $h_i(z)$ are real valued convex functions of the parameter vector $z \in \mathbb{R}^\ell$. By change of variables $y_i = h_i(z)$, the above problem can be reduced to the problem (9.24), and hence the optimal value $v(z)$ of problem (9.26) is equal to $\vartheta(H(z))$, where $\vartheta(\cdot)$ is the optimal value of problem (9.24) and $H(z) := (h_1(z),\ldots,h_m(z))$. Note that if $y \leq y'$, then the feasible set of problem (9.24) corresponding to y' is included in the feasible set corresponding to y, and hence $\vartheta(y) \leq \vartheta(y')$, i.e., $\vartheta(\cdot)$ is componentwise nondecreasing function. It follows that the optimal value function $v(z)$ is convex, which is also not difficult to show directly.

Suppose that the functions $h_i(z)$, $i = 1,\ldots,m$, are differentiable and consider a point $\bar{z} \in \mathbb{R}^\ell$. Then, since $\vartheta(\cdot)$ is componentwise nondecreasing and, by convexity of $h_i(z)$, $H(\bar{z}+z)$ is componentwise greater than or equal to $H(\bar{z}) + \nabla H(\bar{z})z$ for any $z \in \mathbb{R}^\ell$, we have that for $\bar{y} := H(\bar{z})$ and any $y^* \in \partial\vartheta(\bar{y})$ the following inequalities hold

$$v(\bar{z}+z) = \vartheta(H(\bar{z}+z)) \geq \vartheta(H(\bar{z}) + \nabla H(\bar{z})z) \geq v(\bar{z}) + (y^*)^T \nabla H(\bar{z})z.$$

It follows that $\nabla H(\bar{z})^T y^* \in \partial v(\bar{z})$, or in other words that

$$\nabla H(\bar{z})^T \partial\vartheta(\bar{y}) \subset \partial v(\bar{z}). \tag{9.27}$$

As a consequence of the above inclusion we obtain that if $\vartheta(\cdot)$ is subdifferentiable at $\bar{y}$, then $v(\cdot)$ is subdifferentiable at $\bar{z}$. If, moreover, the constraint qualification

$$0 \in \operatorname{int}\{H(\bar{z}) + \nabla H(\bar{z})\mathbb{R}^\ell - \operatorname{dom}\vartheta\} \tag{9.28}$$

is satisfied, then it is possible to show that the inverse of the inclusion (9.28) also holds, and hence in that case

$$\nabla H(\bar{z})^T \partial\vartheta(\bar{y}) = \partial v(\bar{z}). \tag{9.29}$$

The constraint qualification (9.28) holds, in particular, if the Jacobian matrix $\nabla H(\bar{z})$ has full row rank m, or if $\bar{y} = H(\bar{z})$ belongs to the interior of the domain of ϑ. Note that $\operatorname{dom}\vartheta$ is formed by such vectors y that the corresponding problem (9.24) is feasible.

9.3 Probability and measure

It is said that an m-dimensional random vector $V = V(\omega)$ has an absolutely continuous distribution if for any set $\mathcal{A} \subset \mathbb{R}^m$ of Lebesgue measure zero the event $\{V \in \mathcal{A}\}$ has zero probability. The distribution of V is absolutely continuous iff it has a density, i.e., there exists a real valued function $g(\cdot)$ such that for any Borel set $\mathcal{A} \subset \mathbb{R}^m$ probability of the event $\{V \in \mathcal{A}\}$ is equal to the integral $\int_{\mathcal{A}} g(x)\,\mathrm{d}x$. The function $g(\cdot)$ is called the *probability density function*.

Let $f_n(\omega)$ be a sequence of real valued measurable functions on a probability space $(\Omega, \mathcal{F}, P)$. By $f_n \uparrow f$ a.e. we mean that for almost every $\omega \in \Omega$ the sequence $f_n(\omega)$ is monotonically nondecreasing and hence converges to a limit denoted $f(\omega)$, where $f(\omega)$ can be equal to $+\infty$.

Theorem 57 (Monotone Convergence Theorem). *Suppose that $f_n \uparrow f$ a.e. and there exists a P-integrable function $g(\omega)$ such that $f_n(\cdot) \geq g(\cdot)$. Then $\int_\Omega f\,\mathrm{d}P$ is well defined and $\int_\Omega f_n\,\mathrm{d}P \uparrow \int_\Omega f\,\mathrm{d}P$.*

Theorem 58 (Fatou's lemma). *Suppose that there exists a P-integrable function $g(\omega)$ such that $f_n(\cdot) \geq g(\cdot)$. Then*

$$\int_\Omega \liminf_{n\to\infty} f_n\,\mathrm{d}P \leq \liminf_{n\to\infty} \int_\Omega f_n\,\mathrm{d}P. \tag{9.30}$$

Theorem 59 (Lebesgue Dominated Convergence Theorem). *Suppose that there exists a P-integrable function $g(\omega)$ such that $|f_n| \leq g$ a.e., and that $f_n(\omega)$ converges to $f(\omega)$ for almost every $\omega \in \Omega$. Then $\int_\Omega f_n \, dP$ is well defined and $\int_\Omega f_n \, dP \to \int_\Omega f \, dP$.*

The following result can be proved by induction in m.

Theorem 60 (Richter–Rogosinsky). *Let $(\Omega, \mathcal{F})$ be a measurable space, $f_1, \ldots, f_m$ be measurable on $(\Omega, \mathcal{F})$ real valued functions, and μ be a nonnegative finite measure on $(\Omega, \mathcal{F})$ such that $f_1, \ldots, f_m$ are μ-integrable. Suppose that every finite subset of Ω is $\mathcal{F}$-measurable. Then there exists a nonnegative measure η on $(\Omega, \mathcal{F})$ with a finite support of at most m points such that $\int_\Omega f_i d\mu = \int_\Omega f_i d\eta$ for all $i = 1, \ldots, m$.*

10 Bibliographic notes

Basic properties of expectations, defined as integrals on probability spaces, can be found in numerous books on probability and measure theory. For example, we may refer to Billingsley (1995) where the interested reader can find proofs of the Lebesgue and Monotone Convergence Theorems and Fatou's Lemma. Differentiability properties of the expectation functions were discussed by many authors. Proposition 2 follows easily from the Lebesgue Dominated Convergence Theorem (cf., e.g., Rubinstein and Shapiro (1993)). For a thorough development of integration of multifunctions and integral functionals we refer to Rockafellar (1976) and Chapter 8 in Ioffe and Tihomirov (1979), where some additional references can be found. The interchangeability formula (1.8) for the subdifferential and expectation operators, given in Theorem 9, is taken from Theorem 4, p. 351, of Ioffe and Tihomirov (1979). In the convex case it follows from the interchangeability formula (1.8) that the expected value function $f(x)$ is differentiable at a point x_0 iff $\partial F(x_0, \omega)$ is a singleton for a.e. $\omega \in \Omega$. We derived this result in a more direct way in Proposition 4.

Properties of the optimal value $Q(x, \xi)$ of the second stage linear programming problem were studied by Walkup and Wets (1967, 1969), Wets (1966, 1974) and Kall (1976), so most of the material of Sections 2.1 and 2.2 can be found there. Example 15 is discussed in Birge and Louveaux (1997). Polyhedral and convex two-stage problems, discussed in Sections 2.3 and 2.4, are natural extensions of the linear two-stage problems. The conjugate duality, used in Proposition 25 is based on Fenchel duality (Fenchel (1953)) and was developed by Rockafellar (1974).

Optimality conditions of the type used in Proposition 32 are well known (see, e.g., Chapter 1 in Ioffe and Tihomirov (1979)). See also Hiriart-Urruty (1978) and Flåm (1992, 1995) for the analysis in the case of Lipschitz continuous functions.

Duality analysis of stochastic problems, and in particular dualization of the nonanticipativity constraints was developed by Eisner and Olsen (1975), Wets (1976), Rockafellar and Wets (1975, 1976a,b,c, 1977) (see also Wets (1980) and Klein Haneveld (1986) and Rockafellar (1999) and the references therein). We tried to keep the presentation in Section 6 and Section 7 relatively elementary without an extensive use of functional analysis.

The min–max approach to stochastic programming, discussed in Section 8, was investigated extensively by Dupačová (1977, 1978, 1987).

For a thorough treatment of the convex analysis theory we refer to Rockafellar (1970). Theorem 53 is due to Hoffman (1952). For a proof of Richter-Rogosinsky Theorem (theorem 60) see Rogosinsky (1958).

References

Billingsley, P. (1995). *Probability and Measure*, John Wiley & Sons, New York.

Birge, J.R., F. Louveaux (1997). *Introduction to Stochastic Programming*, Springer-Verlag, New York, NY.

Dupačová, J. (1977). The minimax problem of stochastic linear programming and the moment problem. *Ekonom.-Mater. Obzor* 13, 279–307.

Dupačová, J. (1978). Minimax approach to stochastic linear programming and the moment problem. Selected results. *Z. Angew. Math. Mech.* 58, T466–T467.

Dupačová, J. (1987). The minimax approach to stochastic programming and an illustrative application. *Stochastics* 20, 73–88.

Eisner, M., P. Olsen (1975). Duality for stochastic programming interpreted as l.p. in Lp-space. *SIAM J. Appl. Math.* 28, 775–792.

Fenchel, W. (1953). *Convex Cones, Sets, and Functions, lecture notes*, Princeton University.

Flåm, S.D. (1992). Lagrange multipliers in stochastic programming. *SIAM Journal on Control and Optimization* 30, 1–10.

Flåm, S.D. (1995). Corrigendum: Lagrange multipliers in stochastic programming. *SIAM Journal on Control and Optimization* 33, 667–671.

Hiriart-Urruty, J.-B. (1978). Conditions nécessaires d'optimalité pour un programme stochastique avec recours. *SIAM Journal on Control and Optimization* 16, 317–329.

Hoffman, A. (1952). On approximate solutions of systems of linear inequalities. *Journal of Research of the National Bureau of Standards, Section B. Math. Sci.* 49, 263–265.

Ioffe, A.D., V.M. Tihomirov (1979). *Theory of Extremal Problems*, North–Holland, Amsterdam.

Kall, P. (1976). *Stochastic Linear Programming*, Springer-Verlag, Berlin.

Klein Haneveld, W.K. (1986). *Duality in Stochastic Linear and Dynamic Programming, Lecture Notes in Economic and Mathematical Systems*, Vol. 274, Springer-Verlag, New York.

Rockafellar, R.T. (1970). *Convex Analysis*, Princeton University Press, Princeton, NJ.

Rockafellar, R.T. (1974). Conjugate Duality and Optimization, Regional Conference Series in Applied Mathematics, SIAM, Philadelphia.

Rockafellar, R.T. (1976). Integral functionals, normal integrands and measurable selections. *Nonlinear Operators and the Calculus of Variations, Lecture Notes in Mathematics*, Vol. 543, Springer-Verlag, Berlin, pp. 157–207.

Rockafellar, R.T. (1999). Duality and optimality in multistage stochastic programming. *Annals of Operations Research* 85, 1–19.

Rockafellar, R.T., R.J-B. Wets (1975). Stochastic convex programming: Kuhn-Tucker conditions. *J. Math. Econom.* 2, 349–370.

Rockafellar, R.T., R.J.-B. Wets (1976a). Stochastic convex programming: basic duality. *Pacific J. Math.* 62, 173–195.

Rockafellar, R.T., R.J.-B. Wets (1976b). Stochastic convex programming: singular multipliers and extended duality, singular multipliers and duality. *Pacific J. Math.* 62, 507–522.

Rockafellar, R.T., R.J.-B. Wets (1976c). Stochastic convex programming: relatively complete recourse and induced feasibility. *SIAM Journal on Control and Optimization* 14, 574–589.

Rockafellar, R.T., R.J-B. Wets (1977). Measures as Lagrange multipliers in multistage stochastic programming. *J. Math. Anal. Appl.* 60, 301–313.

Rogosinsky, W.W. (1958). Moments of non-negative mass. *Proc. Roy. Soc. London Ser. A* 245, 1–27.

Rubinstein, R.Y., A. Shapiro (1993). *Discrete Event Systems: Sensitivity Analysis and Stochastic Optimization by the Score Function Method*, John Wiley & Sons, Chichester.

Walkup, D., R.J.-B. Wets (1967). Stochastic programs with recourse. *SIAM J. Appl. Math.* 15, 1299–1314.

Walkup, D., R.J.-B. Wets (1969). Stochastic programs with recourse II: on the continuity of the objective. *SIAM J. Appl. Math.* 15, 1299–1314.

Wets, R.J.-B. (1966). Programming under uncertainty: the equivalent convex program. *SIAM J. Appl. Math.* 14, 89–105.

Wets, R.J.-B. (1974). Stochastic programs with fixed recourse: the equivalent deterministic program. *SIAM Review* 16, 309–339.

Wets, R.J.-B. (1976). Duality relations in stochastic programming, in: *Symposia Mathematica*, Vol. XIX (Convegno sulla Programmazione Matematica e sue Applicazioni, INDAM, Rome, 1974), Academic Press, London, pp. 341–355.

Wets, R.J.-B. (1980). Stochastic multipliers, induced feasibility and nonanticipativity in stochastic programming: *Stochastic Programming* (*Proc. Internat. Conf., Univ. Oxford, Oxford, 1974*), Academic Press, London, pp. 137–146.

A. Ruszczyński and A. Shapiro, Eds., *Handbooks in OR & MS, Vol. 10*

Chapter 3

Decomposition Methods

Andrzej Ruszczyński

Department of Management Science and Information Systems, Rutgers University, 94 Rockafeller Rd, Piscataway, NJ 08854, USA

Abstract

Two- and multistage stochastic programming problems have very large dimension and characteristic structures which are tractable by decomposition. We present cutting plane methods, nested decomposition methods, regularized decomposition methods, trust region methods, augmented Lagrangian methods, and splitting methods for convex stochastic programming problems.

Key words: Stochastic programming, decomposition, primal methods, dual methods, operator splitting.

1 Introduction

Two- and multistage stochastic programming problems have very specific structures which can be exploited by decomposition. The objective of this chapter is to provide a detailed description and analysis of the main decomposition methods used in stochastic programming.

To illustrate the main concepts, let us consider the two-stage stochastic programming problem, which has been extensively analyzed in Chapters 1 and 2. Recall that there are two groups of decision variables in the two-stage problem: the first stage decisions $x \in \mathbb{R}^{n_1}$, which are deterministic, and the second stage decisions $y \in \mathbb{R}^{n_2}$, which are allowed to depend on the random problem data, ξ. The linear two-stage problem has the form

$$\begin{aligned} &\text{Min } \{c^T x + \mathbb{E}Q(x, \xi)\} \\ &\text{s.t.} \quad Ax = b,\ x \geq 0, \end{aligned} \tag{1.1}$$

where $\xi = (q, W, h, T)$ is the vector of (possibly random) problem data, and

$$Q(x, \xi) := \inf_{y \in \mathbb{R}^{n_2}_+} \{q^T y \mid Wy = h - Tx\}. \tag{1.2}$$

We assume that the probability space Ω is finite, we denote by ω^s, $s = 1, \ldots, S$, all elementary events and by p_s, $s = 1, \ldots, S$, their probabilities. We also use the notation $T^s = T(\omega^s)$, $W^s = W(\omega^s)$, $h^s = h(\omega^s)$, $y^s = y(\omega^s)$, $q^s = q(\omega^s)$. Then we can rewrite (1.1) and (1.2) as

$$\operatorname*{Min}_{x \in X} \left\{ f(x) := c^T x + \sum_{s=1}^{S} p_s Q^s(x) \right\}, \tag{1.3}$$

with

$$Q^s(x) := \inf_{y \in \mathbb{R}^{n_2}_+} \{(q^s)^T y \mid W^s y = h^s - T^s x\} \tag{1.4}$$

and

$$X = \{x \in \mathbb{R}^{n_1} : Ax = b, \, x \geq 0\}.$$

The polyhedral structure of $Q^s(\cdot)$ is characterized in Proposition 11 in Chapter 2. To avoid unnecessary technical complications we assume that $Q^s(x) > -\infty$ for all $x \in X$.

The main idea of *primal decomposition methods* is to address the problem in its form (1.3). These methods solve many subproblems of form (1.4) to construct some models (approximations) of the recourse costs $Q^s(\cdot)$ and of the expected recourse cost. These models are used in a *master problem*, which generates approximations of the first stage solution, x. The differences between various primal methods lie mainly in the way the master problem is constructed and solved. We shall discuss several versions of primal methods in Sections 2–4. In Section 5 we shall extend these ideas to multistage problems.

Problem (1.3)–(1.4) can be rewritten as a large scale linear programming problem

$$\begin{aligned} \operatorname{Min} \; & \left\{ c^T x + \sum_{s=1}^{S} p_s (q^s)^T y^s \right\} \\ \text{s.t.} \quad & Ax = b, \\ & T^s x + W^s y^s = h^s, \quad s = 1, \ldots, S, \\ & x \geq 0, \, y^s \geq 0, \qquad s = 1, \ldots, S. \end{aligned} \tag{1.5}$$

Splitting the fist stage decisions x into copies $x^1,\ldots,x^S$ corresponding to scenarios we can reformulate (1.5) as follows:

$$\begin{aligned}
&\text{Min} \sum_{s=1}^{S} p_s(c^T x^s + (q^s)^T y^s) \\
&\text{s.t.} \quad Ax^s = b, && s = 1,\ldots,S, \\
&\qquad T^s x + W^s y^s = h^s, && s = 1,\ldots,S, \\
&\qquad x^s \geq 0,\ y^s \geq 0, && s = 1,\ldots,S, \\
&\qquad x^s = x^\sigma, && s,\sigma = 1,\ldots,S.
\end{aligned} \tag{1.6}$$

The idea of *dual methods* is to relax the last group of constraints, which are called in stochastic programming the *nonanticipativity constraints*, by assigning to them some Lagrange multipliers. In the most straightforward approach a multiplier $\lambda^{s,\sigma}$ is associated with each scenario pair[1] and we formulate the Lagrangian

$$L(x, y, \lambda) = \sum_{s=1}^{S} p_s\left(c^T x^s + (q^s)^T y^s\right) + \sum_{s=1}^{S}\sum_{\sigma=1}^{s} (\lambda^{s,\sigma})^T (x^s - x^\sigma).$$

The problem of minimizing the Lagrangian, subject to all the remaining constraints, splits into S independent *scenario subproblems*, each for vectors x^s and y^s. Their solutions depend on the multipliers and the role of the *master problem* is to find the optimal values of these multipliers, so that the x-parts of the scenario solutions will become identical. Dual decomposition methods differ in the ways the nonanticipativity constraints are spelled out, the ways the Lagrangian is defined, the subproblems constructed, and multipliers changed. We shall present them in Sections 6–9.

There are two fundamental advantages of decomposition methods. First, they replace a large and difficult stochastic programming problem by a collection of smaller and easier problems. This allows for solving extremely large models, which are intractable otherwise. Secondly, the subproblems involved in decomposition methods are usually standard linear, quadratic or nonlinear problems, which need to be developed and solved for the simplest deterministic versions of the model. Consequently, standard off-the-shelf optimization software may be used for the solution of these subproblems. As a result, decomposition methods provide a highly efficient and specialized methodology for solving very large and difficult stochastic programming problems by employing readily available tools.

[1] A smaller set of constraints can express nonanticipativity, but this form is sufficient to introduce the idea of dualization.

It should be stressed that the form of the constraints $Ax = b$ and $x \geq 0$ has been chosen only for the simplicity of presentation. We may have here arbitrary constraints of form $x \in X$, where X is a convex closed polyhedron. Similarly, instead of the conditions $y \geq 0$ in the second stage problems, we may require $y^s \in Y^s$, where Y^s is a convex closed polyhedron for each scenario $s = 1, \ldots, S$. Also, the constraints linking the first and the second stage decisions in (1.1) or (1.2) may have the form of an arbitrary combinations of linear equations and inequalities.

2 The cutting plane method

2.1 The main concepts

The idea of *cutting plane methods* is to construct a sequence $\{x^k\}$ of approximations to the solution of (1.3) in the following way. The known pieces of the functions $Q^s(\cdot)$ and facets of their domains are used to construct an approximation of the expected recourse cost $\sum_{s=1}^{S} p_s Q^s(x)$. This approximation is employed to compute the values of the first stage variables, x^k, at the current iteration. At these values of the first stage decisions subproblems (1.4) yield new information about the shape of $Q^s(\cdot)$. This is used to calculate x^{k+1}, and the iteration continues.

To see how the new information can be obtained, let us suppose that $Q^s(x^k) < \infty$. Then, as described in Proposition 12 of Chapter 2,

$$\partial Q^s(x^k) = -(T^s)^T \mathcal{D}^s(x^k), \tag{2.1}$$

where

$$\mathcal{D}^s(x^k) := \arg\max_{(W^s)^T \pi \leq q^s} \pi^T (h^s - T^s x^k) \tag{2.2}$$

is the set of optimal solutions of the dual to problem in (1.4) at $x = x^k$. Let $\pi^{k,s}$ be one of these optimal dual solutions. Then, by the definition of the subdifferential,

$$Q^s(x) \geq Q^s(x^k) - \left\langle (T^s)^T \pi^{k,s}, x - x^k \right\rangle \quad \text{for all} \quad x \in \mathbb{R}^{n_1}.$$

This inequality defines an *objective cut*:

$$Q^s(x) \geq \alpha^{k,s} + (g^{k,s})^T x, \quad \text{for all} \quad x \in \mathbb{R}^{n_1}, \tag{2.3}$$

where

$$g^{k,s} = -(T^s)^T \pi^{k,s}, \tag{2.4}$$

$$\alpha^{k,s} = Q^s(x^k) + (\pi^{k,s})^T T^s x^k = (h^s)^T \pi^{k,s}. \tag{2.5}$$

In the last equation (2.5) we have used the duality relation between (1.4) and the optimal value of problem (2.2):

$$Q^s(x^k) = (\pi^{k,s})^T (h^s - T^s x^k).$$

It follows that to obtain an objective cut for $Q^s(\cdot)$ at x^k we need to solve problem (1.4) at $x = x^k$, retrieve Lagrange multipliers $\pi^{k,s}$ associated with its constraints, and apply formulas (2.3)–(2.5). Equivalently, we may solve the dual problem appearing in (2.2). Moreover, if we restrict our attention to basic solutions of these problems, the multipliers $\pi^{k,s}$ will be chosen from a finite set which does not depend on x^k: the set of basic feasible solutions of the dual problem. We shall call such cuts *basic objective cuts*.

If subproblem (1.4) at $x = x^k$ is infeasible, we can derive an inequality that must be satisfied by every $x \in \operatorname{dom} Q^s$:

$$\beta^{k,s} + (r^{k,s})^T x \le 0, \tag{2.6}$$

and which is violated at x^k. We shall call it a *feasibility cut*, and we shall say that it cuts x^k off. To see how such a feasibility cut can be obtained, consider the Phase I problem corresponding to (1.4):

$$\begin{aligned} &\operatorname*{Min}_{y,z} \|z\| \\ &\text{s.t.} \quad W^s y + z = h^s - T^s x, \\ &\qquad y \ge 0. \end{aligned} \tag{2.7}$$

Here $z = (z_1, \ldots, z_m)$ is a vector of artificial variables and $\|\cdot\|$ denotes a norm on the space $\mathbb{R}^m$. For technical reasons we will use the ℓ_1 norm $\|z\|_1 := |z_1| + \cdots + |z_m|$, or the max-norm $\|z\|_\infty := \max\{|z_1|, \ldots, |z_m|\}$. Note that both norms $\|\cdot\|_1$ and $\|\cdot\|_\infty$ are polyhedral in the sense that they can be represented as a maximum of a finite number of linear functions. Let us denote by $U^s(x)$ the optimal value of (2.7). It is clear that problem (2.7) is always feasible and its optimal value is finite, and moreover $Q^s(x) < \infty$ if and only if $U^s(x) = 0$. We have, therefore, $U^s(x^k) > 0$.

The function $U^s(x)$ is an optimal value of a convex problem having x as a parameter in the constraint right hand side. Moreover, if the norm $\|\cdot\|$ is

polyhedral, then (2.7) can be formulated as a linear programming problem. The dual of (2.7) is given by the max–min problem

$$\max_{\eta} \min_{z,y\geq 0} \{\|z\| + \eta^T(h^s - T^s x - W^s y - z)\}.$$

By calculating the minimum in the above problem, we can write the dual in the following form

$$\begin{aligned} &\max_{\eta} \ \eta^T(h^s - T^s x) \\ &\text{s.t.} \quad (W^s)^T\eta \leq 0, \quad \|\eta\|^* \leq 1, \end{aligned} \tag{2.8}$$

where $\|\cdot\|^*$ denotes the dual of the norm $\|\cdot\|$ (recall that the norms $\|\cdot\|_1$ and $\|\cdot\|_\infty$ are dual to each other). We have, by the standard theory of linear programming, that in the case of a polyhedral norm $\|\cdot\|$ there is no duality gap between problems (2.7) and (2.8) and both problems have optimal solutions. Furthermore, the function $U^s(\cdot)$ is a piecewise linear convex function and its subdifferential $\partial U^s(x)$ is equal to $-(T^s)^T\Delta^s(x)$, where $\Delta^s(x)$ is the set of optimal solutions of the dual problem (2.8).

Therefore, we can construct for $U^s(\cdot)$ an objective cut $(\beta^{k,s}, r^{k,s})$ in a way similar to (2.3):

$$\begin{aligned} &r^{k,s} \in \partial U^s(x^k), \\ &\beta^{k,s} = U^s(x^k) - \langle r^{k,s}, x^k\rangle. \end{aligned}$$

Then for every $x \in \mathbb{R}^{n_1}$

$$U^s(x) \geq U^s(x^k) + \langle r^{k,s}, x - x^k\rangle = \beta^{k,s} + (r^{k,s})^T x,$$

and the above relation turns into equality at $x = x^k$. Since $U^s(x) = 0$ at all feasible x, the last inequality implies (2.6). Moreover, $U^s(x^k) > 0$, so (2.6) cuts x^k off, as promised. Again, by restricting the dual solutions to basic solutions we can guarantee that the number of all possible feasibility cuts is finite. We shall call such cuts *basic feasibility cuts*.

Summing up, the epigraph of $Q^s(\cdot)$ is a closed convex polyhedron defined by finitely many basic objective and feasibility cuts.

The objective cuts at x^k (if they can be computed successfully) yield a cut for the *expected* second stage cost

$$Q(x) = \sum_{s=1}^{S} p_s Q^s(x) \geq \overline{\alpha}^k + (\overline{g}^k)^T x, \tag{2.9}$$

where

$$\overline{g}^k = \sum_{s=1}^{S} p_s g^{k,s}, \quad \overline{\alpha}^k = \sum_{s=1}^{S} p_s \alpha^{k,s}. \tag{2.10}$$

Let $k = 1, 2, \ldots$ be the iteration number and let J^k_{obj} be the set of iteration numbers $j \le k$ such that the cut (2.10) could be calculated. Similarly, let $J^k_{\text{fea}}(s)$ be the set of iterations when a feasibility cut was generated for scenario s.

The cuts constructed so far are used in the *master problem*

$$\text{Min } \{c^T x + v\} \tag{2.11}$$

$$\text{s.t.} \quad \overline{\alpha}^j + (\overline{g}^j)^T x \le v, \quad j \in J^k_{\text{obj}}, \tag{2.12}$$

$$\beta^{j,s} + (r^{j,s})^T x \le 0, \quad j \in J^k_{\text{fea}}(s), \quad s = 1, \ldots, S, \tag{2.13}$$

$$Ax = b, \quad x \ge 0, \tag{2.14}$$

whose solution (x^{k+1}, v^{k+1}) is the next approximation to the solution of (1.3) and a lower bound for $Q(\cdot)$.

To describe the cutting plane method and analyze its properties, let us assume at first that the initial point, x^1, is such that $Q^s(x^1) < \infty$ for all $s = 1, \ldots, S$. We also assume that the set

$$X = \{x \in \mathbb{R}^{n_1} : Ax = b, x \ge 0\}$$

is bounded. The operation of the Cutting Plane Method is presented in Fig. 1.

Step 0. Set $k = 1$, $J^0_{\text{obj}} = \emptyset$, $J^0_{\text{fea}}(s) = \emptyset$, $s = 1, \ldots, S$, $v^1 = -\infty$.
Step 1. For $s = 1, \ldots, S$ solve subproblem (1.4) with $x = x^k$.

(a) If $Q^s(x^k) < \infty$, construct the objective cut (2.3) and set $J^k_{\text{fea}}(s) = J^{k-1}_{\text{fea}}(s)$;
(b) If $Q^s(x^k) = \infty$ (i.e., problem (1.4) is infeasible), construct the feasibility cut (2.6) and set $J^k_{\text{fea}}(s) = J^{k-1}_{\text{fea}}(s) \cup \{k\}$.

If $Q^s(x^k) < \infty$ for all $s = 1, \ldots, S$, construct the aggregate objective cut (2.9) and set $J^k_{\text{obj}} = J^{k-1}_{\text{obj}} \cup \{k\}$; otherwise set $J^k_{\text{obj}} = J^{k-1}_{\text{obj}}$.

Step 2. If $Q(x^k) = v^k$ then stop (optimal solution has been found); otherwise continue.
Step 3. Solve the master problem (2.11)–(2.14). If it is infeasible, stop (the original problem has no feasible solutions). Otherwise, denote by (x^{k+1}, v^{k+1}) its solution, increase k by one, and go to Step 1.

Fig. 1. The cutting plane algorithm.

2.2 Convergence

Let us denote by f^* the optimal value of the original two-stage problem (1.3), with the convention that $f^* = \infty$ if this problem is infeasible.

The key property of the master problem is that its optimal value provides a lower bound for the optimal value of the original problem. To see this, let us consider the function

$$\underline{Q}^k(x) := \max_{j \in J^k_{\text{obj}}} [\overline{\alpha}^j + (\overline{g}^j)^T x]. \tag{2.15}$$

By (2.9),

$$\underline{Q}^k(x) \le Q(x), \quad \text{for all } x \text{ and all } k = 1, 2, \ldots, \tag{2.16}$$

Let us assume that x is fixed in (2.11)–(2.14), so that the optimization is carried out with respect to the variable v. The optimal value of v is then clear from (2.12):

$$\hat{v}(x) = \max_{j \in J^k_{\text{obj}}} [\overline{\alpha}^j + (\overline{g}^j)^T x] = \underline{Q}^k(x).$$

It follows that the master problem is equivalent to minimizing $c^T x + \underline{Q}^k(x)$ subject to the constraints (2.13)–(2.14).

Let us also introduce the sets

$$X^k := \{x : \text{constraints (2.13) hold}\}. \tag{2.17}$$

By the construction of the feasibility cuts

$$X^k \supseteq \operatorname{dom} Q(\cdot), \quad k = 1, 2, \ldots \tag{2.18}$$

Consequently, the master problem (2.11)–(2.14) is equivalent to

$$\text{Min } \{c^T x + \underline{Q}^k(x)\} \tag{2.19}$$

$$\text{s.t.} \quad x \in X^k, \tag{2.20}$$

$$Ax = b,\ x \ge 0. \tag{2.21}$$

By virtue of (2.16) and (2.17) we have the following result.

Lemma 1. *If the master problem is feasible at iteration k, then* $c^T x^{k+1} + v^{k+1} \leq f^*$. *If the master problem is infeasible at iteration k, then* $f^* = +\infty$.

Consequently, the infeasibility test at Step 3 is correct.

We can now prove the convergence of the cutting plane algorithm in its simplest form.

Theorem 2. *Assume that the set X is bounded and that* $Q(x^1) < \infty$. *Moreover, let all cuts constructed at Step 1a and Step 1b be basic objective and feasibility cuts. Then after finitely many iterations the Cutting Plane Algorithm finds an optimal solution of* (1.3).

Proof. If $Q(x^k) = v^k$ then

$$c^T x^k + Q(x^k) = c^T x^k + \underline{Q}^k(x^k) \leq c^T x + \underline{Q}^k(x) \leq c^T x + Q(x)$$

for all $x \in X^k \cap X$. By virtue of (2.18), the point x^k is optimal for (1.3). If $v^k < Q(x^k) < \infty$, the new objective cut (2.9) cuts the point (x^k, v^k) off the set of feasible solutions of the master problem. If $Q^s(x^k) = \infty$ for some s, the new feasibility cut (2.6) cuts x^k off the set of feasible solutions of the master problem. In any case, if x^k is not optimal, Step 2 generates a new cut which is different from the cuts present in the master problem. Since the number of different basic objective and feasibility cuts is finite, the algorithm must stop. □

The assumption that the cuts are basic has been made only for simplicity. We can prove exactly the same result without it, and we shall do it right now. Let us recall that we assume that the set X is bounded. Also, to allow infeasible starting points we add the constraint $v \geq -M$ to (2.14), where $-M$ is a lower bound for the optimal cost.

Lemma 3. *For every* $s = 1, \ldots, S$ *the number of iterations for which* $U^s(x^k) > 0$ *is finite*.

Proof. Since $U^s(\cdot)$ is polyhedral and its domain is the whole $\mathbb{R}^{n_1}$, it may be represented as

$$U^s(x) = \max\left(0, \max_{j \in J}(\delta_j + (d_j)^T x)\right),$$

where J is a finite set. For each x we can define the set

$$I(x) := \{ j : \delta_j + (d_j)^T x = U^s(x) \};$$

(it may be empty when $U^s(x)=0$). Clearly, there may be only a finite number of different sets $I(x)$, and we shall denote them by $I_1, I_2, \ldots, I_L$. Each of the sets I_l defines a cell:

$$C_l := \{x\colon \delta_j + (d_j)^T x = U^s(x), j \in I_l;\ \delta_j + (d_j)^T x < U^s(x), j \in J\backslash I_l\}.$$

Suppose that our assertion is false, and $U^s(x^k)>0$ for infinitely many k. Let us consider an iteration k such that $U^s(x^k)>0$, and let C_l be the cell containing x^k. Then the feasibility cut generated at x^k,

$$U^s(x^k) + \langle r^{k,s}, x - x^k\rangle \le 0, \tag{2.22}$$

has

$$r^{k,s} = \sum_{j\in I_l} \lambda_j d_j, \quad \lambda_j \ge 0, \quad \sum_{j\in I_l} \lambda_j = 1.$$

For all $x \in C_l$ we have:

$$\begin{aligned} U^s(x) &= \sum_{j\in J_l} \lambda_j(\delta_j + (d_j)^T x) \\ &= \sum_{j\in J_l} \lambda_j(\delta_j + (d_j)^T x^k + (d_j)^T(x - x^k)) \\ &= U^s(x^k) + \langle r^{k,s}, x - x^k\rangle. \end{aligned}$$

By (2.22) no point x^m generated by the algorithm for $m>k$ may belong to the set $\{x \in C_l\colon U^s(x) > 0\}$. Since the number of cells is finite, we obtain a contradiction. □

We are now ready to state our main result.

Theorem 4. *After finitely many iterations the Cutting Plane Algorithm either discovers infeasibility or finds an optimal solution of* (1.3).

Proof. By Lemma 3, after finitely many iterations the method either discovers infeasibility, or continues without any new feasibility cuts added, i.e., $x^k \in \operatorname{dom} Q(\cdot)$ for all sufficiently large k.

Since $Q(\cdot)$ is polyhedral it may be represented as

$$Q(x) = \max_{j\in J}(\gamma_j + (u_j)^T x), \quad \text{for all} \quad x \in \operatorname{dom} Q,$$

where J is a finite set. For each x we can define the set

$$I(x) := \{ j: \gamma_j + (u_j)^T x = Q(x)\}.$$

There may be only a finite number of different sets $I(x)$, and we shall denote them by $I_1, I_2, \ldots, I_L$. Each of the sets I_l defines a cell:

$$C_l := \{x: \gamma_j + (u_j)^T x = Q(x), j \in I_l;\ \gamma_j + (u_j)^T x < Q(x), j \in J \backslash I_l\}.$$

Suppose that our assertion is false, and $f(x^k) > f^*$ for infinitely many k. Let $f(x^k) > f^*$ for some $x^k \in C_l$. If $x^k \in \operatorname{int} \operatorname{dom} Q$, the objective cut generated at x^k has the form

$$Q(x^k) + \langle \overline{g}^k, x - x^k \rangle \leq v, \tag{2.23}$$

with the subgradient

$$\overline{g}^k = \sum_{j \in I_l} \lambda_j u_j, \quad \lambda_j \geq 0, \quad \sum_{j \in I_l} \lambda_j = 1.$$

For all $x \in C_l \cap \operatorname{int} \operatorname{dom} Q$ we have:

$$\begin{aligned} Q(x) &= \sum_{j \in J_l} \lambda_j (\gamma_j + (u_j)^T x) \\ &= \sum_{j \in J_l} \lambda_j (\gamma_j + (u_j)^T x^k + (u_j)^T (x - x^k)) \\ &= Q(x^k) + \langle \overline{g}^k, x - x^k \rangle. \end{aligned}$$

By (2.23), for every $m > k$ such that $x^m \in C_l \cap \operatorname{int} \operatorname{dom} Q$ we must have

$$Q(x^m) \leq v^m,$$

so no new cut will be generated at x^m.

Let us now consider the case when $x^k \in C_l$ and x^k is a boundary point of $\operatorname{dom} Q$. Since $\operatorname{dom} Q$ is a convex closed polyhedron, it may be represented as

$$\operatorname{dom} Q = \{x: w_j^T x \leq \sigma_j, j \in J'\},$$

for some vectors w_j and constants σ_j. Its boundary is a finite collection of facets defined as

$$\Gamma_\nu := \{x: w_j^T x = \sigma_j, j \in J_\nu;\ w_j^T x < \sigma_j, j \in J' \backslash J_\nu\},$$

where J_ν, $\nu = 1,\ldots,M$, are all nonempty subsets of J' for which the above formula defines a nonempty set. Every subgradient of $Q(\cdot)$ at x^k has the form

$$\overline{g}^k = \sum_{j\in I_l} \lambda_j u_j + z^k, \quad \lambda_j \geq 0, \quad \sum_{j\in I_l} \lambda_j = 1.$$

where z^k is a normal vector to $\operatorname{dom} Q$ at x^k. By the definition of the normal vector

$$\langle z^k, x - x^k\rangle = 0, \quad \text{for all } x \in \Gamma_\nu.$$

Thus, for all $x \in C_l \cap \Gamma_\nu$ we have:

$$\begin{aligned} Q(x) &= \sum_{j\in J_l} \lambda_j(\gamma_j + (u_j)^T x) \\ &= \sum_{j\in J_l} \lambda_j(\gamma_j + (u_j)^T x^k + (u_j)^T(x - x^k)) \\ &= Q(x^k) + \langle \overline{g}^k - z^k, x - x^k\rangle \\ &= Q(x^k) + \langle \overline{g}^k, x - x^k\rangle. \end{aligned}$$

By (2.23), no new cuts will be generated at iteration $m > k$ such that $x^m \in C_l \cap \Gamma_\nu$.

Thus, after visiting finitely many cells and intersections of cells with facets the algorithm will satisfy the stopping test of Step 2. □

Simplicity is the main advantage of the Cutting Plane Method. However, the number of cuts in the master problem grows and there is no easy way to keep it bounded. A natural idea would be to drop inactive cuts, that is, these objective and feasibility cuts which are satisfied as sharp inequalities at the current solution (x^k, v^k) of the master problem.

If we use the method with basic cuts only (as discussed in Theorem 2), we may drop inactive cuts whenever the optimal value of the master program increases.

We note at first that when inactive cuts are deleted, no decrease of the optimal value of the master problem may result. Thus the sequence of optimal values of the master problem, $\{c^T x^k + v^k < f^*\}$ is monotone. To prove that it is convergent to f^*, suppose that $c^T x^k + v^k < f^*$ for some k and that no increase of the master's objective occurs for all $m > k$. Then no deletion takes place, and Theorem 2 guarantees the convergence of $c^T x^m + v^m$ to f^*, a contradiction. Therefore, an increase in the master's objective must occur at some $m \geq k$. The number of different optimal values of the master problem is finite, because there exists a finite number of different sets of basic objective

and feasibility cuts. Therefore, an increase in this value can occur only finitely many times, and the method must stop at an optimal solution after finitely many iterations.

In the version with arbitrary subgradient cuts we have no guaranteed finiteness of the sets of objective and feasibility cuts, and it is difficult to propose a useful and reliable rule for deleting inactive cuts.

Actually, in both cases, deleting *all* inactive cuts is not a good idea, because experience shows that many of them will have to be reconstructed.

2.3 The multicut version

Convergence properties of the Cutting Plane Method can be improved by using the objective cuts in their original form, without the averaging operation (2.9)–(2.10). For each s we define a lower approximation $\underline{Q}^{k,s}$ of Q^s as follows:

$$\underline{Q}^{k,s}(x) := \begin{cases} \max\limits_{j \in J^k_{\text{obj}}(s)} \left(\alpha^{j,s} + (g^{j,s})^T x\right), & \text{if } \beta^{j,s} + (r^{j,s})^T x \le 0,\ j \in J^k_{\text{fea}}(s), \\ +\infty, & \text{otherwise.} \end{cases} \tag{2.24}$$

Here $J^k_{\text{obj}}(s)$ is the subset of $\{1,\ldots,k\}$ corresponding to iterations at which cuts (2.3) were obtained. These models enter the master problem

$$\operatorname*{Min}_{x \in X} \left\{ c^T x + \sum_{s=1}^{S} p_s \underline{Q}^{k,s}(x) \right\}, \tag{2.25}$$

which is an approximation of (1.3) from below.

A more explicit form of (2.25), similar to (2.11)–(2.14), can be written as follows:

$$\operatorname{Min} \left\{ c^T x + \sum_{s=1}^{S} p_s v^s \right\} \tag{2.26}$$

$$\alpha^{j,s} + (g^{j,s})^T x \le v^s, \quad j \in J^k_{\text{obj}}(s), \quad s = 1,\ldots,S, \tag{2.27}$$

$$\beta^{j,s} + (r^{j,s})^T x \le 0, \quad j \in J^k_{\text{fea}}(s), \quad s = 1,\ldots,S, \tag{2.28}$$

$$Ax = b,\ x \ge 0. \tag{2.29}$$

The multicut method requires more memory, but it uses the previously-collected data more efficiently, because cuts from different scenarios can be combined in various ways.

Step 0. Set $k=1$, $J^0_{\text{obj}}(s)=\emptyset$, $J^0_{\text{fea}}(s)=\emptyset$, $v^{1,s}=-\infty$, $s=1,\dots,S$.
Step 1. For $s=1,\dots,S$ solve subproblem (1.4) with $x=x^k$.

(a) If $Q^s(x^k)<\infty$ then set $J^k_{\text{fea}}(s)=J^{k-1}_{\text{fea}}(s)$. If $Q^s(x^k)>v^{k,s}$ then construct the objective cut (2.3) and set $J^k_{\text{obj}}(s)=J^{k-1}_{\text{obj}}(s)\cup\{k\}$; otherwise set $J^k_{\text{obj}}(s)=J^{k-1}_{\text{obj}}(s)$.
(b) If $Q^s(x^k)=\infty$ (i.e., problem (1.4) is infeasible), construct the feasibility cut (2.6) and set $J^k_{\text{obj}}(s)=J^{k-1}_{\text{obj}}(s)$, $J^k_{\text{fea}}(s)=J^{k-1}_{\text{fea}}(s)\cup\{k\}$.

Step 2. If $Q(x^k)=\sum_{s=1}^S p_s v^{k,s}$ then stop (optimal solution has been found); otherwise continue.
Step 3. Solve the master problem (2.26)–(2.29). If it is infeasible, stop (the original problem has no feasible solutions). Otherwise, denote by (x^{k+1},v^{k+1}) its solution, increase k by one, and go to Step 1.

Fig. 2. The multicut algorithm.

The algorithm is presented in detail in Fig. 2. To allow starting the method without any cuts available, we may add the constraint $v\geq -M$ to the master, where $-M$ is a lower bound for the optimal cost.

Theoretical convergence properties of the Multicut Method are exactly the same as the properties of the Cutting Plane Method. Theorem 2 remains valid, and its proof is the same. Lemma 3 is also true, because the feasibility cuts are used in the same way in both versions. To prove the analogue of Theorem 4 we need only technical adjustments.

Theorem 5. *After finitely many iterations the Multicut Algorithm either discovers infeasibility or finds an optimal solution of* (1.3).

Proof. Since each $Q^s(\cdot)$ is polyhedral it may be represented as

$$Q^s(x)=\max_{j\in J^s}(\gamma^s_j+(u^s_j)^T x),\quad \text{for all } x\in \operatorname{dom} Q^s,$$

where J^s is a finite set. For each x and every s we define

$$I^s(x):=\{j\in J^s\colon \gamma^s_j+(u^s_j)^T x=Q^s(x)\}.$$

There may be only a finite number of different sets $I^s(x)$, and we shall denote them by $I^s_1,I^s_2,\dots,I^s_{L^s}$. Each of the sets I^s_l defines a cell:

$$C^s_l:=\{x\colon \gamma^s_j+(u^s_j)^T x=Q^s(x),\ j\in I^s_l;\ \gamma^s_j+(u^s_j)^T x<Q^s(x),\ j\in J^s\backslash I^s_l\}.$$

Next, similarly to the proof of Theorem 4, let Γ_ν, $\nu=1,\dots,M$, be the facets of $\operatorname{dom} Q$, and let $\Gamma_0=\operatorname{int}\operatorname{dom} Q$.

Let $f(x^k) > f^*$ for some x^k in the intersection $C^1_{l_1} \cap C^2_{l_2} \cap \cdots \cap C^S_{l_S} \cap \Gamma_\nu$. Thus, for at least one s we must have $Q^s(x^k) > v^{k,s}$. Proceeding exactly like in the proof of Theorem 4 we can prove that for every $m > k$ such that x^m is in the same intersection of cells and a facet, we have

$$Q^s(x^m) \leq v^{m,s}.$$

Consequently, the intersection $C^1_{l_1} \cap C^2_{l_2} \cap \cdots \cap C^S_{l_S} \cap \Gamma_\nu$ may be visited at most S times.

Since there are finitely many possible intersections of cells and a facet and each of them may be visited only finitely many times, the algorithm will satisfy the stopping test of Step 2. □

Our remarks about deleting inactive cuts made at the end of the preceding section remain valid for the Multicut Method. If basic cuts are used, inactive cuts may be removed whenever the optimal value of the master problem increases. If general subgradient cuts are employed, no easy rule can be found.

2.4 *Estimating objective cuts*

So far we have assumed that the number of elementary events ω in (1.1)–(1.2) is finite, and that we are able to solve all possible realizations of subproblem (1.2) for $\omega \in \Omega$. In many cases these assumptions are not satisfied. For example, if the right hand side in (1.2) has m independent components and each may have r different realizations, the number of elementary events equals $S = r^m$, which is a very large number, even for moderate values of r and m. A good example here is the capacity expansion problem of Chapter 1, Section 2, in which the number of possible scenarios is so large that the solution of all possible realizations of the second stage problem is practically impossible.

There are two ways around this difficulty. One is to generate a *sample* $\xi^1, \ldots, \xi^N$ of the problem data, having cardinality N which is manageable. Then we can solve the two-stage problem with this sample rather than with the true distribution of the data, hoping that the sample is representative enough to lead to a good approximation of the true solution. In Chapter 5 we discuss in detail the properties of this method.

The second approach is to work with *estimated* cuts rather than with exact cuts at each iteration of the Cutting Plane Method. Let us illustrate this approach on the case of a finite sample space Ω of very large cardinality S. The main difficulty is then the necessity to solve at each iteration of the method S subproblems (1.4), corresponding to data realizations $\xi^1, \ldots, \xi^S$. These solutions are needed to calculate the objective value at x^k:

$$Q(x^k) = \sum_{s=1}^{S} p_s Q(x^k, \xi^s) \tag{2.30}$$

and the objective cut (2.9)–(2.10):

$$Q(x) \geq \overline{\alpha}^k + (\overline{g}^k)^T x, \tag{2.31}$$

where

$$\overline{g}^k = \sum_{s=1}^{S} p_s g^{k,s}, \quad \overline{\alpha}^k = \sum_{s=1}^{S} p_s \alpha^{k,s}. \tag{2.32}$$

Let us assume that the second stage problem (1.4) is solvable for all x satisfying the first stage constraints and for all $\xi^1, \ldots, \xi^S$, that is, we deal with a problem with relatively complete recourse.

In order to estimate the quantities defined by (2.30)–(2.32), at iteration k we draw independent realizations $\xi^1, \ldots, \xi^N$ of random problem data, which will be, in general, different at each iteration. They correspond to a random sample $s^1, \ldots, s^N$ of the scenario numbers: $\xi^\nu = \xi^{s^\nu}$.

Then we solve (1.4) only for these sampled scenarios and we construct the Monte Carlo estimates

$$\tilde{Q}(x^k) := \frac{1}{N} \sum_{\nu=1}^{N} Q(x^k, \xi^\nu),$$

$$\tilde{g}^k := \frac{1}{N} \sum_{\nu=1}^{N} g^{s^\nu}. \tag{2.33}$$

Unfortunately, for reasonable sample sizes N the variance of these estimates may be large. The chapter on Monte Carlo techniques will discuss in much detail methods to deal with this difficulty, such as the *importance sampling* method. Here we provide only a superficial characterization of this approach.

Let $\tilde{\mathbb{P}}$ be another probability distribution on Ω, that is, let $\tilde{p}_s$ be some new probabilities assigned to the scenarios $\xi^1, \ldots, \xi^S$. We assume that the recourse cost $Q(x, \xi)$ is nonnegative for feasible first stage decisions.[2] We can rewrite the expression for the expected value of the second stage cost as follows:

$$\overline{Q}(x) = \sum_{s=1}^{S} \tilde{p}_s [Q(x, \xi^s) p_s / \tilde{p}_s]. \tag{2.34}$$

[2] We can treat in a similar way any recourse cost that is bounded from below.

This formula can be interpreted as the expectation of the function in brackets with respect to the distribution $\tilde{\mathbb{P}}$. The Monte Carlo estimate of (2.34) takes on the form

$$\tilde{Q}(x) = \frac{1}{N} \sum_{\nu=1}^{N} Q(x, \xi^{\nu}) p_{\nu} / \tilde{p}_{\nu}, \tag{2.35}$$

where $\xi^1, \ldots, \xi^N$ are independent observations drawn according to the new distribution $\tilde{\mathbb{P}}$. We can now choose $\tilde{\mathbb{P}}$ to decrease the variability of $Q(x, \xi^s) p_s / \tilde{p}_s$. In fact, setting $\tilde{p}_s = p_s Q(x, \xi^s) / \overline{Q}(x)$, we can make this function constant with respect to s, and the Monte Carlo estimate (2.35)—perfect for any N. This, however, involves the expected cost that we want to compute. Still, our hypothetical construction suggests a way to proceed: use

$$\tilde{p}_s = p_s W(x, \xi^s) / \overline{W}(x), \tag{2.36}$$

with some approximation W of Q. The approximation should be, of course, nonnegative, easy to compute, and easy to integrate. Clearly, the way in which such an approximation may be constructed depends on the structure of the original problem. However, there is a big danger of numerical instability here (see the chapter on Monte Carlo methods).

In the derivation of $\tilde{\mathbb{P}}$ we paid much attention to the accuracy of the recourse cost estimation (2.35). The same probabilities, though, and the same observations ξ^{ν} can be used to construct a subgradient estimate $\tilde{g}^k = (1/N)((\sum_{\nu=1}^{N} g_{\nu}^k p_{\nu}) / \tilde{p}_{\nu})$, which is consistent with (2.35).

2.5 *Extension to convex problems*

The Cutting Plane Method can be extended to two-stage problems with convex objectives of the following form

$$\operatorname*{Min}_{x \in X} \left\{ f(x) := f_1(x) + \sum_{s=1}^{S} p_s Q^s(x) \right\}, \tag{2.37}$$

where $Q^s(x)$ is the optimal value of the second stage problem

$$\operatorname*{Min}_{y} f_2^s(y) \quad \text{subject to} \quad T^s x + W^s y = h^s, \; y \in Y^s. \tag{2.38}$$

We assume that the functions $f_1 : \mathbb{R}^{n_1} \to \mathbb{R}$ and $f_2^s : \mathbb{R}^{n_2} \to \mathbb{R}$, $s = 1, \ldots, S$, are convex and the sets X and Y^s are bounded convex polyhedra. Our model (2.37)–(2.38) is a special case of the model (2.50)–(2.51) from Chapter 2, but

with linear constraints linking the first and the second stage variables, and with polyhedral domains of the objectives at both stages. Since the sets Y^s are bounded, for every x the second stage problem (2.38) either has an optimal solution or is infeasible.

The dual to the second stage problem (2.38) has the form

$$\operatorname*{Max}_{\pi} \left\{ \pi^T (h^s - T^s x) - (\overline{f}_2^{\,s})^* \big((W^s)^T \pi \big) \right\}, \tag{2.39}$$

where $(\overline{f}_2^{\,s})^*$ is the convex conjugate of the second stage objective,

$$\overline{f}_2^{\,s}(y) := \begin{cases} f_2^s(y) & \text{if } y \in Y^s \\ +\infty & \text{otherwise.} \end{cases}$$

We refer the reader to Chapter 2, Sections 2.3 and 2.4, for the derivation of the dual problem. By the duality theory of convex programming we have that if, for some x, problem (2.38) has a solution, then the dual problem (2.39) has a solution, and the optimal values of both problems are equal.[3]

By Proposition 25 of Chapter 2, each function $Q^s(\cdot)$ is convex and lower semicontinuous. Proposition 26 of Chapter 2 provides us with the general form of a subgradient of $Q^s(x)$ at points x at which (2.38) has a solution. Due to the linearity of the constraints and the finiteness of f_2^s we do not need additional constraint qualification conditions and we get

$$\partial Q^s(x) = -(T^s)^T \mathcal{D}^s(x), \tag{2.40}$$

where $\mathcal{D}^s(x)$ is the set of solutions to the dual problem (2.39). Of course, the elements of $\mathcal{D}^s(x)$ are the values of Lagrange multipliers associated with the constraints of (2.38).

It follows that at every point x^k at which the second stage problem (2.38) has a solution, we can construct an objective cut in a way similar to (2.3):

$$Q^s(x) \geq \alpha^{k,s} + (g^{k,s})^T x, \quad \text{for all } x \in \mathbb{R}^{n_1}, \tag{2.41}$$

where

$$g^{k,s} = -(T^s)^T \pi^{k,s},$$
$$\alpha^{k,s} = Q^s(x^k) + (\pi^{k,s})^T T^s x^k,$$

[3] By the polyhedrality of the feasible sets and by the finiteness of f_2^s we do not need additional constraint qualification conditions here.

and $\pi^{k,s} \in \mathcal{D}^s(x^k)$. We can also calculate a cut for $f_1(\cdot)$ at x^k:

$$f_1(x) \geq \alpha^{k,0} + (g^{k,0})^T x, \quad \text{for all } x \in \mathbb{R}^{n_1},$$

where

$$g^{k,0} \in \partial f_1(x^k),$$
$$\alpha^{k,0} = f_1(x^k) - (g^{k,0})^T x^k.$$

All these objective cuts at x^k (if they can be computed successfully) yield a cut for the overall objective

$$f(x) \geq \overline{\alpha}^k + (\overline{g}^k)^T x, \tag{2.42}$$

where

$$\overline{g}^k = g^{k,0} + \sum_{s=1}^{S} p_s g^{k,s}, \quad \overline{\alpha}^k = \alpha^{k,0} + \sum_{s=1}^{S} p_s \alpha^{k,s}. \tag{2.43}$$

If the second stage problem (2.38) is infeasible, we can derive the feasibility cut (2.6) exactly as in the linear case, just with the constraint $y \in Y^s$ instead of $y \geq 0$ in the Phase I problem (2.7).

The Cutting Plane Method for convex problems is almost identical with the method for linear problems. It has the master problem

$$\begin{aligned} &\text{Min } v \\ &\text{s.t. } \quad \overline{\alpha}^j + (\overline{g}^j)^T x \leq v, \quad j \in J^k_{\text{obj}}, \\ &\qquad \beta^{j,s} + (r^{j,s})^T x \leq 0, \quad j \in J^k_{\text{fea}}(s), \quad s = 1, \ldots, S, \\ &\qquad x \in X, \end{aligned} \tag{2.44}$$

whose solution is denoted (x^{k+1}, v^{k+1}). The detailed algorithm is presented in Fig. 3.

To allow starting from infeasible points, we may add the constraint $v \geq -M$ to (2.44), where $-M$ is a lower bound for the optimal cost.

To carry out the convergence analysis of the Cutting Plane Method in the convex case, we shall need an additional assumption.

Assumption 6. There exists a constant C such that $\|g^{k,s}\| \leq C$ for all $k = 1, 2, \ldots$ and all $s = 1, \ldots, S$, whenever $Q^s(x^k) < \infty$.

Step 0. Set $k=1$, $J^0_{\text{obj}}=\emptyset$, $J^0_{\text{fea}}(s)=\emptyset$, $s=1,\dots,S$, $v^1=-\infty$.
Step 1. For $s=1,\dots,S$ solve subproblem (2.38) with $x=x^k$.

(a) If $Q^s(x^k)<\infty$, construct the objective cut (2.41) and set $J^k_{\text{fea}}(s)=J^{k-1}_{\text{fea}}(s)$;
(b) If $Q^s(x^k)=\infty$ (i.e., problem (2.38) is infeasible), construct the feasibility cut (2.6) and set $J^k_{\text{fea}}(s)=J^{k-1}_{\text{fea}}(s)\cup\{k\}$.

If $Q^s(x^k)<\infty$ for all $s=1,\dots,S$, construct the aggregate objective cut (2.42) and set $J^k_{\text{obj}}=J^{k-1}_{\text{obj}}\cup\{k\}$; otherwise set $J^k_{\text{obj}}=J^{k-1}_{\text{obj}}$.

Step 2. If $Q(x^k)=v^k$ then stop (optimal solution has been found); otherwise continue.
Step 3. Solve the master problem (2.44). If it is infeasible, stop (the original problem has no feasible solutions). Otherwise, denote by (x^{k+1}, v^{k+1}) its solution, increase k by one, and go to Step 1.

Fig. 3. The cutting plane algorithm for convex problems.

Due to the boundedness of the sets X and Y^s, the linearity of second stage constraints and the finiteness of $f^s_2(\cdot)$, this assumption can always be satisfied. Indeed, we already know that $Q^s(\cdot)$ is lower semicontinuous and subdifferentiable at every point of its domain. Moreover, dom Q^s is a compact set, and the existence of bounded subgradients follows from general properties of convex functions. We can show it in our case directly, by showing the existence of uniformly bounded Lagrange multipliers $\pi^{k,s}$ at the solutions to the dual problem (2.38).

Theorem 7. *If problem (2.37)–(2.38) has no feasible solutions the Cutting Plane Method will stop at Step 3 after finitely many iterations. If problem (2.37)–(2.38) has feasible solutions then the Cutting Plane Method either stops at Step 2 at an optimal solution, or generates a sequence of points $\{x^k\}$ such that*

$$\lim_{k\to\infty} f(x^k)=f^*.$$

Proof. Since the master problem is a relaxation of (2.37), if the method stops at Step 3, the original problem is infeasible. Also, we always have $v^k\le f^*$, so the method can stop at Step 2 only if x^k is optimal. It remains to analyze the case of infinitely many steps.

The construction and the use of feasibility cuts is the same as in the linear case, and Lemma 3 remains valid. Thus, if the problem has no feasible solutions, the method will discover this after finitely many iterations. Moreover, if feasible points exist and the method does not stop at an optimal solution, we shall have $f(x^k)<\infty$ for all sufficiently large k.

For $\varepsilon>0$ we define

$$\mathcal{K}_\varepsilon=\{k: f^*+\varepsilon<f(x^k)<+\infty\}.$$

Let $k_1, k_2 \in \mathcal{K}_\varepsilon$ with $k_1 < k_2$. Since $f(x^{k_1}) > f^* + \varepsilon$ and $f^* \geq v^{k_1}$ there will be a new objective cut generated at x^{k_1}. It will be in the master from k_1 on, so it has to be satisfied at x^{k_2}:

$$f(x^{k_1}) + \langle \overline{g}^{k_1}, x^{k_2} - x^{k_1} \rangle \leq v^{k_2} \leq f^*.$$

On the other hand, $\varepsilon < f(x^{k_2}) - f^*$, which combined with the last inequality yields

$$\varepsilon < f(x^{k_2}) - f(x^{k_1}) - \langle \overline{g}^{k_1}, x^{k_2} - x^{k_1} \rangle.$$

The function $f(\cdot)$ is subdifferentiable in its domain and X is compact, so there is a constant C such that $f(x_1) - f(x_2) \leq C\|x_1 - x_2\|$, for all $x_1, x_2 \in \operatorname{dom} f \cap X$. By Assumption 6 we can choose C big enough so that $\|\overline{g}^k\| \leq C$ for all k. It follows that

$$\varepsilon < 2C\|x^{k_1} - x^{k_2}\| \quad \text{for all } k_1, k_2 \in \mathcal{K}_\varepsilon.$$

Since the set X is compact, the last inequality implies that the set $\mathcal{K}_\varepsilon$ is finite for each $\varepsilon > 0$. □

Similarly to the linear case, we can develop a multicut version of the method. It is virtually identical to the method discussed in the preceding section; the only difference is that we also need to construct and store cuts for the first stage objective. We leave to the reader the obvious technical details. Its theoretical convergence properties are exactly the same as those of the method with aggregate cuts discussed here.

The proof of convergence of the cutting plane method in the convex case indicates that the problem of deleting inactive cuts is even more acute than in the linear case. In fact, no reliable rule exists for the deletion of inactive cuts in the general convex case.

3 Regularized decomposition

3.1 The idea of regularization

The principal difficulty associated with cutting plane methods is the growth of the number of cuts that need to be stored in the master problem. Also, there is no easy way to make use of a good starting solution.

To mitigate these difficulties we add a quadratic regularizing term to the polyhedral model used in the master problem (2.25). We obtain the following *regularized master problem*:

$$\operatorname*{Min}_{x\in X}\left\{\frac{\rho}{2}\,\|x-w^k\|^2+c^Tx+\sum_{s=1}^{S}p_s\underline{Q}^{k,s}(x)\right\}, \tag{3.1}$$

where the models $\underline{Q}^{k,s}(\cdot)$ are defined in (2.24). In the proximal term $(\rho/2)\|x-w^k\|^2$, where $\rho>0$, the center w^k is updated depending on the relations between the value of $f(x^{k+1})$ at the master's solution, x^{k+1}, and its prediction provided by the current model:

$$f^k(x^{k+1}) := c^Tx^{k+1}+\sum_{s=1}^{S}p_s\underline{Q}^{k,s}(x^{k+1}). \tag{3.2}$$

Recall that $f(\cdot)$ is our true objective (see (1.3)). If these values are equal or close, we set $w^{k+1}=x^{k+1}$ (*serious step*); otherwise $w^{k+1}=w^k$ (*null step*). In any case, the collections of objective and feasibility cuts are updated, and the iteration continues.

We present the Regularized Decomposition Method in its most efficient multicut version. A method with averaged cuts can be developed and analyzed in an identical way.

The regularized master can be equivalently formulated as a quadratic programming problem

$$\operatorname{Min}\left\{c^Tx+\sum_{s=1}^{S}p_sv^s+\frac{\rho}{2}\,\|x-w^k\|^2\right\} \tag{3.3}$$

$$\alpha^{j,s}+(g^{j,s})^Tx\le v^s,\quad j\in J^k_{\text{obj}}(s),\quad s=1,\dots,S, \tag{3.4}$$

$$\beta^{j,s}+(r^{j,s})^Tx\le 0,\quad j\in J^k_{\text{fea}}(s),\quad s=1,\dots,S, \tag{3.5}$$

$$Ax=b,\ x\ge 0. \tag{3.6}$$

The detailed algorithm is stated in Fig. 4. As before, we assume that the initial point x^1 is such that $Q(x^1)<\infty$. Also, $\gamma\in(0,1)$ is a fixed constant used to compare the observed improvement in the objective value to the predicted improvement.

Step 0. Set $k=1$, $J^0_{\text{obj}}(s)=\emptyset$, $J^0_{\text{fea}}(s)=\emptyset$, $v^{1,s}=-\infty$, $s=1,\dots,S$.
Step 1. For $s=1,\dots,S$ solve subproblem (1.4) with $x=x^k$.

(a) If $Q^s(x^k)<\infty$ then set $J^k_{\text{fea}}(s)=J^{k-1}_{\text{fea}}(s)$. If $Q^s(x^k)>v^{k,s}$ then construct the objective cut (2.3) and set $J^k_{\text{obj}}(s)=J^{k-1}_{\text{obj}}(s)\cup\{k\}$; otherwise set $J^k_{\text{obj}}(s)=J^{k-1}_{\text{obj}}(s)$.
(b) If $Q^s(x^k)=\infty$ (i.e., problem (1.4) is infeasible), construct the feasibility cut (2.6) and set $J^k_{\text{obj}}(s)=J^{k-1}_{\text{obj}}(s)$, $J^k_{\text{fea}}(s)=J^{k-1}_{\text{fea}}(s)\cup\{k\}$.

Step 2. If $k=1$ or if

$$f(x^k)\le(1-\gamma)f(w^{k-1})+\gamma f^{k-1}(x^k),$$

then set $w^k=x^k$; otherwise set $w^k=w^{k-1}$.
Step 3. Solve the master problem (3.3)–(3.6). If it is infeasible, stop (the original problem has no feasible solutions). Otherwise, denote by (x^{k+1},v^{k+1}) its solution and set $f^k(x^{k+1})=c^Tx^{k+1}+\sum_{s=1}^S p_s v^{k+1,s}$.
Step 4. If $f^k(x^{k+1})=f(w^k)$ then stop (w^k is an optimal solution); otherwise continue.
Step 5. Remove from the sets of cuts $J^k_{\text{obj}}(s)$ and $J^k_{\text{fea}}(s)$, $s=1,\dots,S$, some (or all) cuts whose Lagrange multipliers at the solution of (3.3)–(3.6) were 0. Increase k by one, and go to Step 1.

Fig. 4. The regularized decomposition algorithm.

3.2 Relation to the proximal point method

To understand the mechanism of convergence of the Regularized Decomposition Method, let us suppose that the models $\underline{Q}^{k,s}(\cdot)$ in the master problem (3.1) are exact: $\underline{Q}^{k,s}(x)=Q^s(x)$ for all $x\in\mathbb{R}^{n_1}$ and all $s=1,\dots,S$. Consider the optimal value of such an ideal master problem

$$f_\rho(w):=\min_{x\in X}\left\{\frac{\rho}{2}\,\|x-w\|^2+c^Tx+\sum_{s=1}^S p_sQ^s(x)\right\}. \tag{3.7}$$

The function $f_\rho(\cdot)$ is called the Moreau–Yosida regularization of $f(\cdot)$. If $\operatorname{dom}f\cap X\neq\emptyset$, its regularization has many remarkable properties: convexity, continuous differentiability, Lipschitz continuity with constant 1, etc. For our purposes, its relations to f are of primary importance.

Lemma 8. *Suppose that there exists $\tilde{x}\in X$ such that $f(\tilde{x})<f(w)$. Then*

$$f_\rho(w)\le f(w)-\rho\|\tilde{x}-w\|^2\varphi\left(\frac{f(w)-f(\tilde{x})}{\rho\|\tilde{x}-w\|^2}\right),$$

where

$$\varphi(\tau) = \begin{cases} 0 & \text{if } \tau < 0, \\ \tau^2 & \text{if } 0 \le \tau \le 1, \\ -1 + 2\tau & \text{if } \tau > 1. \end{cases}$$

Proof. By convexity, the entire segment containing points $x = w + t(\tilde{x} - w)$ with $0 \le t \le 1$ is feasible for (3.7). Restricting the minimization to these x will provide an upper bound:

$$\begin{aligned} f_\rho(w) &\le \min_{0 \le t \le 1} \left[f((1-t)w + t\tilde{x}) + \frac{\rho t^2}{2} \|\tilde{x} - w\|^2 \right] \\ &\le f(w) + \min_{0 \le t \le 1} \left[t(f(\tilde{x}) - f(w)) + \frac{\rho t^2}{2} \|\tilde{x} - w\|^2 \right]. \end{aligned}$$

In the last estimate we also used the convexity of $f(\cdot)$. The value of t that minimizes the above expression is equal to

$$\hat{t} = \min\left(1, \frac{f(w) - f(\tilde{x})}{\rho \|\tilde{x} - w\|^2}\right).$$

Our assertion follows now from a straightforward calculation. □

At the solution $x(w)$ of problem (3.7) we shall have

$$f(x(w)) \le f_\rho(w) \le f(w) - \rho \|\tilde{x} - w\|^2 \varphi\left(\frac{f(w) - f(\tilde{x})}{\rho \|\tilde{x} - w\|^2}\right).$$

Therefore, if a better point exists, the ideal master (3.7) will find a better point. Consequently, $x = w$ is the minimizer in (3.7) if and only if w is a minimizer of f.

In fact, the *Proximal Point Method*,

$$w^{k+1} = x(w^k), \quad k = 1, 2, \ldots \tag{3.8}$$

must converge to an optimal solution, if an optimal solution exists.

Theorem 9. *Suppose that problem (1.3) has an optimal solution. Then the sequence $\{w^k\}$ generated by the Proximal Point Method is convergent to an optimal solution of (1.3).*

Proof. Let x^* be an optimal solution. We have the identity

$$\|w^{k+1}-x^*\|^2 = \|w^k - x^*\|^2 + 2\langle w^{k+1} - w^k, w^{k+1} - x^*\rangle - \|w^{k+1} - w^k\|^2. \tag{3.9}$$

The necessary condition of optimality for (3.7) at the solution $w^{k+1} = x(w^k)$ yields

$$0 \in \partial\Big[f(x) + \frac{\rho}{2}\ \|x - w^k\|^2\Big], \quad \text{at } x = w^{k+1}.$$

Thus,

$$-\rho(w^{k+1} - w^k) \in \partial f(w^{k+1}). \tag{3.10}$$

By the subgradient inequality,

$$f(x^*) \geq f(w^{k+1}) - \rho\langle w^{k+1} - w^k, x^* - w^{k+1}\rangle.$$

Using this inequality in (3.9) (and skipping the last term) we obtain

$$\|w^{k+1} - x^*\|^2 \leq \|w^k - x^*\|^2 - \frac{2}{\rho}(f(w^{k+1}) - f(x^*)). \tag{3.11}$$

Several conclusions follow from this inequality. First, the sequence $\{w^k\}$ is bounded, because the distance to x^* is nonincreasing. Secondly, summing up (3.11) from $k=1$ to ∞, we get

$$\sum_{k=2}^{\infty} (f(w^k) - f(x^*)) \leq \frac{\rho}{2}\ \|w^1 - x^*\|^2,$$

so $f(w^k) \to f(x^*)$. Consequently, for every accumulation point $\tilde{x}$ of $\{w^k\}$ we have $f(\tilde{x}) = f(x^*)$. We choose one such $\tilde{x}$, substitute it for x^* in (3.11) and conclude that the entire sequence $\{w^k\}$ is convergent to $\tilde{x}$. □

It is easy to see that we have not used in our analysis the fact that f is polyhedral; Theorem 9 remains true for any convex function f which has a minimum.[4] For polyhedral f the convergence is finite.

[4] A more general view on the Proximal Point Method is presented in Section 9.2.

Theorem 10. *Suppose that f is a convex polyhedral function and that a minimum of f exists. Then the Proximal Point Method stops after finitely many steps at a minimizer of f.*

Proof. Suppose that the method does not stop. Therefore, $0 \notin \partial f(w^{k+1})$, $k = 1, 2, \ldots$ and thus,

$$0 \notin \bigcup_{k=1}^{\infty} \partial f(w^{k+1}).$$

Since f is polyhedral, only finitely many different subdifferentials $\partial f(w^{k+1})$ exist. Each of them is a convex closed polyhedral set, so the right hand side of the last displayed relation is a union of finitely many closed sets. Thus, it is closed. Consequently, there exists $\varepsilon > 0$ such that the ball $B(0, \varepsilon)$ of radius ε centered at 0 has no common points with this union of subdifferentials. We get

$$B(0, \varepsilon) \cap \partial f(w^{k+1}) = \emptyset, \quad k = 1, 2, \ldots.$$

Since the sequence $\{w^k\}$ is convergent by Theorem 9, we have $w^{k+1} - w^k \to 0$. Therefore, $\rho(w^{k+1} - w^k) \in B(0, \varepsilon)$ for large k and we obtain a contradiction with (3.10). □

3.3 Convergence of the regularized decomposition method

The main difference between the Regularized Decomposition Method (and *bundle methods*, in general) and the Proximal Point Method (see Section 3.2) is that the master problem (3.1) uses a model $f^k(\cdot)$ instead of the true function $f(\cdot)$. Recall that

$$f^k(x) = \begin{cases} c^T x + \sum_{s=1}^{S} p_s \underline{Q}^{k,s}(x) & \text{if } x \in X, \\ +\infty & \text{otherwise.} \end{cases}$$

Its minimizer, x^{k+1}, is no longer guaranteed to be better than w^k. The role of null steps is to correct the model f^k, if x^{k+1} is not better than w^k. We shall see that such model improvements ensure that progress will be made whenever any progress is possible.

The analysis of convergence of the Regularized Decomposition Method requires an additional assumption.

Assumption 11. There exists a constant C such that $\|g^{k,s}\| \le C$ for all $k = 1, 2, \ldots$ and all $s = 1, \ldots, S$, whenever $Q^s(x^k) < \infty$.

It is clearly satisfied if basic objective cuts are employed. Also, if $\operatorname{dom} Q^s = \mathbb{R}^{n_1}$, by the polyhedrality of Q^s, Assumption 11 holds. The only situation, in which a potential for violating this assumption exists, is when x^k is a boundary point of $\operatorname{dom} Q^s$, in which case arbitrary large elements of the normal cone to $\operatorname{dom} Q^s$ may be added to the subgradient of $Q^s(\cdot)$. Technically, in such a situation the active second stage constraints are linearly dependent and arbitrary large multipliers $\pi^{k,s}$ may be used in formula (2.4). If our method for solving the second stage problem does not exhibit such properties, Assumption 11 will hold in this case, too.

By the construction of the method, the sequence $\{f(w^k)\}$ is nonincreasing. Our first result shows that the algorithm cannot get stuck at a nonoptimal point.

This property is obvious when basic cuts are employed, because each new cut added at Step 1 cuts the current master's solution off. As a result, due to the strict convexity of the master's objective, the optimal value of (3.1) increases after each null step. Since the number of possible combinations of basic cuts is finite, either a serious step will occur, or the stopping test of Step 4 will be satisfied.

For general subgradient cuts we need some analysis, but the conclusion remains the same.

Lemma 12. *Suppose that w^k is not an optimal solution of problem (1.3). Then there exist $m > k$ such that $f(w^m) < f(w^k)$.*

Proof. Suppose that null steps are made at iterations $j = k+1, k+2, \ldots$. Thus $w^j = w^k$ and

$$f(x^j) > f^{j-1}(x^j) + (1-\gamma)(f(w^{j-1}) - f^{j-1}(x^j)).$$

Denote $\delta_j = f(w^k) - f^{j-1}(x^j)$. We have

$$f(x^j) > f^{j-1}(x^j) + (1-\gamma)\delta_j. \tag{3.12}$$

We shall show that the optimal value of the master problem increases by a quantity related to δ_j, when new objective cuts are added.

First, the master's optimal value will not change, if we delete inactive cuts. We shall denote the model without inactive cuts by $f^{j-1}(x)$. Subdifferentiating (3.1) we see that $\rho(w^k - x^j)$ is a subgradient of $\underline{f}^{j-1}(\cdot)$ at x^j, because x^j is the master's solution. Thus,

$$f^j(x) \geq \underline{f}^{j-1}(x) \geq f^{j-1}(x^j) + \rho\langle w^k - x^j, x - x^j\rangle, \quad \text{for all } x. \tag{3.13}$$

Secondly, by (3.12), after adding new objective cuts at x^j we shall have

$$f^j(x^j) > f^{j-1}(x^j) + (1-\gamma)\delta_j.$$

Therefore, for all x,

$$f^j(x) > f^{j-1}(x^j) + (1-\gamma)\delta_j + \langle g, x - x^j\rangle,$$

where $g \in \partial f^j(x^j)$. Combining (3.13) with the last inequality we obtain

$$\begin{aligned} f^j(x) &\geq \max(f^{j-1}(x^j) + \rho\langle w^k - x^j, x - x^j\rangle, f^{j-1}(x^j) + (1-\gamma)\delta_j + \langle g, x - x^j\rangle) \\ &\geq f^{j-1}(x^j) + \rho\langle w^k - x^j, x - x^j\rangle \\ &\quad + \max(0, (1-\gamma)\delta_j + \langle g - \rho(w^k - x^j), x - x^j\rangle). \end{aligned}$$

Consequently, the next master's objective can be estimated from below as follows:

$$\begin{aligned} f^j(x) + \frac{\rho}{2}\|x - w^k\|^2 &\geq f^{j-1}(x^j) + \rho\langle w^k - x^j, x - x^j\rangle + \frac{\rho}{2}\|x - w^k\|^2 \\ &\quad + \max(0, (1-\gamma)\delta_j + \langle g - \rho(w^k - x^j), x - x^j\rangle) \\ &= f^{j-1}(x^j) + \frac{\rho}{2}\|x^j - w^k\|^2 + \frac{\rho}{2}\|x - x^j\|^2 \\ &\quad + \max(0, (1-\gamma)\delta_j + \langle g - \rho(w^k - x^j), x - x^j\rangle). \end{aligned}$$

It follows that the master's optimal value,

$$\theta_j = f^j(x^{j+1}) + \frac{\rho}{2}\|x^{j+1} - w^k\|^2,$$

satisfies the inequality

$$\begin{aligned} \theta_j - \theta_{j-1} &\geq \min_{x\in X}\left[\frac{\rho}{2}\|x - x^j\|^2 + \max(0, (1-\gamma)\delta_j + \langle g - \rho(w^k - x^j), x - x^j\rangle)\right] \\ &\geq \min_{h\in\mathbb{R}}\left[\frac{\rho h^2}{2} + \max(0, (1-\gamma)\delta_j - 2Ch)\right], \end{aligned}$$

where in the last relation we have used the estimates $\|g\| \leq C$ and $\|\rho(w^k - x^j)\| \leq C$.

The right hand side of the above expression can be estimated as follows. If $(1-\gamma)\delta_j \leq 4C^2/\rho$ we have

$$h = (1-\gamma)\delta_j/(2C), \quad \theta_j - \theta_{j-1} \geq \rho(1-\gamma)^2\delta_j^2/(8C^2),$$

otherwise

$$h = 2C/\rho, \quad \theta_j - \theta_{j-1} \geq -2C^2/\rho + (1-\gamma)\delta_j \geq (1-\gamma)\delta_j/2.$$

The sequence $\{\theta_j\}$ is increasing and bounded above by $f(w^k)$. If there are no serious steps after iteration k we conclude that $\delta_j \to 0$. Since $f(w^k) \geq \theta_j \geq f(w^k) - \delta_{j+1}$, we have $\theta_j \uparrow f(w^k)$.

On the other hand, the master's objective is bounded above by the Moreau–Yosida regularization (3.7)

$$\theta_j \leq f_\theta(w^j) = f_\theta(w^k).$$

If w^k is not optimal, Lemma 8 yields $f_\theta(w^k) < f(w^k)$ and we obtain a contradiction. □

We are now ready to prove the convergence of the Regularized Decomposition Method. Our analysis will have much in common with the analysis of the Proximal Point Method.

Theorem 13. *Suppose that problem (1.3) has an optimal solution. Then the Regularized Decomposition Method generates a sequence $\{w^k\}$ which is convergent to an optimal solution of (1.3).*

Proof. If w^k is optimal for some k, then $w^{j+1} = w^j$ for $j = k, k+1, \ldots$, and the theorem is true. If w^k is not optimal for any k, then, by Lemma 12, each series of null steps is finite and is followed by a serious step. Thus, the number of serious steps is infinite. Let us denote by $\mathcal{K}$ the set of iterations at which serious steps occur. If $w^{k+1} = x^{k+1}$ is the optimal solution of the master (3.1), we have the necessary condition of optimality

$$0 \in \partial\left[f^k(x) + \frac{\rho}{2}\,\|x - w^k\|^2\right], \quad \text{at } x = w^{k+1}.$$

Thus,

$$-\rho(w^{k+1} - w^k) \in \partial f^k(w^{k+1}).$$

Let x^* be an optimal solution of (1.3). By the subgradient inequality for f^k we get

$$f^k(x^*) \geq f^k(w^{k+1}) - \rho\langle w^{k+1} - w^k,\, x^* - w^{k+1}\rangle. \tag{3.14}$$

There is a serious step from w^k to $w^{k+1} = x^{k+1}$, so the test of Step 2 is satisfied (for $k+1$):

$$f(w^{k+1}) \leq (1-\gamma) f(w^k) + \gamma f^k(w^{k+1}).$$

After elementary manipulations we can rewrite it as

$$f^k(w^{k+1}) \geq f(w^{k+1}) - \frac{1-\gamma}{\gamma}[f(w^k) - f(w^{k+1})]. \tag{3.15}$$

Combining the last inequality with (3.14) and using the obvious relation $f(x^*) \geq f^k(x^*)$ we obtain

$$f(x^*) \geq f(w^{k+1}) + \frac{1-\gamma}{\gamma}[f(w^{k+1}) - f(w^k)] - \rho\langle w^{k+1} - w^k,\, x^* - w^{k+1}\rangle.$$

This can be substituted to the identity (3.9) which, after skipping the last term, yields

$$\begin{aligned}\|w^{k+1} - x^*\|^2 \leq \|w^k - x^*\|^2 &- \frac{2}{\rho}[f(w^{k+1}) - f(x^*)] \\ &+ \frac{2(1-\gamma)}{\gamma\rho}[f(w^k) - f(w^{k+1})] \quad \text{for all } k \in \mathcal{K}.\end{aligned} \tag{3.16}$$

It is very similar to inequality (3.11) in the proof of Theorem 9, and our analysis will follow the same line.

The series $\sum_{k=1}^{\infty}[f(w^k) - f(w^{k+1})]$ is convergent, because $\{f(w^k)\}$ is nonincreasing and bounded from below by $f(x^*)$. Therefore, we obtain from (3.16) that the distance $\|w^{k+1} - x^*\|$ is uniformly bounded, and $\{w^k\}$ must have accumulation points.

Summing up (3.16) for $k \in \mathcal{K}$ we get

$$\sum_{k\in\mathcal{K}} (f(w^{k+1}) - f(x^*)) \leq \frac{\rho}{2}\,\|w^1 - x^*\|^2 + \frac{1-\gamma}{\gamma}\left[f(w^1) - \lim_{k\to\infty} f(w^k)\right],$$

so $f(w^{k+1}) \to f(x^*)$, $k \in \mathcal{K}$. Consequently, at every accumulation point $\tilde{x}$ of $\{w^k\}$ one has $f(\tilde{x}) = f(x^*)$. Since $\tilde{x}$ is optimal, we can substitute it for x^* in (3.16). Skipping the negative term we get

$$\|w^{k+1} - \tilde{x}\|^2 \leq \|w^k - \tilde{x}\|^2 + \frac{2(1-\gamma)}{\gamma\rho}[f(w^k) - f(w^{k+1})].$$

It is true not only for $k \in \mathcal{K}$ but for all k, because at $k \notin \mathcal{K}$ we have a trivial equality here. Summing these inequalities from $k = j$ to $k = l > j$ we get

$$\|w^{l+1} - \tilde{x}\|^2 \leq \|w^j - \tilde{x}\|^2 + \frac{2(1-\gamma)}{\gamma\rho}[f(w^j) - f(w^{l+1})].$$

Since $\tilde{x}$ is an accumulation point, for any $\varepsilon > 0$ we can find j such that $\|w^j - \tilde{x}\| \leq \varepsilon$. Also, if j is large enough, $f(w^j) - f(w^{l+1}) \leq \varepsilon$ for all $l > j$, because $\{f(w^k)\}$ is convergent. Then $\|w^{l+1} - \tilde{x}\|^2 \leq \varepsilon^2 + 2\varepsilon(1-\gamma)/(\gamma\rho)$ for all $l > j$, so the entire sequence $\{w^k\}$ is convergent to $\tilde{x}$. □

Let us now prove a useful technical property of the Regularized Decomposition Method.

Remark 14. *Under the conditions of Theorem 13,*

$$\lim_{k\to\infty} \theta_k = f^*, \tag{3.17}$$

$$\lim_{k\to\infty} f^k(x^{k+1}) = f^*, \tag{3.18}$$

$$\lim_{k\to\infty} (x^{k+1} - w^k) = 0, \tag{3.19}$$

where f^* is the optimal value of our problem.

Proof. We shall prove at first that

$$f(w^k) - \gamma\theta_k \leq f(w^{k-1}) - \gamma\theta_{k-1}, \quad k = 1, 2, \ldots. \tag{3.20}$$

The inequality is true at all null steps, as shown in the proof of Lemma 12. If there is serious step at iteration k, we get from (3.13) that

$$\begin{aligned}\theta_k &\geq \min_x \left[f^k(x) + \frac{\rho}{2}\,\|x - w^k\|^2 \right] \\ &\geq \min_x \left[f^{k-1}(w^k) + \rho\langle w^{k-1} - w^k,\, x - w^k\rangle + \frac{\rho}{2}\,\|x - w^k\|^2 \right] \\ &= f^{k-1}(w^k) - \frac{\rho}{2}\,\|w^k - w^{k-1}\|^2 \\ &= \theta_{k-1} - \rho\|w^k - w^{k-1}\|^2. \end{aligned} \tag{3.21}$$

The test for a serious step is satisfied, so

$$\begin{aligned} f(w^{k-1}) - f(w^k) &\geq \gamma[f(w^{k-1}) - f^{k-1}(w^k)] \\ &= \gamma[f^{k-1}(w^{k-1}) - f^{k-1}(w^k)] \geq \gamma\rho\|w^k - w^{k-1}\|^2, \end{aligned}$$

where in the last transformation we have used (3.13) again. Combining the last relation with (3.21) we obtain (3.20), as required.

The optimal value of the master satisfies the inequality $\theta_k \leq f(w^k)$, so

$$f(w^k) - \gamma\theta_k \geq (1-\gamma) f(w^k) \geq (1-\gamma) f^*.$$

It follows from (3.20) that the sequence $\{f(w^k) - \gamma\theta_k\}$ is convergent, hence $\{\theta_k\}$ is convergent. If there is a serious step at iteration k, inequality (3.15) implies that

$$f(w^k) \geq \theta_k \geq f^k(w^{k+1}) \geq f(w^{k+1}) - \frac{1-\gamma}{\gamma}[f(w^k) - f(w^{k+1})].$$

Both sides converge to f^* as $k \to \infty$, $k \in \mathcal{K}$, so $\theta_k \to f^*$ at serious steps. But the entire sequence $\{\theta_k\}$ is convergent and (3.17) follows.

The objective of the master problem (3.1) is strictly convex. Therefore, its value at w^k can be estimated by using its minimum value, θ_k, and the distance to the minimum, $\|w^k - x^{k+1}\|$, as follows:

$$f(w^k) \geq \theta_k + \frac{\rho}{2}\,\|w^k - x^{k+1}\|^2.$$

Therefore,

$$0 \leq \frac{\rho}{2}\,\|w^k - x^{k+1}\|^2 \leq f(w^k) - \theta_k.$$

Since the right hand side converges to zero, (3.19) holds. Relation (3.18) follows directly from it. □

If problem (1.3) has feasible solutions but is unbounded, the Regularized Decomposition Method generates a sequence $\{w^k\}$ such that $f(w^k) \to -\infty$. To prove this suppose that $f(w^k)$ remains bounded from below. Then we can choose x^* such that $f(x^*) < f(w^k)$ for all k, and inequality (3.16) remains true. Thus, $f(w^k) \to f(x^*)$ as proved in the above theorem. But $f(x^*)$ can be made arbitrarily small by the choice of x^*, and we obtain absurd.

Our analysis, so far, did not rely on the fact that the function $f(\cdot)$ is polyhedral. Actually, the convexity and Assumption 11 are sufficient for the convergence of the Regularized Decomposition Method to a solution, if a solution exists.

In the polyhedral case we can prove finite convergence of the method, provided that basic cuts are employed and that the deletion rules are slightly refined.

Lemma 15. *Suppose that problem (1.3) has an optimal solution. Assume that the Regularized Decomposition Method uses basic objective and feasibility cuts. If the method does not stop then the number of serious steps is infinite and there exists k_0 such that for all $k > k_0$*

$$x^{k+1} = \arg\min f^k(\cdot), \tag{3.22}$$

and

$$f^k(x^{k+1}) = f^*, \tag{3.23}$$

where f^* denotes the optimal value of f.

Proof. By Theorem 13, the sequence $\{w^k\}$ is convergent to some optimal solution x^*. If the number of serious steps is finite, we must have $w^k = x^*$ for all sufficiently large k. We have already discussed such a situation before Lemma 12. Since each new cut added at Step 1 cuts the current master's solution off, the optimal value of (3.1) increases after each null step. The number of possible combinations of basic cuts is finite, so the stopping test of Step 4 must activate. Thus, if the method does not stop, the number of serious steps must be infinite.

Let us now look more closely at the master problem (3.1). The necessary condition of optimality for (3.1) implies

$$-\rho(x^{k+1} - w^k) \in \partial f^k(x^{k+1}). \tag{3.24}$$

There are only finitely many models $f^k(\cdot)$ possible and each of them, as a polyhedral function, has finitely many different subdifferentials. Therefore, the quantity $\operatorname{dist}(0, \partial f^k(x^{k+1}))$ may take only finitely many different values. Since the left hand side of (3.24) converges to zero by Remark 14, we must have

$$0 \in \partial f^k(x^{k+1})$$

for all sufficiently large k, so (3.22) is true.

Since only finitely many different minimum values of models $f^k(\cdot)$ may occur, Remark 14 implies (3.23) for all sufficiently large k. □

It follows that in the case of infinitely many steps, the serious steps of the Regularized Decomposition Method look (for all sufficiently large k) similarly to the steps of the Cutting Plane Method. The only role of the regularizing term at these late iterations is to select the solution of the linear master problem that is closest to the current center w^k. We also see that the minimum value of the linear master does not change and remains equal to the minimum value of the original problem. We need, therefore, to exclude the possibility of infinitely many such degenerate iterations. To achieve this, we need to slightly modify the algorithm.

The simplest modification is to forbid deletion of cuts at any iteration k at which the value of the linear part of the master's objective does not change, i.e., when

$$f^k(x^{k+1}) = f^{k-1}(x^k).$$

Indeed, by Lemma 15, after finitely many steps the Regularized Decomposition Method will enter the phase when (3.23) holds. From then on, no deletion will take place. By (3.22) the optimal solution of each master problem is the same as the optimal solution of the master problem (2.25) of the multicut method. By Theorem 5, the method will stop after finitely many steps.

The possibility to delete inactive cuts is one of the main computational advantages of the Regularized Decomposition Method. It becomes particularly important when the number of scenarios S is much larger than the dimension n_1 of the first stage vector x. The number of linearly independent active cuts in (3.1) cannot exceed $n_1 + S$. Since each of S models $\underline{Q}^{k,s}(\cdot)$ must have at least one active objective cut, it follows that at most n_1 of them will be represented by more than one active cut. We call them *critical scenarios*. All the other scenarios can be sufficiently well represented by linear models. Increasing S can only increase the number of noncritical scenarios, which do not contribute much to the complexity of the problem. Clearly, the set of critical scenarios depends on the current data: ρ, w^k, f^k. Our algorithm can be viewed as an iterative way of updating this set, by introducing new cuts

to some scenarios (possibly making them critical) and removing cuts from some other scenarios (possibly making them noncritical). The notion of critical scenarios can also be exploited in developing methods for the solution of the master problem (3.1).

The efficiency of the Regularized Decomposition Method can be improved by dynamically changing the proximal parameter, ρ. The general principle is clear: if the steps are too long, increase ρ, if they are too short, decrease ρ. A good way to decide whether steps are too long is to observe the difference

$$\Delta_k = f(x^k) - f(w^{k-1}).$$

We know that if it is positive (actually, not sufficiently negative) a null step will be made. If Δ_k is large, for example larger than $f(w^{k-1}) - f^{k-1}(x^k)$, it is advisable to increase ρ. On the other hand, when $f(x^k) = f^{k-1}(x^k)$, we may conclude that the step is too short, because we do not learn new cuts, so ρ has to be decreased. Detailed rules are discussed in the literature listed at the end of this chapter.

Another practical question associated with the Regularized Decomposition Method is the solution of the master problem (3.1). While for linear master problems, like (2.25), commercially available linear programming solvers may be employed, the regularized master requires a quadratic programming solver. Fortunately, the quadratic regularizing term is particularly simple, just the sum of squares, so the problem is very stable.

To allow infeasible starting points, we may add the constraints $v^s \geq -M$ to the master (3.3)–(3.6), where $-M$ is a lower bound for the optimal cost. If no feasible solution exists, the method will discover it in finitely many steps.

4 Trust region methods

One of the advantages of the Regularized Decomposition Method over the Cutting Plane Method is the ability to control the length of the steps made. It avoids making long shots towards minima of poor approximations and it makes good use of a reasonable initial point. Another way to prevent inefficient long steps is to explicitly limit the step size in the master problem (2.11)–(2.14) or in its multicut version (2.25). This is the idea of *Trust Region Methods*. The trust region master has the form

$$\operatorname*{Min}_{x \in X} \left\{ c^T x + \sum_{s=1}^{S} p_s \underline{Q}^{k,s}(x) \right\}, \tag{4.1}$$

$$\text{s.t.} \quad \|x - w^k\|_* \leq \Delta. \tag{4.2}$$

Here, similarly to the Regularized Decomposition Method, w^k is the "best" point found so far. The role of the constraint (4.2) is to keep the master's solution, x^{k+1}, in a neighborhood of w^k: a ball of radius $\Delta > 0$ centered at w^k.

From the theoretical point of view, the norm $\|\cdot\|_*$ may be any norm in $\mathbb{R}^{n_1}$. However, if we use the Euclidean norm, the master problem (4.1)–(4.2) becomes a quadratically constrained optimization problem. There is no advantage of using it instead of the regularized decomposition master (3.1). Indeed, if λ^k is the Lagrange multiplier associated with the constraint

$$\|x - w^k\|^2 \le \Delta^2,$$

which is equivalent to (4.2), then x^{k+1} is also a solution of the regularized master (3.1) with $\rho = 2\lambda^k$. For these reasons we shall discuss the Trust Region Method with

$$\|d\|_* = \max_{1 \le j \le n_1} |d_j|.$$

Then the constraint (4.2) can be represented as simple bounds

$$-\Delta \le x_j - w_j^k \le \Delta, \quad j = 1, \ldots, n_1,$$

and the master problem (4.1)–(4.2) becomes a linear programming problem.

Our presentation of the Trust Region Method will be very similar to the description of the Regularized Decomposition Method. We use $f^k(x)$ to denote the master's objective:

$$f^k(x) = \begin{cases} c^T x + \sum_{s=1}^S p_s \underline{Q}^{k,s}(x) & \text{if } x \in X, \\ +\infty & \text{otherwise.} \end{cases}$$

As before, we assume that the initial point x^1 is such that $Q(x^1) < \infty$. Also, $\gamma \in (0, 1)$ is a parameter of the method used to judge whether the master's solution, x^{k+1}, is significantly better than w^k. The detailed algorithm is presented in Fig. 5.

The analysis of convergence of the Trust Region Method is much easier than in the case of the Regularized Decomposition Method.

Theorem 16. *Suppose that problem (1.3) has an optimal solution. Then the sequence $\{w^k\}$ generated by the Trust Region Method has the property that*

$$\lim_{k \to \infty} f(w^k) = f^*,$$

where f^* is the optimal value of (1.3).

Step 0. Set $k=1$, $J^0_{\text{obj}}(s)=\emptyset$, $J^0_{\text{fea}}(s)=\emptyset$, $v^{1,s}=-\infty$, $s=1,\ldots,S$.
Step 1. For $s=1,\ldots,S$ solve subproblem (1.4) with $x=x^k$.

(a) If $Q^s(x^k)<\infty$ then set $J^k_{\text{fea}}(s)=J^{k-1}_{\text{fea}}(s)$. If $Q^s(x^k)>v^{k,s}$ then construct the objective cut (2.3) and set $J^k_{\text{obj}}(s)=J^{k-1}_{\text{obj}}(s)\cup\{k\}$; otherwise set $J^k_{\text{obj}}(s)=J^{k-1}_{\text{obj}}(s)$.
(b) If $Q^s(x^k)=\infty$ (i.e., problem (1.4) is infeasible), construct the feasibility cut (2.6) and set $J^k_{\text{obj}}(s)=J^{k-1}_{\text{obj}}(s)$, $J^k_{\text{fea}}(s)=J^{k-1}_{\text{fea}}(s)\cup\{k\}$.

Step 2. If $k=1$ or if

$$f(x^k)\le(1-\gamma)f(w^{k-1})+\gamma f^{k-1}(x^k),$$

then set $w^k=x^k$; otherwise set $w^k=w^{k-1}$.

Step 3. Solve the master problem (4.1)–(4.2). If it is infeasible, stop (the original problem has no feasible solutions). Otherwise, denote by (x^{k+1},v^{k+1}) its solution and set $f^k(x^{k+1})=c^Tx^{k+1}+\sum_{s=1}^S p_s v^{k+1,s}$.
Step 4. If $f^k(x^{k+1})=f(w^k)$ then stop (w^k is an optimal solution); otherwise increase k by one, and go to Step 1.

Fig. 5. The trust region algorithm.

Proof. Suppose that the number of serious steps is finite and let w denote the last point to which a serious step has been made. After the last serious step, the Trust Region Method becomes identical with the Cutting Plane Method (in its multicut version), for problem (1.3) with the additional constraint that $\|x-w\|_*\le\Delta$. By Theorem 5 we have that $\{f(x^k)\}$ is convergent to the minimum value of the problem having the additional constraint $\|x-w\|_*\le\Delta$. This minimum value must be equal to $f(w)$, if no serious steps are made after w. Thus $f(\mathrm{w})=f^*$, since otherwise a small step from w towards an optimal solution x^* would guarantee improvement.

Let us now consider the case of infinitely many serious steps. Let x^* be a solution of problem (1.3), and let

$$h_k=\|w^k-x^*\|_*.$$

Suppose that there is a serious step after iteration k, that is $w^{k+1}=x^{k+1}$. Then we have (by the rule of Step 2)

$$f(x^{k+1})-f^*\le(1-\gamma)(f(w^k)-f^*)+\gamma(f^k(x^{k+1})-f^*). \tag{4.3}$$

If $h_k\le\Delta$, then x^* is feasible for the master problem and

$$f^k(x^{k+1})\le f^k(x^*)\le f(x^*).$$

This combined with (4.3) yields

$$f(x^{k+1}) - f^* \le (1-\gamma)(f(w^k) - f^*).$$

Suppose now that $h_k > \Delta$. Consider the point

$$\tilde{x} = \frac{\Delta}{h_k}\, x^* + \left(1 - \frac{\Delta}{h_k}\right) w^k.$$

By construction, its distance to w^k is Δ. Since it is feasible for the master problem, we have

$$f^k(x^{k+1}) - f^* \le f^k(\tilde{x}) - f^* \le \left(1 - \frac{\Delta}{h_k}\right)(f(w^k) - f^*),$$

where we have also used the convexity of $f^k(\cdot)$. Combining this inequality with (4.3) we see that

$$f(x^{k+1}) - f^* \le \left(1 - \frac{\gamma\Delta}{h_k}\right)(f(w^k) - f^*).$$

In both cases, if there is a serious step after iteration k we have

$$f(w^{k+1}) - f^* \le \left(1 - \frac{\gamma\Delta}{\max(\Delta, h_k)}\right)(f(w^k) - f^*).$$

Let the index $l = 1, 2, \ldots$ number the serious steps only and let us write β_l for the value of $f(w^k) - f^*$ at the l-th new center w^k. The last inequality can be then rewritten as

$$\beta_{l+1} \le \left(1 - \frac{\gamma\Delta}{\max(\Delta, h_{k(l)})}\right)\beta_l, \quad l = 1, 2, \ldots,$$

where $k(l)$ is the iteration number at which the l-th serious step is made. By the triangle inequality for the norm $\|\cdot\|_*$ we have

$$h_{k(l)} \le h_1 + l\Delta.$$

Therefore,

$$\beta_{l+1} \leq \left(1 - \frac{\gamma\Delta}{h_1 + l\Delta}\right)\beta_l, \quad l = 1, 2, \ldots, \tag{4.4}$$

The sequence $\{\beta_l\}$ is decreasing and bounded from below by 0. Suppose that $\beta_l \geq \varepsilon > 0$ for all l. Then, summing (4.4) from $l = 1$ to m we obtain

$$0 \leq \beta_{m+1} \leq \beta_1 - \varepsilon\gamma\Delta \sum_{l=1}^{m} \frac{1}{h_1 + l\Delta},$$

which yields a contradiction as $m \to \infty$, because the series $\sum_{l=1}^{\infty} l^{-1}$ is divergent. Thus $\beta_l \to 0$. □

Let us observe that we have not used the polyhedrality of $f(\cdot)$ in our analysis. In fact, Theorem 16 remains true for the convex problem (2.37)–(2.38). The proof can be repeated *verbatim*, only at the beginning, in the case of finitely many serious steps, we have to use Theorem 7 instead of Theorem 5.

Also, similarly to the analysis of the Regularized Decomposition Method, we can prove that $f(w^k) \to \inf f$ even if the problem has no solution.[5] Indeed, suppose that there exists $f^* > \inf f$ such that $f(w^k) \geq f^*$ for all k. Then Theorem 16 implies that $f(w^k) \to f^*$. But f^* can be chosen arbitrarily close to $\inf f$, and the result follows.

In the linear case we can prove the finite convergence of the Trust Region Method.

Theorem 17. *The Trust Region Method finds an optimal solution of problem (1.3) after finitely many steps.*

Proof. If the number of serious steps is finite, the result follows from Theorem 5. Suppose that the number of serious steps is infinite. It follows from Theorem 16 that, for sufficiently large k, there exists an optimal solution in the Δ-neighborhood of w^k. Therefore,

$$f^k(x^{k+1}) \leq f^*,$$

for all sufficiently large k.

Proceeding as in the proof of Theorem 5 we conclude that the number of steps at which new cuts are added to the master problem must be finite, because there are finitely many cells of linearity of $f(\cdot)$. Thus,

$$f(x^{k+1}) = f^k(x^{k+1})$$

[5] $\inf f$ denotes the infimum of f over the feasible set of (1.3).

for all sufficiently large k. Combining the last two relations we see that we must have $f(x^{k+1}) = f^*$ for all sufficiently large k. Consequently, only one serious step can be made after that, a contradiction. □

As in the case of the Cutting Plane Method, deleting inactive cuts is not easy. If we use basic cuts alone, we may afford deleting inactive cuts whenever the optimal value of the master problem increases. For general subgradient cuts, no reliable rule can be found. Things are easier if we use the Euclidean norm for the trust region definition, because the arguments from the analysis of the Regularized Decomposition Method apply here. Using Euclidean norms, though, does not provide any significant benefits over the Regularized Decomposition Method.

The size of the trust region Δ, similarly to the parameter ρ of the Regularized Decomposition Method, can be adjusted in the course of computation. If $f(x^{k+1})$ is significantly larger than $\hat{f}^k(x^{k+1})$, we may decrease Δ to avoid too long steps. If no new cuts are generated, we may increase Δ to allow longer steps.

The Trust Region Method has been defined and analyzed under the assumption that a feasible starting point, x^1, is known. If such a point is not readily available, we may add the constraints $v^s \geq -M$ to (2.14), and start the method with some $\Delta > 0$. If inconsistency is detected (which must happen after finitely many steps by Theorem 5), we may increase Δ, and repeat the procedure. If (1.3) has feasible points, after finitely many such adjustments a feasible point will be found. After that, we may return to a smaller Δ, if we wish.

5 Nested cutting plane methods for multistage problems

5.1 Basic ideas

The idea of cutting plane methods can be extended to linear and polyhedral multistage stochastic programming problems with finitely many scenarios. We shall present it for the polyhedral model analyzed in detail in Chapter 2. The problem has the form:

$$\begin{array}{lllllll}
\text{Min} & \mathbb{E}\Big\{ f_1(x_1,\xi_1) + f_2(x_2,\xi_2) + f_3(x_3,\xi_3) + \cdots + f_T(x_T,\xi_T)\Big\} & & & & & \\
\text{s.t.} & A_{11}(\xi_1)x_1 & & & & & = b_1(\xi_1), \\
& A_{21}(\xi_2)x_1 & + & A_{22}(\xi_2)x_2 & & & = b_2(\xi_2), \\
& & & A_{32}(\xi_3)x_2 & + & A_{33}(\xi_3)x_3 & = b_3(\xi_3), \\
& \cdots\cdots & & \cdots\cdots & & \cdots\cdots & \cdots\cdots \\
& & & A_{T,T-1}(\xi_T)x_{T-1} & + & A_{TT}(\xi_T)x_T & = b_T(\xi_T),
\end{array} \tag{5.1}$$

where f_t, $t = 1,\ldots, T$, are random polyhedral functions, as defined in Chapter 2. In (5.1) each $x_t = x_t(\xi_{[1,t]})$ is a function of $\xi_{[1,t]}$, the random data of the problem. We assume that the process ξ has finitely many scenarios: $\xi^s = (\xi_1^s, \ldots, \xi_T^s)$, each of them with probability $p_s > 0$.

The probabilistic structure of the random data can be given in the form of a scenario tree. It has nodes arranged at levels which correspond to stages $1, 2,\ldots, T$. At level 1 we have only one root node, and we associate with it the value of ξ_1 (which is known at stage 1). At level 2 we have at least as many nodes as many different realizations of ξ_2 may occur. Each of them is connected with the root node by an arc. For each node i at level 2 (which corresponds to a particular realization $\xi_2^{(i)}$ of ξ_2) we create at least as many nodes at level 3 as different values of ξ_3 may follow $\xi_2^{(i)}$, and we connect them with the node i, etc. Generally, nodes at level t correspond to possible values of ξ_t that may occur. Each of them is connected to a unique node at level $t-1$, called the *ancestor* node, which corresponds to the identical first $t-1$ parts of the process $\xi_{[1,t]}$, and is also connected to nodes at level $t+1$, which correspond to possible continuations of $\xi_{[1,t]}$. The set of nodes is denoted $\mathcal{N}$. We refer the reader to Figure 2 in Chapter 1 for an example of a scenario tree.

For each node $i \in \mathcal{N}$ and its ancestor $a = a(i)$ in the scenario tree we denote by ρ_{ai} the probability of moving from node a to node i. Each probability ρ_{ai} can be viewed as the conditional probability of the process being in node i given its history up to the ancestor node $a = a(i)$. We can relate them to scenario probabilities p_s as follows. Every scenario s can be defined by its nodes $i_1, i_2,\ldots, i_T$, arranged in the chronological order, i.e., node i_2 (at level $t = 2$) is connected to the root $i_1 = 1$, node i_3 is connected to the node i_2, etc. The probability of that scenario is then given by the product $p_s = \rho_{i_1 i_2}\rho_{i_2 i_3}\cdots\rho_{i_{T-1} i_T}$.

It is also convenient to introduce node probabilities $p^{(i)}$. Denoting by $\mathcal{B}^{(i)}$ the set of scenarios passing through node i (at level t) of the scenario tree, we let $p^{(i)} := \mathbb{P}[\mathcal{B}^{(i)}]$. If $i_1, i_2,\ldots, i_t$, with $i_1 = 1$ and $i_t = i$, is the history of the process up to node i, then the probability $p^{(i)}$ is given by the product

$$p^{(i)} = \rho_{i_1 i_2}\rho_{i_2 i_3}\cdots\rho_{i_{t-1} i_t}$$

of the corresponding conditional probabilities. We also have the recursive relation: $p^{(i)} = \rho_{ai} p^{(a)}$, where $a = a(i)$ is the ancestor of the node i. This equation defines the conditional probability ρ_{ai} from the probabilities $p^{(i)}$ and $p^{(a)}$.

Proceeding exactly as in Section 3.2 of Chapter 1, we denote the value of x_t associated with node i at level t by $x^{(i)}$. Similarly, let $T^{(i)}$, $W^{(i)}$ and $h^{(i)}$ be the values $A_{t,t-1}$, A_{tt} and b_t in scenarios passing through node i. Finally, let $f^{(i)}(\cdot) = f_t(\cdot, \xi^s)$, where s is a scenario passing through node i (the value

of ξ_t is identical in all these scenarios). We can then rewrite the problem as follows

$$\begin{aligned} &\text{Min} \sum_{i\in\mathcal{N}} p^{(i)} f^{(i)}(x^{(i)}) \\ &\text{s.t.} \quad T^{(i)} x^{(a(i))} + W^{(i)} x^{(i)} = h^{(i)}, \quad i \in \mathcal{N}\backslash\{1\}, \\ &\qquad W^{(1)} x^{(1)} = h^{(1)}. \end{aligned} \tag{5.2}$$

The idea of nested decomposition is embedded in the tree formulation (5.2). For each node i of the scenario tree we define the subtree $\mathcal{T}(i)$ rooted at i and the associated *cost-to-go function*

$$Q^{(i)}(x^{(a(i))}) := \inf\left\{ \sum_{j\in\mathcal{T}(i)} \rho_{ij} f^{(j)}(x^{(j)}) \;\middle|\; T^{(j)} x^{a(j)} + W^{(j)} x^{(j)} = h^{(j)}, j \in \mathcal{T}(i) \right\} \tag{5.3}$$

with $\rho_{ij} = p^{(j)}/p^{(i)}$ denoting the conditional probability of reaching node $j \in \mathcal{T}(i)$ from i. It is the specialization to the tree model of the function defined in Chapter 2, formula (3.3). Our superscript (i) of Q represents in a sufficient way the entire history of ξ before node i has been reached.[6]

As outlined in Chapters 1 and 2, these functions are related through the dynamic programming equation

$$Q^{(i)}(x^{(a(i))}) = \inf_{x^{(i)}} \left\{ f^{(i)}(x^{(i)}) + \sum_{j\in\mathcal{S}(i)} \rho_{ij} Q^{(j)}(x^{(i)}) \;\middle|\; T^{(i)} x^{(a(i))} + W^{(i)} x^{(i)} = h^{(i)} \right\}, \tag{5.4}$$

where $\mathcal{S}(i)$ is the set of successors of node i: such j that $i = a(j)$. Clearly, when i is the leaf node (it corresponds to the last stage T) there are no successors, and the summation is over the empty set. If i is the root node corresponding to the initial stage, there is no ancestor $a(i)$, and Q has no arguments (it is just the optimal value of the entire problem). However, for the uniformity of notation, we shall work with the general form (5.4).

By Propositions 21 and 30 of Chapter 2, each cost-to-go function $Q^{(i)}(\cdot)$, if it is finite at at least one point, is a convex polyhedral function. Therefore, (5.4) is a two-stage problem with convex polyhedral 'second-stage' functions $Q^{(j)}(\cdot)$.

Equation (5.4) carries much information about our problem. In particular, it allows us to construct in a recursive way polyhedral approximations to the

[6] For $i = 1$ the cost-to-go $Q^{(1)}$ has no arguments and represents the optimal value of problem (5.1).

cost-to-go functions at every node, by employing objective and feasibility cuts in a manner similar to the two-stage case. Indeed, let $\underline{Q}^{(j)}(\cdot)$ be lower polyhedral approximations to $Q^{(j)}(\cdot)$, $j \in \mathcal{S}(i)$. Then

$$Q^{(i)}(x^{(a(i))}) \geq \inf_{x^{(i)}} \left\{ f^{(i)}(x^{(i)}) + \sum_{j \in \mathcal{S}(i)} \rho_{ij} \underline{Q}^{(j)}(x^{(i)}) \,\middle|\, T^{(i)} x^{(a(i))} + W^{(i)} x^{(i)} = h^{(i)} \right\}. \tag{5.5}$$

Let $x^{(a(i))}$ be fixed and let the problem at the right hand side of (5.5) have a solution with the optimal value $\underline{v}(x^{(a(i))})$. Denote by $\pi^{(i)}$ the Lagrange multipliers associated with the constraint $T^{(i)} x^{(a(i))} + W^{(i)} x^{(i)} = h^{(i)}$. Then, by Proposition 21 of Chapter 2, we can construct an objective cut, similarly to (2.3):

$$Q^{(i)}(x) \geq \underline{v}(x^{(a(i))}) - \langle [T^{(i)}]^T \pi^{(i)}, x - x^{(a(i))} \rangle \quad \text{for all } x.$$

This inequality defines an *objective cut*:

$$Q^{(i)}(x) \geq \alpha^{(i)} + (g^{(i)})^T x, \quad \text{for all } x, \tag{5.6}$$

where

$$g^{(i)} = -(T^{(i)})^T \pi^{(i)}, \tag{5.7}$$

$$\alpha^{(i)} = \underline{v}(x^{(a(i))}) + (\pi^{(i)})^T T^{(i)} x^{(a(i))}. \tag{5.8}$$

Thus, to obtain an objective cut for $Q^{(i)}(\cdot)$ at $x^{(a(i))}$ we need to solve the problem at the right hand side of (5.5), retrieve the Lagrange multipliers $\pi^{(i)}$ associated with its constraints, and apply formulas (5.6)–(5.8). In general, we have

$$\underline{v}(x^{(a(i))}) \leq Q^{(i)}(x^{(a(i))}),$$

and our objective cut does not have to support $Q^{(i)}(\cdot)$ at $x^{(a(i))}$. However, if the models $\underline{Q}^{(j)}(x^{(i)})$, $j \in \mathcal{S}(i)$, are exact, then our objective cut is exact at $x^{(a(i))}$.

If the problem in (5.5) is infeasible, we can construct a feasibility cut,

$$\beta^{(i)} + (r^{(i)})^T x \leq 0, \tag{5.9}$$

similarly to the way described in Section 2.1. If $f^{(i)}(\cdot)$ is a linear function, we proceed identically as in Section 2.1. If $f^{(i)}(\cdot)$ is a more complex convex

polyhedral function, we convert the problem in a usual way to a problem with a linear objective, by adding one new variable and new constraints describing the pieces of $f^{(i)}(\cdot)$.

It is important to remember that the right hand side of (5.5) is a lower bound for $Q^{(i)}(\cdot)$, so the feasibility cuts remain valid for the true cost-to-go function.

5.2 *The nested cutting plane method*

With each node of the scenario tree we can associate an approximate problem $P(i)$ of the form appearing at the right hand side of (5.5). Each of these problems maintains and updates the following data: its current solution $x^{(i)}$, convex polyhedral models of the cost-to-go functions $\underline{Q}^{(j)}(\cdot)$ of its successors $j \in \mathcal{S}(i)$ (if any), and the current approximation $v^{(i)}$ of the optimal value of its own cost-to-go function. The operation of each subproblem is formalized in Fig. 6.

It remains to describe the way in which these subproblems are initiated, activated in the course of the solution procedure, and terminated.

We assume that we know a sufficiently large number M such that each cost-to-go function can be bounded from below by $-M$. Our initial approximations of the successors' functions are just

$$\underline{Q}^{(j)}(\cdot) = -M.$$

Step 1. If i is not the root node, retrieve from the ancestor problem $P(a(i))$ its current approximate solution $x^{(a(i))}$.

Step 2. If i is not a leaf node, retrieve from each successor problem $P(j)$, $j \in \mathrm{S}(i)$, all new objective and feasibility cuts and update the approximations of their cost-to-go functions $\underline{Q}^{(j)}(\cdot)$.

Step 3. Solve the problem

$$\begin{aligned} \text{Min} \quad & \left\{ f^{(i)}(x^{(i)}) + \sum_{j \in \mathcal{S}(i)} \rho_{ij} \underline{Q}^{(j)}(x^{(i)}) \right\} \\ \text{s.t.} \quad & T^{(i)} x^{(a(i))} + W^{(i)} x^{(i)} = h^{(i)}. \end{aligned}$$

(a) If it is solvable, replace $x^{(i)}$ by the new solution and $v^{(i)}$ by the optimal value. If i is not the root node and $v^{(i)}$ increased, construct a new objective cut (5.6)–(5.8).

(b) If the problem is infeasible, and i is not the root node, construct a new feasibility cut (5.9). If i is the root node, then stop, because the entire problem is infeasible.

Step 4. Wait for the command to activate again, and then go to Step 1.

Fig. 6. Subproblem $P(i)$ of the nested decomposition method.

At the beginning, no ancestor solutions are available, but we can initiate each subproblem with some arbitrary point $x^{(i)} \in \operatorname{dom} f^{(i)}$.

There is much freedom in determining the order in which the subproblems are solved. Three rules have to be observed.

I. There is no sense to activate a subproblem $P(i)$ whose ancestor's solution $x^{(a(i))}$ did not change, and whose successors $P(j)$, $j \in \mathcal{S}(i)$, did not generate any new cuts since this problem was activated last.

II. If a subproblem $P(i)$ has a new solution $x^{(i)}$, each of its successors $P(j)$, $j \in \mathcal{S}(i)$ has to be activated some time after this solution has been obtained.

III. If a subproblem $P(i)$ generates a new cut, i.e., if it is infeasible or has a new optimal value $v^{(i)}$, its ancestor $P(a(i))$ has to be activated some time after this cut has been generated.

We shall terminate the method if Rule I applies to all subproblems, in which case we claim that the current solutions $x^{(i)}$ constitute the optimal solution of the entire problem. The other stopping test is the infeasibility test at Step 3(a) for the root node. It is obvious, because we operate with relaxations here, and if a relaxation is infeasible, so is the true problem.

Theorem 18. *The Nested Cutting Plane Method after finitely many subproblem solutions either discovers the infeasibility of problem* (5.1) *or stops at its optimal solution.*

Proof. Arguing exactly as in the proofs of Theorems 4 and 5 we can prove that each leaf subproblem can generate only finitely many different objective and feasibility cuts. Thus, each of its predecessors can have only finitely many different polyhedral models of the leaves' optimal value functions. Consequently, it can also generate only finitely many different objective and feasibility cuts. Proceeding in this way from the leaves to the root we conclude that the root subproblem can be activated only finitely many times, because it may receive only finitely many different cuts. This, however, implies that its successors can be activated only finitely many times, by new root's solutions or by new cuts coming from their successors. Continuing this argument from the root to the leaves we deduce that each subproblem can be activated only finitely many times. The method must stop.

Suppose that the method stops as a result of Rule I. Consider the immediate predecessors of the leaf nodes. Arguing as in the proofs of Theorems 4 and 5 we have

$$x^{(i)} = \arg\min \left\{ f^{(i)}(x) + \sum_{j \in \mathcal{S}(i)} \rho_{ij} Q^{(j)}(x) \mid T^{(i)} x^{(a(i))} + W^{(i)} x = h^{(i)} \right\}. \tag{5.10}$$

Moreover, we know that the cuts generated at these solutions support $Q^{(i)}$ at $x^{(a(i))}$. Thus, $Q^{(i)}(x^{(a(i))}) = \underline{Q}^{(i)}(x^{(a(i))})$. Consequently, relations (5.10) must hold at the ancestor $a(i)$. By induction, these relations are true at all nodes. The optimality of the current solution follows then from Theorem 34 of Chapter 2. □

5.3 Modifications and extensions

There are many ways in which the general idea of nested decomposition can be modified and adapted to a particular problem.

Regularization and trust regions

First, we may use the concepts of regularization or trust regions to stabilize the iterates and to facilitate the convergence. The best place to introduce these modifications is the root problem, associated with the first stage. The operation of the method is then almost identical with the two-stage case: the root node is the regularized (trust region) master, and the subtrees rooted at its successors are giant second stage problems. The protocol for processing the subproblems cannot be so flexible as in the purely linear method: we need an *exact* solution to the 'second stage problem' in order to decide whether a serious step can be made. In other words, we solve the 'second stage problems' by a linear nested decomposition method with a flexible protocol, and only after they are solved to optimality, we make adjustments at the root node.

Regularization or trust regions at lower level nodes introduce more complication. In general, optimal values of these modified subproblems are no longer lower bounds for the true cost-to-go functions. Therefore, these subproblems cannot generate objective cuts for their predecessors so easily as linear subproblems could. Only when the entire subtree rooted at such a subproblem is solved to optimality, a valid cut can be generated. It follows that using regularization or trust regions at lower levels of the tree restricts the processing order of the subproblems to *depth-first* protocols, in which a subproblem is processed only if all its successors are solved to optimality.

Cut sharing

The number of cuts that are generated and stored by the nested decomposition method may easily become very large. In some cases we may reduce this number by using the similarity between the cost-to-go functions $Q^{(i)}$ corresponding to different nodes. The most dramatic simplification occurs when the parts of the data vector ξ_t, corresponding to different time stages $t = 1,\ldots,T$ are statistically independent, in which case the distribution of $\xi_{[t+1,T]}$ does not depend on $\xi_{[1,t]}$. In simple words, the data subtrees rooted at nodes i at level t may differ only by the data at level t and are

identical otherwise. If the data at two nodes i and j at level t are the same, their cost-to-go functions $Q^{(i)}(\cdot)$ and $Q^{(j)}(\cdot)$, defined by (5.3), are *identical*. It follows that every cut generated for a particular function $Q^{(i)}(\cdot)$ is valid for all other functions at this level which happen to have identical data $f^{(i)}(\cdot)$, $T^{(i)}$, $W^{(i)}$ and $h^{(i)}$.

Estimating cuts

Similarly to the two-stage case discussed in Section 2.4, we may work with estimated expected value cuts rather than with exact averages. However, in the nested method the estimation errors propagate quickly in the subproblem tree. For more information see the chapter on Monte Carlo methods.

6 Introduction to dual methods

Dual methods for stochastic programming problems are based on duality relations associated with the nonanticipativity constraints. Since the methods are essentially the same for two-stage and multistage models, we shall present them here only for the general multistage case.

Let us consider the polyhedral multistage problem:

$$
\begin{aligned}
&\text{Min } \mathbb{E}\{f_1(x_1,\xi_1)+f_2(x_2,\xi_2)+f_3(x_3,\xi_3)+\cdots+f_T(x_T,\xi_T)\}\\
&\text{s.t.}\quad A_{11}(\xi_1)x_1 \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad = b_1(\xi_1),\\
&\qquad A_{21}(\xi_2)x_1 + A_{22}(\xi_2)x_2 \qquad\qquad\qquad\qquad\qquad\qquad = b_2(\xi_2),\\
&\qquad\qquad\qquad A_{32}(\xi_3)x_2 + A_{33}(\xi_3)x_3 \qquad\qquad\qquad = b_3(\xi_3),\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \cdots\cdots\\
&\qquad\qquad\qquad A_{T,T-1}(\xi_T)x_{T-1} + A_{TT}(\xi_T)x_T \qquad = b_T(\xi_T),
\end{aligned}
\tag{6.1}
$$

with each $x_t = x_t(\xi_{[1,t]})$ being a function of $\xi_{[1,t]}$. We assume throughout this chapter that the process ξ_t, $t=1,\ldots,T$, has finitely many realizations, ξ^s, $s=1,\ldots,S$, attained with probabilities $p_1,\ldots,p_s$. The objective parts associated with the successive stages, $f_t(x_t,\xi_t)$, $t=1,\ldots,T$, are *random polyhedral* functions.[7]

In the dual approach we assume that each decision x_t may depend on *all* random data, ξ. Since ξ has finitely many realizations we may model our assumption by assigning a decision sequence,

$$x^s = (x_1^s,\ldots,x_T^s),$$

[7] See Chapter 2, Section 2.3.

to the sth realization of ξ. The problem takes on the form

$$\begin{aligned}
&\text{Min} \sum_{s=1}^{S} p_s[f_1(x_1^s, \xi_1^s) + f_2(x_2^s, \xi_2^s) + f_3(x_3^s, \xi_3^s) + \cdots + f_T(x_T^s, \xi_T^s)] \\
&\text{s.t.} \quad A_{11}(\xi_1^s)x_1^s \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad = b_1(\xi_1^s), \\
&\qquad A_{21}(\xi_2^s)x_1^s \quad + \quad A_{22}(\xi_2^s)x_2^s \qquad\qquad\qquad\qquad = b_2(\xi_2^s), \\
&\qquad\qquad\qquad A_{32}(\xi_3^s)x_2^s \qquad\quad + \quad A_{33}(\xi_3^s)x_3^s \qquad = b_3(\xi_3^s), \\
&\qquad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
&\qquad\qquad\qquad A_{T,T-1}(\xi_T^s)x_{T-1}^s \quad + \quad A_{TT}(\xi_T^s)x_T^s \quad = b_T(\xi_T^s), \\
&\qquad\qquad\qquad s = 1, \ldots, S.
\end{aligned} \tag{6.2}$$

As discussed extensively in Chapters 1 and 2, this formulation is not equivalent to the original problem (6.1), unless we introduce additional constraints that limit the dependence of x_t on ξ to the information that is available up to time t. These conditions take the form of *nonanticipativity constraints*,

$$x_t^s = x_t^\sigma \quad \text{for all } s, \sigma \text{ for which } \xi_{[1,t]}^s = \xi_{[1,t]}^\sigma, \quad t = 1, \ldots, T. \tag{6.3}$$

Abstractly, they define a subspace $\mathcal{W}$ of *implementable* policies. For numerical methods it is convenient to describe this subspace by a set of algebraic equations. Clearly, (6.3) is such a set, but there is much redundancy in it. Let us describe one way in which the nonanticipativity constraints can be written explicitly.

Let I_t be the set of nodes at level t. For a node $i \in I_t$ we denote by $\mathcal{B}(i)$ the set of scenarios that pass through node i and are, therefore, indistinguishable on the basis of the information available up to time t. The sets $\mathcal{B}(i)$ for all $i \in I_t$ are the atoms of the sigma-subalgebra $\mathcal{F}_t$ associated with the time stage t. We denote them by $\mathcal{B}_t^1, \ldots, \mathcal{B}_t^{\gamma_t}$.

Let us assume that all scenarios are ordered in such a way that each set $\mathcal{B}_t^\nu$ is a set of consecutive numbers l_t^ν, $l_t^\nu + 1, \ldots, r_t^\nu$. Then nonanticipativity can be expressed by the system of equations

$$x_t^s - x_t^{s+1} = 0, \quad s = l_t^\nu, \ldots, r_t^\nu - 1, \; t = 1, \ldots, T-1, \; \nu = 1, \ldots, \gamma_t. \tag{6.4}$$

In other words, each decision is related to its neighbors from the left and from the right, if they correspond to the same node of the scenario tree. The coefficients of constraints (6.4) define a giant matrix

$$G = [G^1 \; \cdots \; G^S]$$

whose rows have two nonzeros each: 1 and −1. Thus, we obtain an algebraic description of the nonanticipativity constraints:

$$Gx = 0.$$

Example 19. Consider the scenario tree depicted in Fig. 7.

Let us assume that the scenarios are numbered from the left to the right. Our nonanticipativity constraints take on the form:

$$x_1^1 - x_1^2 = 0,\ x_1^2 - x_1^3 = 0,\ \ldots,\ x_1^7 - x_1^8 = 0,$$
$$x_2^1 - x_2^2 = 0,\ x_2^2 - x_2^3 = 0,\ x_2^3 - x_2^4 = 0,$$
$$x_2^5 - x_2^6 = 0,\ x_2^6 - x_2^7 = 0,\ x_2^7 - x_2^8 = 0,$$
$$x_3^2 - x_3^3 = 0,$$
$$x_3^5 - x_3^6 = 0,\ x_3^6 - x_3^7 = 0.$$

Using I to denote an identity matrix of an appropriate dimension, we may write the constraint matrix G as shown in Fig. 8. G is always a very sparse matrix, because each of its rows has only two nonzeros.

Let us define the objective functions associated with the scenarios:

$$f^s(x^s) := \begin{cases} \sum_{t=1}^{T} f_t(x_t^s, \xi_t^s) & \text{if the constraints of (6.2) are satisfied} \\ & \text{for scenario } s, \\ +\infty & \text{otherwise.} \end{cases}$$

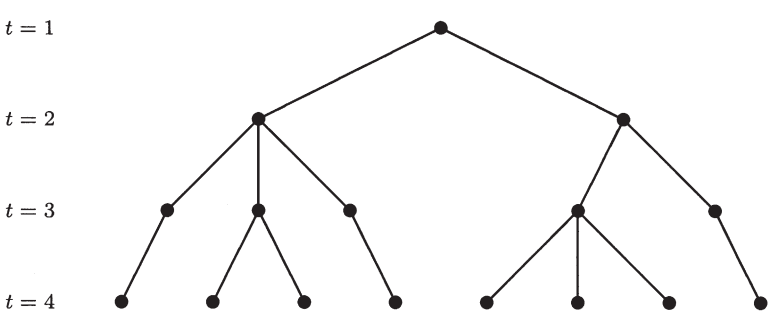

Fig. 7. Example of a scenario tree.

$$
G = \begin{bmatrix}
I & -I & & & & & & & & \\
 & I & -I & & & & & & & \\
 & & I & -I & & & & & & \\
 & & & I & -I & & & & & \\
 & & & & I & -I & & & & \\
 & & & & & I & -I & & & \\
 & & & & & & I & -I & & \\
 & & & & & & & I & -I & \\
 & & & & & & & & I & -I \\
I & -I & & & & & & & & \\
 & I & -I & & & & & & & \\
 & & I & -I & & & & & & \\
 & & & I & -I & & & & & \\
 & & & & I & & & & & \\
 & & & & & I & -I & & & \\
 & & & & & & I & -I & & \\
 & & & & & & & I & -I & \\
 & & & & & & & & I & \\
 & & I & -I & & & & & & \\
 & & & & & I & -I & & & \\
 & & & & & & I & & & \\
 & & & & & & & I & -I & \\
 & & & & & & & & I &
\end{bmatrix}
$$

Fig. 8. The constraint matrix corresponding to the scenario tree from Fig. 7. The subdivision corresponds to scenario subvectors $x^1, \ldots, x^8$.

Problem (6.2)–(6.3) can be now written compactly as

$$
\begin{aligned}
&\text{Min} \left\{ f(x) := \sum_{s=1}^{S} p_s f^s(x^s) \right\} \\
&\text{s.t.} \quad Gx = 0.
\end{aligned} \tag{6.5}
$$

The optimality conditions and the duality theory for problem (6.5) have been studied in Chapter 2. Let us recall briefly the main concepts and results.

Consider the Lagrangian of (6.5):

$$
\begin{aligned}
L(x, \lambda) &:= f(x) + \langle \lambda, Gx \rangle \\
&= \sum_{s=1}^{S} p_s f^s(x^s) + \sum_{s=1}^{S} \langle \lambda, G^s x^s \rangle.
\end{aligned}
$$

The associated dual function has the form

$$
D(\lambda) := \inf_x L(x, \lambda) = \sum_{s=1}^{S} \inf_{x^s} [\, p_s f^s(x^s) + \langle \lambda, G^s x^s \rangle]. \tag{6.6}
$$

Thus,

$$D(\lambda) = \sum_{s=1}^{S} D^s(\lambda)$$

with each $D^s(\cdot)$ defined by a *scenario subproblem*:

$$D^s(\lambda) = \inf_{x^s}[\, p_s f^s(x^s) + \langle \lambda, G^s x^s \rangle], \quad s = 1, \ldots, S. \tag{6.7}$$

We see that

$$D(\lambda) = -\sup_{x}[\langle -G^T\lambda, x\rangle - f(x)] = -f^*(-G^T\lambda),$$

where $f^*(\cdot)$ is the Fenchel conjugate of $f(\cdot)$. Thus, $D(\cdot)$ is concave and polyhedral. By the Fenchel–Moreau Theorem, at any point λ at which $D(\lambda) > -\infty$ we have

$$\partial D(\lambda) = \{Gx \colon x \in \hat{X}(\lambda)\}, \tag{6.8}$$

where $\hat{X}(\lambda)$ is the set of minimizers in (6.6).[8] Clearly,

$$\hat{X}(\lambda) = \hat{X}^1(\lambda) \times \cdots \times \hat{X}^S(\lambda)$$

with $\hat{X}^s(\lambda)$ denoting the solution set of the sth subproblem in (6.7).

We shall restrict our presentation to the case when the sets $\operatorname{dom} f^s$ are bounded for each s. In this case the dual functional $D(\cdot)$ is finite everywhere and each set $\hat{X}(\lambda)$ is a bounded convex polyhedron.

It follows that in order to calculate the value and a subgradient of $D(\lambda)$ we need to solve scenario subproblems (6.7) and aggregate their values to $D(\lambda)$ and their solutions to a subgradient of D at λ.

Also, we see that

$$D^s(\lambda) = -\sup_{x^s}[\langle -(G^s)^T\lambda, x^s\rangle - p_s f^s(x^s)] = -p_s(f^s)^*(-G^T\lambda/p_s),$$

Thus,

$$\partial D^s(\lambda) = \{G^s x^s \colon x^s \in \hat{X}^s(\lambda)\}. \tag{6.9}$$

[8] Since $D(\cdot)$ is concave, $\partial D(\lambda)$ is the set of vectors g such that $D(\lambda') \le D(\lambda) + \langle g, \lambda' - \lambda\rangle$ for all λ'.

The dual problem has the form

$$\text{Max } D(\lambda). \tag{6.10}$$

By the duality theory (cf. Theorems 46 and 47 of Chapter 2) we know that if the primal problem (6.5) has an optimal solution then the dual problem (6.10) has an optimal solution, optimal values of both problems coincide, and for any optimal solution $\hat{\lambda}$ of the dual problem and any optimal solution $\hat{x}$ of the primal problem we have

$$\hat{x} \in \hat{X}(\hat{\lambda}),$$

i.e., the primal solution solves the scenario subproblems (6.7).

This is the fundament of the dual methods. Their main structure is the same: employ an iterative method for solving (6.10) and recover the primal solution from the scenario subproblems (6.7). In the solution of the dual problem we shall take advantage of the decomposable structure of the dual function.

7 The dual cutting plane method

The structure of the dual problem (6.10) is very similar to the form of the primal problem (1.3) discussed in connection with the two-stage model: we have to maximize a sum of many concave polyhedral functions. Therefore, we may adapt to our case the Cutting Plane Method presented in Section 2.

At any point λ^k generated by the method we can solve each scenario subproblem (6.7) and obtain a solution $x^{k,s}$. They define objective cuts:

$$D^s(\lambda) \le D^s(\lambda^k) + \langle G^s x^{k,s}, \lambda - \lambda^k \rangle, \quad s = 1, \ldots, S,$$

for each scenario subproblem, and the objective cut

$$D(\lambda) \le D(\lambda^k) + \langle G x^k, \lambda - \lambda^k \rangle \tag{7.1}$$

for the entire dual function.

Since the feasible sets X^s are assumed to be bounded, for any λ the scenario subproblems have optimal solutions and the dual function is bounded. Therefore, we do not need to operate with feasibility cuts here.

We shall assume that we know a polyhedron Λ which contains the optimal solution of the dual problem. It may be a sufficiently large box. Let J^k be a subset of $\{1,\ldots,k\}$. The cuts collected at iterations $j \in J^k$ are included into the *master problem*:

$$\text{Max } v \tag{7.2}$$

$$\text{s.t.} \quad D(\lambda^j) + \langle Gx^j, \lambda - \lambda^j \rangle \geq v, \quad j \in J^k, \tag{7.3}$$

$$\lambda \in \Lambda, \tag{7.4}$$

whose decision variables are λ and $v \in \mathbb{R}$. To illustrate its meaning, let us fix λ and carry out the maximization in v. We see that the optimal value of v is equal to

$$\overline{D}^k(\lambda) = \min_{j \in J^k} [D(\lambda^j) + \langle Gx^j, \lambda - \lambda^j \rangle].$$

By the subgradient inequality (7.1), $\overline{D}^k(\lambda) \geq D(\lambda)$ for all λ. Thus, the master problem (7.2)–(7.4) is the problem to maximize a certain upper estimate of the dual functional. After this problem is solved, the true value of the dual functional is calculated, the approximation $D^k(\cdot)$ improved, and the iteration continues. The algorithm in its simplest form is presented in Fig. 9.

The convergence properties of the method have been analyzed in Section 2. Theorem 4 can be rephrased here as follows.

Theorem 20. *If the dual problem has an optimal solution in the set* Λ, *then the Dual Cutting Plane Algorithm finds an optimal solution of the dual problem in finitely many iterations*.

In the presentation above we have assumed that all cuts collected at earlier iterations are retained in the master problem, that is, $J^k = \{1, \ldots, k\}$. We would like to decrease their number by deleting some or all inactive cuts. As discussed in Section 2 it is safe in the version of the method which uses so-called *basic cuts*. In our case, basic cuts correspond to optimal solutions of the scenario subproblems which are drawn from a finite set. Since the feasible sets of the scenario subproblems are bounded convex polyhedra, such finite sets exist: they are basic feasible solutions of the scenario subproblems. So, if our method generates only basic solutions to scenario subproblems, we may drop inactive cuts, whenever the optimal value of the master problem *decreases*.

Step 0. Set $k = 1$, $J^0 = \emptyset$, $v^1 = \infty$.
Step 1. For $s = 1, \ldots, S$ solve subproblem (6.7) with $\lambda = \lambda^k$, construct the objective cut (7.1) and set $J^k = J^{k-1} \cup \{k\}$.
Step 2. If $D(\lambda^k) = v^k$ then stop (optimal dual solution has been found); otherwise continue.
Step 3. Solve the master problem (7.2)–(7.4), denote by (λ^{k+1}, v^{k+1}) its solution, increase k by one, and go to Step 1.

Fig. 9. The dual cutting plane algorithm.

On the other hand, after the dual problem has been solved, we need to recover the optimal solution of the primal problem. In general, it will not be composed of basic solutions of the scenario subproblems. Combinations of basic solutions are needed.

To illustrate how such a combination can be constructed, let us assume that the optimal solution $\hat{\lambda}$ of the dual problem has been found by the Dual Cutting Plane Method at iteration k, and that it is an interior point of the set Λ. It follows that

$$0 \in \partial \overline{D}^k(\hat{\lambda})$$

(because $\hat{\lambda}$ is the maximizer of $\overline{D}^k(\cdot)$). Therefore, there exists a subset $\hat{J}$ of cardinality at most $m+1$ among the constraints (7.3), which has the following properties. First, these cuts are active at $\hat{\lambda}$, i.e.,

$$D(\hat{\lambda}) = D(\lambda^j) + \langle Gx^j, \hat{\lambda} - \lambda^j \rangle, \quad j \in \hat{J}. \tag{7.5}$$

Secondly, the convex hull of their gradients contains 0, that is, there exist nonnegative multipliers μ_j, $j \in \hat{J}$, such that

$$\sum_{j \in \hat{J}} \mu_j G x^j = 0, \tag{7.6}$$

$$\sum_{j \in \hat{J}} \mu_j = 1. \tag{7.7}$$

The multipliers μ_j are Lagrange multipliers associated with the cuts indexed by $j \in \hat{J}$ at the optimal solution to the master problem. Consider the point

$$\hat{x} = \sum_{j \in \hat{J}} \mu_j x^j.$$

By (7.6) it satisfies the nonanticipativity condition $Gx = 0$.

Let us denote by $\hat{f}$ the common optimal value of the primal and the dual problem. Adding equations (7.5) multiplied by μ_j and using (7.6)–(7.7) we get

$$\hat{f} = \sum_{j \in \hat{J}} \mu_j [D(\lambda^j) - \langle Gx^j, \lambda^j \rangle] = \sum_{j \in \hat{J}} \mu_j f(x^j) \geq f(\hat{x}),$$

where in the last inequality we have used the convexity of f. Since $\hat{x}$ satisfies the nonanticipativity constraint (see (7.6)), it is optimal.

The dual problem, however, is rather difficult to solve, because of the large dimension of the dual vector λ. It might be interesting to compare (6.10) with the problem (1.3) arising in the primal approach to the two-stage problem. Both involve sums of many polyhedral functions, but in the two-stage case the dimension of the decision vector x does not grow with the scenario number S, which allows for efficient solution techniques exploiting the notion of critical scenarios. On the other hand, in (6.10) the dimension of λ grows with the number of scenarios.

We can use more sophisticated algorithms to solve the dual problem (6.10): regularized (bundle) methods or trust region methods. Their analysis is the same as in Sections 3 and 4.

8 The augmented Lagrangian method

8.1 The basic method

The high dimension of the dual vector λ in (6.10) suggests another approach to (6.2)–(6.3): the augmented Lagrangian method. Let us consider the compact formulation (6.5) and define the augmented Lagrangian as follows:

$$L_\rho(x, \lambda) := f(x) + \langle \lambda, Gx \rangle + \frac{\rho}{2} \|G(x)\|^2, \tag{8.1}$$

where $\rho > 0$ is a penalty coefficient. The *Multiplier Method* applied to (6.5) carries out for $k = 1, 2, \ldots$ the following iteration:

(a) given λ^k, find $x^k = \arg\min L_\rho(x, \lambda^k)$;
(b) set $\lambda^{k+1} = \lambda^k + \rho G x^k$.

The convergence of this algorithm is analyzed in the following theorem.

Theorem 21. *Assume that the dual problem (6.10) has an optimal solution. Then the sequence $\{\lambda^k\}$ generated by the Multiplier Method after finitely many steps finds an optimal solution of (6.10).*

Proof. We shall show that the Multiplier Method is equivalent to the Proximal Point Method[9] for solving the dual problem:

$$\lambda^{k+1} = \arg\max \left\{ D(\lambda) - \frac{1}{2\rho} \|\lambda - \lambda^k\|^2 \right\}, \quad k = 1, 2, \ldots. \tag{8.2}$$

[9] The Proximal Point Method is discussed in Sections 3.2 and 9.2.

The problem above can be written as a max-min problem:

$$\operatorname*{Max}_{\lambda} \operatorname*{Min}_{x\in X}\left\{f(x) + \langle \lambda, Gx\rangle - \frac{1}{2\rho}\,\|\lambda - \lambda^k\|^2\right\}.$$

Let us interchange the 'Max' and 'Min' operators (we shall soon show that it is legitimate):

$$\operatorname*{Min}_{x\in X} \operatorname*{Max}_{\lambda}\left\{f(x) + \langle \lambda, Gx\rangle - \frac{1}{2\rho}\,\|\lambda - \lambda^k\|^2\right\}.$$

Now we can calculate the maximum with respect to λ in a closed form:

$$\lambda(x) = \lambda^k + \rho Gx. \tag{8.3}$$

After substituting the optimal λ, the min-max problem becomes equivalent to

$$\operatorname*{Min}_{x\in X} L_\rho(x, \lambda^k).$$

Its solution is the point x^k calculated in Step (a) of the Multiplier Method. The corresponding multiplier value follows from (8.3) and is equal to λ^{k+1}, as defined in Step (b) of the Multiplier Method. Now we can provide a proof for the validity of the interchange of the 'Min' and 'Max' operations, by verifying that the pair (x^k, λ^{k+1}) is a saddle point of the function

$$\Phi^k(x, \lambda) = f(x) + \langle \lambda, Gx\rangle - \frac{1}{2\rho}\,\|\lambda - \lambda^k\|^2.$$

Indeed, $\Phi^k(x^k, \lambda^{k+1}) = L_\rho(x^k, \lambda^k)$ and x^k minimizes $L_\rho(x, \lambda^k)$ in X, while λ^{k+1} maximizes $\langle \lambda, Gx\rangle - (1/2\rho)\|\lambda - \lambda^k\|^2$, as shown in (8.3). Thus the saddle point (x^k, λ^{k+1}) is a solution of both 'min-max' and 'max-min' problems.

The convergence of the Multiplier Method follows now from Theorems 9 and 10. □

In fact, the same result is true for the convex case (but without finite convergence, in general).

The utmost simplicity of the multiplier update (b) is the main advantage of this approach. On the other hand, the minimization step (a) cannot be easily decomposed into scenario subproblems, as it could be done in (6.6) for the ordinary Lagrangian, because of the quadratic term $\|Gx\|^2$. We shall address this issue in the next section.

8.2 The separable approximation

One possibility to overcome the nonseparability of the augmented Lagrangian is to apply an iterative nonlinear Jacobi method to the minimization of (8.1). This method uses, at iteration j, a certain approximation $\tilde{x}^{k,j}$ of the minimizer x^k in (a), and solves for each scenario s simplified problems:

$$x^{k,j+1,s} = \arg\min_{x^s \in X^s} \left\{ p_s f^s(x^s) + \langle \lambda^k, G^s x^s \rangle + \frac{\rho}{2} \left\| G^s x^s + \sum_{\sigma \neq s} G^\sigma \tilde{x}^{k,j,\sigma} \right\|^2 \right\}. \tag{8.4}$$

In other words, the augmented Lagrangian (8.1) is minimized with respect to the decisions associated with scenario s, while keeping the decisions associated with other scenarios $\sigma \neq s$ fixed at $\tilde{x}^{k,j,\sigma}$. This is done for all scenarios s. Then, with some stepsize $\tau \in (0, 1)$, the reference point is updated

$$\tilde{x}^{k,j+1} = (1 - \tau)\tilde{x}^{k,j} + \tau x^{k,j},$$

and the iteration continues. This general scheme converges for sufficiently small stepsizes τ, but the convergence may be slow. What makes it particularly useful in our case is its application together with the construction of the constraint matrix G in (6.4).

The quadratic term of the augmented Lagrangian has then the form

$$\|Gx\|^2 = \sum_{t=1}^{T-1} \sum_{v=1}^{\gamma_t} \sum_{s=l^v}^{r^v-1} \|x_t^s - x_t^{s+1}\|^2.$$

The minimization in (8.4) involves at most two simple quadratic terms for each subvector x_t^s, $t = 1, \ldots, T-1$, relating it to the reference values at its neighbors: $\tilde{x}_t^{s-1}$ and $\tilde{x}_t^{s+1}$. Not only makes it the subproblems easier to manipulate and solve, but it has a positive impact on the speed of convergence.

Denote the functions minimized at (8.4) by

$$\tilde{L}^s(x^s, \tilde{x}, \lambda) = p_s f^s(x^s) + \langle \lambda, G^s x^s \rangle + \frac{\rho}{2} \left\| G^s x^s + \sum_{\sigma \neq s} G^\sigma \tilde{x}^\sigma \right\|^2,$$

and let

$$\tilde{L}(x, \tilde{x}, \lambda) = \sum_{s=1}^{S} \tilde{L}^s(x^s, \tilde{x}, \lambda) - \frac{\rho}{2}(S-1) \left\| \sum_{\sigma=1}^{S} G^\sigma \tilde{x}^\sigma \right\|^2.$$

Clearly, $\tilde{L}(\cdot, \tilde{x}, \lambda)$ is the function minimized in parallel in (8.4). The error of the approximation to L_ρ can be estimated as follows.

Lemma 22. *For all x, $\tilde{x}$ and λ the following inequality is true:*

$$|L_\rho(x, \lambda) - \tilde{L}(x, \tilde{x}, \lambda)| \le \frac{\rho}{2} \sum_{s=1}^{S} \|G^s(x^s - \tilde{x}^s)\|^2.$$

Proof. By direct calculation we obtain

$$L_\rho(x, \lambda) - \tilde{L}(x, \tilde{x}, \lambda) = \frac{\rho}{2} \sum_{s=1}^{S} \sum_{\sigma \neq s} \langle G^s(x^s - \tilde{x}^s), G^\sigma(x^\sigma - \tilde{x}^\sigma)\rangle. \tag{8.5}$$

The scalar products above can be nonzero only for $\sigma = s-1$ and $\sigma = s+1$ (otherwise $(G^s)^T G^\sigma = 0$). For such (s, σ) pairs we have

$$\langle G^s(x^s - \tilde{x}^s), G^\sigma(x^\sigma - \tilde{x}^\sigma)\rangle = \sum_{i=1}^{n} \sum_{l=1}^{n} \langle G_i^s(x_i^s - \tilde{x}_i^s), G_l^\sigma(x_l^\sigma - \tilde{x}_l^\sigma)\rangle. \tag{8.6}$$

Let us denote by $J(s, \sigma)$ the set of variables linked by nonanticipativity constraints in scenarios s and σ. We have

$$|\langle G_i^s, G_l^\sigma\rangle| = \begin{cases} 1 & \text{if } l = i \text{ and } i \in J(s, \sigma), \\ 0 & \text{otherwise,} \end{cases}$$

and

$$\begin{aligned} |\langle G^s(x^s - \tilde{x}^s), G^\sigma(x^\sigma - \tilde{x}^\sigma)\rangle| &\le \sum_{i \in J(s,\sigma)} (x_i^s - \tilde{x}_i^s)(x_i^\sigma - \tilde{x}_i^\sigma) \\ &\le \frac{1}{2} \sum_{i \in J(s,\sigma)} [(x_i^s - \tilde{x}_i^s)^2 + (x_i^\sigma - \tilde{x}_i^\sigma)^2]. \end{aligned}$$

Using this in (8.5) we obtain the inequality

$$|L_\rho(x, \lambda) - \tilde{L}(x, \tilde{x}, \lambda)| \le \frac{\rho}{4} \sum_{s=1}^{S} \sum_{\sigma \neq s} \sum_{i \in J(s,\sigma)} [(x_i^s - \tilde{x}_i^s)^2 + (x_i^\sigma - \tilde{x}_i^\sigma)^2].$$

Each term $(x_i^s - \tilde{x}_i^s)^2$ appears in this sum at most $2\|G_i^s\|^2$ times and we can continue our estimate as follows:

$$|L_\rho(x, \lambda) - \tilde{L}(x, \tilde{x}, \lambda)| \le \frac{\rho}{2} \sum_{s=1}^{S} \sum_{i \in J(s,\sigma)} \|G_i^s\|^2 (x_i^s - \tilde{x}_i^s)^2.$$

Noting that the columns G_i^s, $i = 1, \ldots, n$, are orthogonal, we obtain the required result. □

Our next result estimates the progress that is made within each subproblem.

Lemma 23. *Suppose that x^s minimizes in X^s the function $\tilde{L}^s(\cdot, \tilde{x}, \lambda)$. Then*

$$\tilde{L}^s(x^s, \tilde{x}, \lambda) - \tilde{L}^s(\tilde{x}^s, \tilde{x}, \lambda) \le -\frac{\rho}{2} \|G^s(x^s - \tilde{x}^s)\|^2.$$

Proof. For every $g^s \in \partial \tilde{L}^s(x^s, \tilde{x}, \lambda)$ we have

$$\tilde{L}^s(\tilde{x}^s, \tilde{x}, \lambda) - \tilde{L}^s(x^s, \tilde{x}, \lambda) \ge \langle g^s, \tilde{x}^s - x^s \rangle + \frac{\rho}{2} \|G^s(x^s - \tilde{x}^s)\|^2.$$

Since x^s is a minimizer, there exists a subgradient g^s of $\tilde{L}^s(\cdot, \tilde{x}, \lambda)$ at x^s such that

$$\langle g^s, \tilde{x}^s - x^s \rangle \ge 0,$$

and the result follows. □

We are now ready to prove the convergence of the Jacobi method.

Theorem 24. *Assume that the sets X^s, $s = 1, \ldots, S$ are bounded. Then*

(a) *For all $s = 1, \ldots, S$ we have $\lim_{j \to \infty} G^s(x^{s,j} - \tilde{x}^{s,j}) = 0$;*
(b) *Every accumulation point of the sequence $\tilde{x}^{k,j}$, $j = 1, 2, \ldots$, is a minimizer of $L_\rho(x, \lambda^k)$ over X.*

Proof. Let us estimate the progress made for the true augmented Lagrangian function. By Lemma 22 we have

$$L_\rho(\tilde{x}^{k,j} + \tau(x^{k,j} - \tilde{x}^{k,j}), \lambda^k) - \tilde{L}(\tilde{x}^{k,j} + \tau(x^{k,j} - \tilde{x}^{k,j}), \lambda^k)$$
$$\le \frac{1}{2} \rho\tau^2 \sum_{s=1}^{S} \|G^s(x^{k,j,s} - \tilde{x}^{k,j,s})\|^2.$$

By Lemma 23 and by the convexity of $\tilde{L}$,

$$\tilde{L}(\tilde{x}^{k,j}+\tau(x^{k,j}-\tilde{x}^{k,j}),\lambda^k)-\tilde{L}(\tilde{x}^{k,j},\lambda^k)\leq-\frac{1}{2}\,\rho\tau\sum_{s=1}^{S}\|G^s(x^{k,j,s}-\tilde{x}^{k,j,s})\|^2.$$

Combining the last two inequalities we obtain

$$L_\rho(\tilde{x}^{k,j+1},\lambda^k)-L_\rho(\tilde{x}^{k,j},\lambda^k)\leq\frac{1}{2}\,\rho\tau(1-\tau)\sum_{s=1}^{S}\|G^s(x^s-\tilde{x}^s)\|^2. \tag{8.7}$$

This proves (a). At any accumulation point x^* we then must have that $x^{*,s}$ is a minimizer of $\tilde{L}^s(x^s,x^*,\lambda^k)$ over X^s. Thus, x^* minimizes $L_\rho(x,\lambda^k)$, as required. □

It follows from our analysis that the sparsity of the nonanticipativity constraints allows for using relatively large stepsizes in the Jacobi method. The best estimate of the speed of convergence is obtained in (8.7) for $\tau=1/2$.

Another advantage is the minimal amount of communication between scenario subproblems within the Jacobi method. In fact, each subproblem s needs to communicate with at most two subproblems: $s-1$ and $s+1$, which is very important for parallel and distributed computation.

Also, both iterative processes: the outer iterations for the multipliers and the inner iterations of the Jacobi method, are very simple and they can be carried out for a very large number of scenarios, and for large scenario subproblems.

9 Progressive hedging

9.1 The method

Let us write the multistage stochastic programming problem (6.5) in an abstract form

$$\operatorname*{Min}_{x\in\mathcal{W}}\sum_{s=1}^{S}p_s f^s(x^s), \tag{9.1}$$

where $\mathcal{W}$ denotes the linear subspace of nonanticipative policies, that is, policies satisfying the nonanticipativity constraints $Gx=0$. As before, x^s denotes the sequence of decisions associated with scenario s. Using

$$f(x)=\sum_{s=1}^{S}p_s f^s(x^s)$$

and the indicator function of the subspace of nonanticipative policies,

$$\delta_{\mathcal{W}}(x) = \begin{cases} 0 & \text{if } x \in \mathcal{W}, \\ +\infty & \text{otherwise,} \end{cases}$$

we can rewrite (9.1) as follows:

$$\operatorname{Min}\{f(x) + \delta_{\mathcal{W}}(x)\}. \tag{9.2}$$

Let us note that each of the components individually is easy to deal with: f by decomposition into scenarios, and $\mathcal{W}$ by linear algebra. Problems of this form are well understood and a number of *operator splitting methods* have been suggested to exploit the properties of the two components in solving the whole. One of the most general is the *Douglas–Rachford method.*

The method generates a sequence (x^k, u^k) such that

$$u^k \in \partial\delta_{\mathcal{W}}(x^k),$$

which means two things: x^k is an element of $\mathcal{W}$, and $u^k \in \mathcal{W}^{\perp}$. This is accomplished in the following way ($\mu > 0$ is a fixed parameter):

(a) find y^k and $g^k \in \partial f(y^k)$ such that

$$y^k + \mu g^k = x^k - \mu u^k; \tag{9.3}$$

(b) find $x^{k+1} \in \mathcal{W}$ and $u^{k+1} \in \mathcal{W}^{\perp}$ such that

$$x^{k+1} + \mu u^{k+1} = y^k + \mu u^k. \tag{9.4}$$

Step (a) amounts to finding

$$y^k = \arg\min_y \left\{ f(y) + \langle u^k, y\rangle + \frac{1}{2\mu}\,\|y - x^k\|^2 \right\}, \tag{9.5}$$

and g^k is the subgradient of f that appears in the necessary conditions of optimality for this problem. Problem (9.5) decomposes into individual scenarios

$$\operatorname*{Min}_{y^s} \left\{ p_s f^s(y^s) + \langle u^{k,s}, y^s\rangle + \frac{1}{2\mu}\,\|y^s - x^{k,s}\|^2 \right\}.$$

Their solutions $y^{k,s}$ form a policy y^k which is feasible with respect to scenario constraints, but not necessarily nonanticipative. The nonanticipativity is restored in step (b).

Since both u^k and u^{k+1} are in $\mathcal{W}^\perp$, step (b) implies that

$$x^{k+1} - y^k \perp \mathcal{W}.$$

Therefore,

$$x^{k+1} = \Pi_{\mathcal{W}}(y^k),$$

where $\Pi_{\mathcal{W}}$ is the orthogonal projection on the subspace of nonanticipative policies. To perform this operation, for each time stage t we consider bundles $\mathcal{B}_t^1, \ldots, \mathcal{B}_t^{\gamma_t}$ of scenarios which cannot be distinguished up to time t. They correspond to nodes of the scenario tree located at level t. For each of these bundles, say $\mathcal{B}_t^j$, we replace all y_t^s, $s \in \mathcal{B}_t^j$, by their average on $\mathcal{B}_t^j$:

$$x_t^{k+1,s} = \frac{1}{|\mathcal{B}_t^j|} \sum_{\sigma \in \mathcal{B}_t^j} y_t^{k,\sigma}. \tag{9.6}$$

Finally, directly from (b),

$$u^{k+1} = u^k + \mu^{-1}(y^k - x^{k+1}) = u^k + \mu^{-1}\Pi_{\mathcal{W}^\perp}(y^k).$$

We can specialize and simplify this general technique by taking into account the specific nature of our objective function f and of the subspace $\mathcal{W}$. We change the scalar product in the entire decision space to

$$\langle x, u\rangle_{\mathbb{P}} = \sum_{s=1}^{S} p_s \langle x^s, u^s\rangle. \tag{9.7}$$

This changes in a corresponding manner $\mathcal{W}^\perp$ and the orthogonal projection operation. Then problem (9.5) (with the new scalar product and the corresponding norm) decomposes into scenario subproblems

$$\operatorname*{Min}_{y^s} \left\{ f^s(y^s) + \langle u^{k,s}, y^s\rangle + \frac{1}{2\mu} \|y^s - x^{k,s}\|^2 \right\}. \tag{9.8}$$

Its solution, as before, will be denoted $y^{k,s}$. The orthogonal projection on $\mathcal{W}$ (in the new geometry) can be calculated in way similar to (9.6), but with the conditional expectation on $\mathcal{B}_t^j$, instead of the plain average:

$$x_t^{k+1,s} = \frac{1}{\mathbb{P}(\mathcal{B}_t^j)} \sum_{\sigma \in \mathcal{B}_t^j} p_\sigma y_t^{k,\sigma}. \tag{9.9}$$

The formula for the multiplier,

$$u^{k+1} = u^k + \mu^{-1}(y^k - x^{k+1}), \tag{9.10}$$

remains the same. Note that u^k is always orthogonal to x^k in the new geometry: $\langle u^k, x^k \rangle_{\mathbb{P}} = 0$. The Progressive Hedging Method is presented in Fig. 10.

The main difficulty associated with the application of the Progressive Hedging Method is the absence of a merit function whose improvements can be monitored from iteration to iteration. This makes adjustments in the penalty parameter μ harder than in the case of other decomposition methods.

9.2 Convergence

To prove the convergence of the Progressive Hedging Method, we need to put it into a more abstract framework of the theory of maximal monotone operators. For a multivalued operator $M: \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ we define its *domain* as the set of x for which $M(x) \neq \emptyset$. Its graph is the set

$$\mathcal{G}(M) := \{(x, y) : y \in M(x)\}.$$

An operator M is called *monotone*, if

$$\langle x' - x, y' - y \rangle \geq 0, \quad \text{for all } x, x' \in \operatorname{dom} M,\ y \in M(x),\ y' \in M(x').$$

Step 0. Set $x^0 = \mathrm{W}$ and $u^0 = \mathrm{W}^\perp$. Set $k = 0$.
Step 1. For $s = 1, \ldots, S$ solve the scenario subproblems (9.8).
Step 2. For each stage $t = 1, \ldots, T$ and for each bundle $\mathcal{B}_t^j$ of scenarios which cannot be distinguished up to stage t, calculate the corresponding components of $x_t^{k+1,s}$ for $s \in \mathcal{B}_t^j$ by (9.9).
Step 3. Calculate new multipliers by (9.10).
Step 4. Increase k by one, and go to Step 1.

Fig. 10. The progressive hedging algorithm.

It is *maximal monotone* if its graph is not contained in the graph of another monotone operator. An example of a maximal monotone operator is the subdifferential of a proper convex function.

Together with a maximal monotone operator M we shall consider its *resolvent*:

$$J_{\mu M} := (I + \mu M)^{-1},$$

where $\mu > 0$. It is well defined on the whole space $\mathbb{R}^n$, that is, for every $z \in \mathbb{R}^n$ there exists a unique $x \in \operatorname{dom} M$ and a unique $y \in M(x)$ such that $z = x + \mu y$.

Consider the inclusion

$$0 \in M(x).$$

If x is its solution, then

$$x = J_{\mu M}(x),$$

i.e., x is a fixed point of the resolvent. The opposite is also true: every fixed point of the resolvent is a zero of M.

Let us introduce yet another operator derived from M:

$$O_{\mu M} := 2J_{\mu M} - I.$$

By the maximal monotonicity of M, every $z \in \mathbb{R}^n$ can be uniquely represented as $z = x + \mu y$ with $y \in M(x)$. Then $J_{\mu M}(z) = x$ and we get

$$O_{\mu M}(x + \mu y) = x - \mu y. \tag{9.11}$$

This identity can be used as follows. Let $z = x + \mu y$ with $y \in M(x)$, $z' = x' + \mu y'$ with $y' \in M(x')$. We have

$$\begin{aligned} \|O_{\mu M}(z) - O_{\mu M}(z')\|^2 &= \|(x - \mu y) - (x' - \mu y')\|^2 \\ &= \|(x + \mu y) - (x' + \mu y')\|^2 - 4\mu \langle x - x', y - y' \rangle. \end{aligned}$$

By the monotonicity of M, we have $\langle x - x', y - y' \rangle \geq 0$, and the last displayed inequality yields

$$\|O_{\mu M}(z) - O_{\mu M}(z')\| \leq \|z - z'\| \quad \text{for all } z, z' \in \mathbb{R}^n.$$

This means that the operator $O_{\mu M}$ is *nonexpansive*.

For any nonexpansive $D: \mathbb{R}^n \to \mathbb{R}^n$ and every $\lambda \in (0, 1)$ the classical method

$$x^{k+1} = (1 - \lambda)x^k + \lambda D(x^k), \quad k = 1, 2, \ldots, \tag{9.12}$$

is convergent to a fixed point of D, if a fixed point of D exists. By setting $D = O_{\mu M}$ and $\lambda = 1/2$, we get

$$z^{k+1} = (1 - \lambda)z^k + \lambda O_{\mu M}(z^k) = J_{\mu M}(z^k). \tag{9.13}$$

The above iteration is called the Proximal Point Method for maximal monotone operators. We have just proved that it is convergent to a fixed point of $O_{\mu M}$, which, by (9.11), must be a zero of M. In the special case when M is the subdifferential of a convex function, algorithm (9.13) becomes identical with the method analyzed in Section 3.2, with $\rho = 1/\mu$.

We can use all these observations in the analysis of the Douglas–Rachford method. Consider

$$M = A + B,$$

where A and B are maximal monotone. We define the operator D as the functional composition of $O_{\mu A}$ and $O_{\mu B}$:

$$D = O_{\mu A} \circ O_{\mu B}.$$

It is nonexpansive, because both $O_{\mu A}$ and $O_{\mu B}$ are nonexpansive. The method (9.12) for $\lambda = 1/2$ takes on the form:

$$z^{k+1} = \frac{1}{2}\, O_{\mu A}(O_{\mu B}(z^k)) + \frac{1}{2}\, z^k. \tag{9.14}$$

It is called the Douglas–Rachford method for finding a zero of $A + B$.

Iterations (9.14) can be carried out as follows. We find the unique representation $z^k = x^k + \mu u^k$ with $u^k \in B(x^k)$. Then, by (9.11),

$$O_{\mu B}(z^k) = x^k - \mu u^k.$$

Next, we find the unique y^k and $g^k \in A(y^k)$ for which

$$y^k + \mu g^k = x^k - \mu u^k. \tag{9.15}$$

This gives us

$$O_{\mu A}(O_{\mu B}(z^k)) = y^k - \mu g^k.$$

The next point $z^{k+1} = x^{k+1} + \mu u^{k+1}$ is defined as

$$z^{k+1} = \frac{1}{2}(y^k - \mu g^k) + \frac{1}{2}(x^k + \mu u^k) = y^k + \mu u^k. \tag{9.16}$$

Theorem 25. *Assume that A and B are maximal monotone and a zero of $A+B$ exists. Then the sequence $\{z^k\}$ generated by the Douglas–Rachford method (9.14) is convergent to a point $\hat{z} = \hat{x} + \mu\hat{u}$, $\hat{u} \in B(\hat{x})$, such that $\hat{x}$ is a zero of $A+B$.*

Proof. Let us assume that a fixed point of D exists. Then, by the convergence theory of method (9.12) we know that the sequence $\{z^k\}$ is convergent to a fixed point $\hat{z} = \hat{x} + \mu\hat{u}$ of D, $\hat{u} \in B(\hat{x})$.

It remains to characterize fixed points of D. If $z^1 = \hat{z}$, operations (9.15)–(9.16) generate for all k points $x^k = \hat{x}$, $u^k = \hat{u}$, $y^k = \hat{y}$ and $g^k = \hat{g}$. Since $z^{k+1} = z^k$ we have $\hat{x} + \mu\hat{u} = \hat{y} + \mu\hat{u}$, so $\hat{x} = \hat{y}$. From (9.15) we see that $\hat{u} = -\hat{g}$. In other words,

$$0 = \hat{u} + \hat{g},$$

where $\hat{u} \in A(\hat{x})$ and $\hat{g} \in B(\hat{x})$. Thus $\hat{x}$ is a zero of $A+B$. The converse is also true. If x is a zero of $A+B$, then there exists $u \in A(x) \cap [-B(x)]$ and $x + \mu u$ is a fixed point of D. Therefore, our assumption that a zero of $A+B$ exists implies the existence of fixed point of D. Consequently, the sequence $\{x^k\}$ converges to a zero $A+B$. □

In fact, the operator $\frac{1}{2}(D+I)$ is a resolvent of another maximal monotone operator N and the entire Douglas–Rachford method can be interpreted as a proximal point method for N.

Now it is easy to rephrase the Progressive Hedging Method in terms of the abstract Douglas–Rachford method for maximal monotone operators. By setting

$$A(x) = \partial f(x), \quad B(x) = \partial\delta_{\mathcal{W}}(x),$$

we immediately see that (9.3)–(9.4) are identical with (9.15)–(9.16). Therefore, the method (9.3)–(9.4) generates a sequence $\{x^k\}$ which is convergent to a point $\hat{x}$ at which

$$0 \in \partial f(\hat{x}) + \partial\delta_{\mathcal{W}}(\hat{x}).$$

It is an optimal solution of (9.1) if

$$\partial(f + \delta_{\mathcal{W}}) = \partial f + \partial \delta_{\mathcal{W}},$$

which can be guaranteed by the constraint qualification condition

$$\text{ri dom } f \cap \mathcal{W} \neq \emptyset. \tag{9.17}$$

The same analysis can be carried out for the final version of the Progressive Hedging Method, developed with the new scalar product (9.7) and described in Fig. 10. Let us define the matrix

$$P = \text{diag}\{\, p_s,\, s = 1, \ldots, S\}.$$

The operator

$$A_P(x) = P^{-1}\partial f(x),$$

which is the subdifferential of f in the new geometry, is maximal monotone with respect to the scalar product (9.7). The operator

$$B_P(x) = \begin{cases} P^{-1}\mathcal{W}^{\perp} & \text{if } x \in \mathcal{W}, \\ \emptyset & \text{otherwise,} \end{cases}$$

is maximal monotone in the new geometry, too. For these two operators steps (9.15)–(9.16) are equivalent to (9.8)–(9.9). Thus, under the constraint qualification condition (9.17), the sequence $\{x^k\}$ is convergent to a solution of (9.1), provided that a solution exists.

10 Bibliographic notes

The cutting plane method for two-stage linear programming is due to Benders (1962). Wets (1966) analyzed two-stage stochastic linear programming problems. Van Slyke and Wets (1969) developed for these problems the cutting plane method. The multicut version was analyzed by Birge and Louveaux (1988). Our proof of finite termination with nonbasic cuts is original. The cutting plane method for stochastic programming is dual to the decomposition method of Dantzig and Wolfe (1960). It was discussed by Dantzig and Madansky (1961). The cutting plane method for convex programming was invented by Kelley (1960). Our analysis of the general convex case follows Hiriart-Urruty and Lemaréchal (1993). Solution methods for many similar scenario subproblems were developed by Wallace (1986a,b), Haugland and Wallace (1988) and Wets (1988). Zakeri et al. (2000) discuss the

case of inexact cuts derived from approximate solutions of subproblems. Importance sampling was discussed by Glynn and Iglehart (1989). Stochastic cutting plane methods were developed by Higle and Sen (1996).

The Regularized Decomposition Method was proposed by Ruszczyński (1986). The proximal point method is due to Martinet (1970). Rockafellar (1976a) analyzed it in the context of monotone operators. Bundle methods were developed by Lemaréchal (1978), Mifflin (1982) and Kiwiel (1983) (see also the monographs by Hiriart-Urruty and Lemaréchal, 1993, and by Kiwiel, 1985). Algorithmic advances were discussed in Kiwiel (1990) and Ruszczyński and Świętanowski (1997).

Trust region methods are standard techniques in nonlinear and nonsmooth optimization (see, e.g., Moré, 1983; Kiwiel, 1996). In stochastic programming a trust region method was proposed and implemented in parallel by Linderoth and Wright (2001). Our version and its analysis are original.

Nested decomposition for multistage linear programming was introduced by Glassey (1973) and Ho and Manne (1974). For multistage stochastic programming problems the method was proposed by Birge (1985), Noël and Smeers (1986, 1987). Parallel versions were analyzed by Ruszczyński (1993), Birge et al. (1996), Nielsen and Zenios (1996) and Dempster and Thompson (1998). Chen and Powell (1999) use statistical estimates within nested decomposition. Cut sharing was discussed by Infanger and Morton (1996). Algorithmic refinements, bounds and applications are discussed by Frauendorfer (1992) and Infanger (1994).

Decomposition based on Lagrangian duality with bundle methods for solving the dual was analyzed by Chun and Robinson (1995). Bacaud et al. (2001) suggest a specialized preconditioning scheme in such approaches. General augmented Lagrangian methods are due to Hestenes (1969), Powell (1969) and Haarhoff and Buys (1970). For a modern theory with connections to proximal point methods see Rockafellar (1976b) and Bertsekas (1982). Augmented Lagrangian decomposition techniques were analyzed by Stephanopoulos and Westerberg (1975) and Cohen (1980). Scenario decomposition for stochastic programming is due to Mulvey and Ruszczyński (1995). Parallel versions were developed by Berger et al. (1994). Our analysis exploiting the sparsity of the nonanticipativity constraints follows Ruszczyński (1995).

The Progressive Hedging Method is due to Rockafellar and Wets (1991). The operator splitting method is due to Douglas and Rachford (1956). The proximal point method is due to Martinet (1970). Relations with augmented Lagrangians were analyzed by Lions and Mercier (1979) and Gabay (1983). Relations of operator splitting and proximal point methods were analyzed by Eckstein and Bertsekas (1992). The method (9.12) for nonexpansive operators in Banach spaces was developed by Opial (1967). Lawrence and Spingarn (1987) analyzed its application to operator splitting methods. For further modifications of operator splitting and their relations to bundle methods see Kiwiel et al. (1999).

References

Bacaud, L., C. Lemaréchal, A. Renaud, C. Sagastizábal (2001). Bundle methods in stochastic optimal power management: A disaggregated approach using preconditioners. *Computational Optimization and Applications* 20, 227–244.

Benders, J.F. (1962). Partitioning procedures for solving mixed-variables programming problems. *Numerische Mathematik* 4, 238–252.

Berger, A.J., J.M. Mulvey, A. Ruszczyński (1994). An extension of the DQA algorithm to convex stochastic programs. *SIAM Journal on Optimization* 4, 735–753.

Bertsekas, D.P. (1982). *Constrained Minimization and Lagrange Multiplier Methods*, Academic Press, New York.

Birge, J.R. (1985). Decomposition and partitioning methods for multistage stochastic linear programs. *Operations Research* 33, 989–1007.

Birge, J.R., C.J. Donohue, D.F. Holmes, O.G. Svintsitski (1996). A parallel implementation of the nested decomposition method for multistage stochastic linear programs. *Mathematical Programming, Series B* 75, 327–352.

Birge, J.R., F.V. Louveaux (1988). A multicut algorithm for two-stage stochastic linear programs. *European Journal of Operational Research* 34, 384–392.

Chen, Z.L., W.B. Powell (1999). Convergent cutting-plane and partial-sampling algorithm for multistage stochastic linear programs with recourse. *Journal of Optimization Theory and Applications* 102, 497–524.

Chun, B.J., S.M. Robinson (1995). Scenario analysis via bundle decomposition. *Annals of Operations Research* 56, 39–63.

Cohen, G. (1980). Auxiliary problem principle and decomposition of optimization problems. *Journal of Optimization Theory and Applications* 32, 277–305.

Dantzig, G., A. Madansky (1961). On the solution of two-stage linear programs under uncertainty, in: *Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, Berkeley, pp. 165–176.

Dantzig, G., P. Wolfe (1960). Decomposition principle for linear programs. *Operations Research* 8, 101–111.

Dempster, M.A.H., R.T. Thompson (1998). Parallelization and aggregation of nested Benders decomposition. *Annals of Operations Research* 81, 163–187.

Douglas, J., H.H. Rachford (1956). On the numerical solution of heat conduction problems in two and three space variables. *Transactions of the American Mathematical Society* 82, 421–439.

Eckstein, J., D.P. Bertsekas (1992). On the Douglas–Rachford splitting method and the proximal point algorithm for maximal monotone operators. *Mathematical Programming* 55, 293–318.

Frauendorfer, K. (1992). Stochastic Two-Stage Programming. *Lecture Notes in Economics and Mathematical Systems* 392, Springer, Berlin.

Gabay, D. (1983). Applications of the method of multipliers to variational inequalities, in: M. Fortin, R. Glowinski (eds.), *Augmented Lagrangian Methods: Applications to the Solution of Boundary-Value Problems*, North-Holland, Amsterdam.

Glassey, C.R. (1973). Nested decomposition and multistage linear programs. *Management Science* 20, 282–292.

Glynn, P.W., D.L. Iglehart (1989). Importance sampling for stochastic simulation. *Management Science* 35, 1367–1392.

Haarhoff, P.C., J.D. Buys (1970). A new method for the minimization of a nonlinear function subject to nonlinear constraints. *Computer Journal* 13, 178–184.

Haugland, D., S.W. Wallace (1988). Solving many linear programs that differ only in the righthand side. *European Journal of Operational Research* 37, 318–324.

Hestenes, M.R. (1969). Multiplier and gradient methods. *Journal of Optimization Theory and Applications* 4, 303–320.

Higle, J.L., S. Sen (1996). *Stochastic Decomposition: A Statistical Method for Large Scale Stochastic Linear Programming*, Kluwer, Dordrecht.

Hiriart-Urruty, J.-B., C. Lemaréchal (1993). *Convex Analysis and Minimization Algorithms*, Springer-Verlag, Berlin.

Ho, J.K., A.S. Manne (1974). Nested decomposition for dynamic models. *Mathematical Programming* 6, 121–140.

Infanger, G. (1994). *Planning under Uncertainty: Solving Large-Scale Stochastic Linear Programs*, Boyd & Fraser, Danvers.

Infanger, G., D.P. Morton (1996). Cut sharing for multistage stochastic linear programs with interstage dependency. *Mathematical Programming* 75, 241–256.

Kelley, J.E. (1960). The cutting plane method for solving convex programs. *J. SIAM* 8, 703–712.

Kiwiel, K.C. (1983). An aggregate subgradient method for nonsmooth convex minimization. *Mathematical Programming* 27, 320–341.

Kiwiel, K.C. (1985). *Methods of Descent for Nondifferentiable Optimization*, Lecture Notes in Mathematics 1133, Springer-Verlag, Berlin.

Kiwiel, K.C. (1990). Proximity control in bundle methods for convex nondifferentiable optimization. *Mathematical Programming* 46, 105–122.

Kiwiel, K.C. (1996). Restricted step and Levenberg–Marquardt techniques in proximal bundle methods for nonconvex nondifferentiable optimization. *SIAM Journal on Optimization* 6, 227–249.

Kiwiel, K.C., C.H. Rosa, A. Ruszczyński (1999). Decomposition via alternating linearization. *SIAM Journal on Optimization* 9, 153–172.

Lawrence, J., J.E. Spingarn (1987). On fixed points of nonexpansive piecewise isometric mappings. *Proceedings of the London Mathematical Society* 55, 605–624.

Lemaréchal, C. (1978). *Nonsmooth optimization and descent methods*. Research Report 78-2, International Institute for Applied Systems Analysis, Laxenburg.

Linderoth, J.T., S.J. Wright (2001). *Implementing a decomposition algorithm for stochastic programming on a computational grid*. Preprint ANL/MCS-P875-0401, Mathematics and Computer Science Division, Argonne National Laboratory, April 2001.

Lions, P.L., B. Mercier (1979). Splitting algorithms for the sum of two nonlinear operators. *SIAM Journal on Numerical Analysis* 16, 964–979.

Martinet, B. (1970). Régularisation d'inéquations variationelles par approximations successives. *Revue Franc. Rech. Opér.* R3, 154–158.

Mifflin, R. (1982). A modification and an extension of Lemaréchal's algorithm for nonsmooth optimization. *Mathematical Programming Study* 17, 77–90.

Moré, J.J. (1983). Recent developments in algorithms and software for trust region methods, in: A. Bachem, M. Grötschel, B. Korte (eds.), *Mathematical Programming, The State of the Art*, Springer, Berlin, pp. 258–287.

Mulvey, J.M., A. Ruszczyński (1995). A new scenario decomposition method for large-scale stochastic optimization. *Operations Research* 43, 477–490.

Nielsen, S.S., S.A. Zenios (1996). Solving multistage stochastic network programs on massively parallel computers. *Mathematical Programming* 73, 227–250.

Noël, M.-C., Y. Smeers (1986). On the use of nested decomposition for solving nonlinear multistage stochastic programs, in: *Stochastic Programming (Gargnano, 1983)*, *Lecture Notes in Control and Inform. Sci.*, Vol. 76, Springer, Berlin, pp. 235–246.

Noël, M.-C., Y. Smeers (1987). Nested decomposition of multistage nonlinear programs with recourse. *Mathematical Programming* 37, 131–152.

Opial, Z. (1967). Weak convergence for the sequence of successive approximations for nonexpansive mappings. *Bulletin of the American Mathematical Society* 73, 591–597.

Powell, M.J.D. (1969). A method for nonlinear constraints in minimizing problems, in: R. Fletcher (ed.), *Optimization*, Academic Press, New York, pp. 283–298.

Rockafellar, R.T. (1976). Monotone operators and the proximal point algorithm. *SIAM Journal on Control and Optimization* 14, 877–898.

Rockafellar, R.T. (1976). Augmented Lagrangians and applications of the proximal point algorithm in convex programming. *Mathematics of Operations Research* 1, 97–116.

Rockafellar, R.T., R.J.-B. Wets (1991). Scenarios and policy aggregation in optimization under uncertainty. *Mathematics of Operations Research* 16, 1–23.

Ruszczyński, A. (1986). A regularized decomposition method for minimizing a sum of polyhedral functions. *Mathematical Programming* 35, 309–333.

Ruszczyński, A. (1993). Parallel decomposition of multistage stochastic programming problems. *Mathematical Programming* 58, 201–228.

Ruszczyński, A. (1995). On convergence of an augmented Lagrangian decomposition method for sparse convex optimization. *Mathematics of Operations Research* 20, 634–656.

Ruszczyński, A., A. Świętanowski (1997). Accelerating the regularized decomposition method for two stage stochastic linear problems. *European Journal of Operational Research* 101, 328–342.

Stephanopoulos, G., W. Westerberg (1975). The use of Hestenes' method of multipliers to resolve dual gaps in engineering system optimization. *Journal of Optimization Theory and Applications* 15, 285–309.

Van Slyke, R., R.J.-B. Wets (1969). L-shaped linear programs with applications to optimal control and stochastic programming. *SIAM Journal on Applied Mathematics* 17, 638–663.

Wallace, S.W. (1986). Solving stochastic programs with network recourse. *Networks* 16, 295–317.

Wallace, S.W. (1986). Decomposing the requirement space of the transportation problem into polyhedral cones. *Mathematical Programming Study* 28, 29–47.

Wets, R.J.-B. (1966). Programming under uncertainty: the equivalent convex program. *SIAM Journal on Applied Mathematics* 14, 89–105.

Wets, R.J.-B. (1988). Large scale linear programming techniques in stochastic programming, in: Yu.M. Ermoliev, R.J.-B. Wets (eds.), *Numerical Techniques for Stochastic Optimization,* Springer-Verlag, Berlin, pp. 65–93.

Zakeri, G., A.B. Philpott, D.M. Ryan (2000). Inexact cuts in Benders decomposition. *SIAM Journal on Optimization* 10, 643–657.

A. Ruszczyński and A. Shapiro, Eds., *Handbooks in OR & MS, Vol. 10*

Chapter 4

Stochastic Integer Programming

François V. Louveaux

Département des Méthodes Quantitatives, Facultés Universitaires Notre Dame de la Paix, Rempart de la Vierge 8, B-5000 Namur, Belgium, E-mail: flouveaux@fundp.ac.be

Rüdiger Schultz

Institut für Mathematik, Universität Duisburg-Essen, Lotharstraße 65, D-47048 Duisburg, Germany, E-mail: schultz@math.uni-duisburg.de

Abstract

When introducing integer variables into traditional linear stochastic programs structural properties and algorithmic approaches have to be rethought from the very beginning. Employing basics from parametric integer programming and probability theory we analyze the structure of stochastic integer programs. In the algorithmic part of the paper we review solution techniques from integer programming and discuss their impact on the specialized structures met in stochastic programming.

Key words: Stochastic integer programs, mixed-integer recourse, simple integer recourse, multi-stage models, decomposition schemes, cutting planes, Lagrangian relaxation, integer *L*-shaped algorithm, sampling methods.

1 Introduction

Like in other branches of mathematical optimization, integer variables are indispensable in many stochastic programming models. Integrality either may occur explicitly via indivisibles or Boolean decisions. Or it may occur in implicit fashion as a modelling tool, for instance when handling disjunctions and discontinuities, or when dealing with nonconvex piecewise linearity.

As an example let us reconsider the popular newsboy problem: a newsboy can purchase from a publisher a number of newspapers with purchase costs

c_i and selling costs q_i, $i=1,\ldots,s$. Each newspaper has a weight a_i, and a total weight b is available to the newsboy. The demand for newspapers, which is unknown to the newsboy at the time of purchase, is described by a random vector $h(\omega)\in\mathbb{R}^s$ on some probability space $(\Omega,\mathcal{F},P)$. Unsold newspapers cannot be returned to the publisher. The newsboy faces the problem to purchase from the publisher in such a way that the profit after selling is maximal. Maximizing the expected profit leads to the stochastic program

$$\min\{c^T x+\mathbb{E}_P[Q(x,h(\omega))]\colon 0\le a^T x\le b\} \tag{1.1}$$

where

$$Q(x,h(\omega))=\min\{-q^T y\colon 0\le y\le h(\omega),\ y\le x\}. \tag{1.2}$$

The newsboy problem is a specimen from a class of cost minimization (or profit maximization) problems where planning decisions must be taken before and operational decisions are taken after observation of a random demand. It is quite common to model these problems under the assumption that the commodities involved are infinitely divisible. As in (1.1), (1.2) this leads to decision variables in the real numbers. However, when dealing with indivisibles, as the newsboy obviously does, the precise modeling would require integer variables. The model (1.1), (1.2) then has to be supplemented by the conditions $x\in\mathbb{Z}^s_+$ and $y\in\mathbb{Z}^s_+$.

Now let us assume that, for renting a newsstand, for instance, the newsboy faces fixed charge costs f when selling the newspapers. Then, disregarding divisibility or indivisibility of goods, the discontinuity of the cost structure leads to an integer variable in (1.2):

$$Q(x,h(\omega))=\min\{-q^T y+fu\colon 0\le y\le h(\omega)u,\ y\le x,\ u\in\{0,1\}\}.$$

In the simple examples discussed above integrality, of course, does not pose a real challenge. This changes drastically if the above modeling paradigms occur in more complex situations. In principle, any integer or mixed-integer programming model arising in operations research may be affected by randomness and hence give rise to a stochastic programming extension.

In the present chapter we will see that the appearance of integer requirements has substantial structural and algorithmic consequences for stochastic programming models as discussed in Chapters 1 and 2, for instance.

2 Structural properties

2.1 Two-stage models with complete integer recourse

Introducing integer requirements into the two-stage stochastic linear program from Section 2 of Chapter 2 leads to the optimization problem

$$\min_x \{c^T x + \mathbb{E}_P[Q(x, \xi(\omega))] : x \in X\} \tag{2.3}$$

where $Q(x, \xi(\omega))$ is the optimal value of the second-stage problem

$$\min_{(y,y')} \{q^T y + q'^T y' : Tx + Wy + W'y' = h(\omega),\ y \in \mathbb{Z}_+^{\bar{n}_2},\ y' \in \mathbb{R}_+^{n_2'}\}. \tag{2.4}$$

We assume that all ingredients above have conformable dimensions, that W, W' are rational matrices, and that $X \subseteq \mathbb{R}^{n_1}$ is a nonempty closed polyhedron, possibly involving integrality constraints on components of the vector x. For ease of exposition, q, q', W, W', T are deterministic, such that the random variable $\xi(\omega) = h(\omega)$, living on some probability space $(\Omega, \mathcal{F}, P)$, is the only stochastic ingredient. So far, most of the subsequent results on stochastic integer programs were obtained for that special case. Where appropriate, we will point to existing results under more general randomness.

As with the two-stage stochastic linear program, the second-stage value function $\Phi : \mathbb{R}^{m_2} \to \mathbb{R}$ with

$$\Phi(t) := \min_{(y,y')} \{q^T y + q'^T y' : Wy + W'y' = t,\ y \in \mathbb{Z}_+^{\bar{n}_2},\ y' \in \mathbb{R}_+^{n_2'}\} \tag{2.5}$$

is of prime importance for the structure of (2.3). Recall that, without integer requirements, Φ is convex on its domain of finiteness. To obtain a first impression on the impact of integrality let us start with two illustrative examples. They are both derived from

$$\Phi(t) = \min\{y^+ + y^- : y^+ - y^- = t,\ y^+ \in \mathbb{R}_+,\ y^- \in \mathbb{R}_+\}. \tag{2.6}$$

This is the classical simple-recourse situation. By duality one immediately obtains that $\Phi(t) = |t|$.

Now we add another variable v which we require to be integral:

$$\Phi(t) = \min\left\{\frac{1}{2}\ v+y^+ +y^- : v + y^+ - y^- = t,\ v \in \mathbb{Z}_+,\ y^+ \in \mathbb{R}_+,\ y^- \in \mathbb{R}_+\right\}$$
$$= \min\left\{\frac{1}{2}\ v + |t - v| : v \in \mathbb{Z}_+\right\}.$$

Understanding $v \in \mathbb{Z}_+$ as indices, the function Φ can be seen as the pointwise minimum of staggered absolute values. Hence, Φ is no longer convex but still continuous, even Lipschitz continuous on $\mathbb{R}$.

The second example is basically derived by turning y^+ and y^- into integer variables:

$$\Phi(t) = \min\{v^+ + v^- : y + v^+ - v^- = t,\ y \in \mathbb{R}_+,\ v^+ \in \mathbb{Z}_+,\ v^- \in \mathbb{Z}_+\}$$
$$= \begin{cases} 0, & t \geq 0 \\ -t, & t < 0. \end{cases}$$

Here $\lceil \cdot \rceil$ denotes the integer round-up operation. Obviously, Φ is no longer continuous, but still lower semicontinuous. Moreover, discontinuities occur in a set of Lebesgue measure zero only, and "jump heights" at discontinuities are globally bounded.

From these observations it becomes quite clear that, with stochastic integer programs, basic properties like convexity and duality, that are so important in the purely linear case, cannot be maintained for reasonable problem classes. Let us now study the mixed-integer value function in more detail.

Proposition 1. *Suppose that the recourse in (2.3)–(2.4) is complete, i.e.,*

$$W(\mathbb{Z}_+^{\bar{n}_2}) + W'(\mathbb{R}_+^{n_2'}) = \mathbb{R}^{m_2}, \tag{2.7}$$

and that the LP relaxation to the second-stage problem has a feasible dual, i.e.,

$$\{u \in \mathbb{R}^{m_2} : W^T u \leq q,\ W'^T u \leq q'\} \neq \emptyset, \tag{2.8}$$

then the value function $\Phi(\cdot)$ is well-defined on $\mathbb{R}^{m_2}$.

Proof. Let $t \in \mathbb{R}^{m_2}$. By (2.7), the mixed-integer program defining $\Phi(t)$ is feasible. By the existence theorem of mixed-integer linear programming this problem then is solvable provided it is bounded. The latter, however, follows from the solvability of the LP relaxation to the mixed-integer program defining $\Phi(t)$ which is a consequence of the primal feasibility implied by (2.7) and the dual feasibility in (2.8). □

Before proceeding further, let us have a quick look at the linear-programming counterpart to Φ that has been studied in more detail in Chapter 2:

$$\Phi_{\text{lin}}(t) := \min\{q'^T y' : W'y' = t,\ y' \in \mathbb{R}^{n_2'}_+\}. \tag{2.9}$$

If we assume that $W'(\mathbb{R}^{n_2'}_+)$ is full-dimensional and that

$$\{u \in \mathbb{R}^{m_2} : W'^T u \le q'\} \neq \emptyset,$$

then the latter set has vertices d_k, $k = 1,\ldots,K$, and it holds by linear programming duality that

$$\Phi_{\text{lin}}(t) = \max\{t^T u : W'^T u \le q'\} = \max_{k=1,\ldots,K} d_k^T t \quad \text{for all } t \in W'(\mathbb{R}^{n_2'}_+).$$

Hence, Φ_{lin} is convex and piecewise linear on its (conical) domain of definition.

Imposing the basic assumptions (2.7) and (2.8) we obtain

$$\begin{aligned}\Phi(t) &= \min\{q^T y + q'^T y' : Wy + W'y' = t,\ y \in \mathbb{Z}^{\bar{n}_2}_+,\ y' \in \mathbb{R}^{n_2'}_+\}\\ &= \min_y\{q^T y + \min_{y'}\{q'^T y' : W'y' = t - Wy,\ y' \in \mathbb{R}^{n_2'}_+\} : y \in \mathbb{Z}^{\bar{n}_2}_+\}\\ &= \min_y\{\Phi_y(t) : y \in \mathbb{Z}^{\bar{n}_2}_+\},\end{aligned} \tag{2.10}$$

where

$$\Phi_y(t) = q^T y + \max_{k=1,\ldots,K} d_k^T(t - Wy) \quad \text{for all } t \in Wy + W'(\mathbb{R}^{n_2'}_+).$$

Here, d_k, $k = 1,\ldots,K$ are the vertices of $\{u \in \mathbb{R}^{m_2} : W'^T u \le q'\}$, and we have applied the argument about Φ_{lin} from the purely linear case. For $t \notin Wy + W'(\mathbb{R}^{n_2'}_+)$ the problem $\min_{y'}\{q'^T y' : W'y' = t - Wy,\ y' \in \mathbb{R}^{n_2'}_+\}$ is infeasible, and we put $\Phi_y(t) = +\infty$. It is convenient to introduce the notation $Y(t) := \{y \in \mathbb{Z}^{\bar{n}_2}_+ : \Phi_y(t) < +\infty\}$.

According to (2.10) the value function Φ is made up by the pointwise minimum of a family of convex, piecewise linear functions whose domains of definition are polyhedral cones arising as shifts of the cone $W'(\mathbb{R}^{n_2'}_+)$. By our basic assumption $W(\mathbb{Z}^{\bar{n}_2}_+) + W'(\mathbb{R}^{n_2'}_+) = \mathbb{R}^{m_2}$, the cone $W'(\mathbb{R}^{n_2'}_+)$ is full-dimensional.

Some first conclusions about the continuity of Φ may be drawn from the above observations:

(1) Suppose that $t \in \mathbb{R}^{m_2}$ does not belong to any boundary of any of the sets $Wy + W'(\mathbb{R}^{n_2'}_+)$, $y \in \mathbb{Z}^{\bar{n}_2}$. Then the same is true for all points τ in some open ball B around t. Hence, $Y(\tau) = Y(t)$ for all $\tau \in B$. With an enumeration $(y_\nu)_{\nu \in \mathbb{N}}$ of $Y(t)$ we consider the functions $\Phi^\kappa(\tau) := \min\{\Phi_{y_\nu}(\tau) : \nu \leq \kappa\}$ for all $\tau \in B$. Then $\lim_{\kappa \to \infty} \Phi^\kappa(\tau) = \Phi(\tau)$ for all $\tau \in B$. Since, for any function Φ_y, its "slopes" are determined by the same, finitely many vectors d_k, $k = 1, \ldots, K$, the functions Φ^κ, $\kappa \in \mathbb{N}$ are all Lipschitz continuous on B with a uniform Lipschitz constant. Thus, the family of functions Φ^κ, $\kappa \in \mathbb{N}$ is equicontinuous on B and has a pointwise limit there. Consequently, this pointwise limit Φ is continuous on B, in fact Lipschitz continuous with the mentioned uniform constant.

(2) Any discontinuity point of Φ must be located at the boundary of some set $Wy + W'(\mathbb{R}^{n_2'}_+)$, $y \in \mathbb{Z}^{\bar{n}_2}$. Hence, the set of discontinuity points of Φ is contained in a countable union of hyperplanes. Since $W'(\mathbb{R}^{n_2'}_+)$ has only finitely many facets, this union of hyperplanes subdivides into finitely many classes, such that, in each class, the hyperplanes are parallel. By the rationality of the matrices W and W', within each class, the pairwise distance of the hyperplanes is uniformly bounded below by some positive number.

(3) Let $t_\nu \to t$ and $y \in \mathbb{Z}^{\bar{n}_2}$ such that $t_\nu \in Wy + W'(\mathbb{R}^{n_2'}_+)$ for all sufficiently large ν. Since the set $Wy + W'(\mathbb{R}^{n_2'}_+)$ is closed, this yields $t \in Wy + W'(\mathbb{R}^{n_2'}_+)$. Therefore, for sufficiently large ν, $Y(t_\nu) \subseteq Y(t)$. This paves the way for showing that $\liminf_{t_\nu \to t} \Phi(t_\nu) \geq \Phi(t)$, which is the lower semicontinuity of Φ at t.

The above analysis can be extended into the following result that dates back to a series of papers by Blair and Jeroslow, out of which we refer to Blair and Jeroslow (1977), and to the monographs Bank et al. (1982) and Bank and Mandel (1988).

Proposition 2. *Let W, W' be matrices with rational entries and assume that $W(\mathbb{Z}^{\bar{n}_2}_+) + W'(\mathbb{R}^{n_2'}_+) = \mathbb{R}^{m_2}$ as well as $\{u \in \mathbb{R}^{m_2} : W^T u \leq q, W'^T u \leq q'\} \neq \emptyset$. Then it holds*

(1) *Φ is real-valued and lower semicontinuous on $\mathbb{R}^{m_2}$,*

(2) *there exists a countable partition $\mathbb{R}^{m_2} = \cup_{i=1}^{\infty} \mathcal{T}_i$ such that the restrictions of Φ to $\mathcal{T}_i$ are piecewise linear and Lipschitz continuous with a uniform constant $L > 0$ not depending on i,*

(3) *each of the sets $\mathcal{T}_i$ has a representation $\mathcal{T}_i = \{t_i + \mathcal{K}\} \setminus \cup_{j=1}^{N} \{t_{ij} + \mathcal{K}\}$ where $\mathcal{K}$ denotes the polyhedral cone $W'(\mathbb{R}^{n_2'}_+)$ and t_i, t_{ij} are suitable points from $\mathbb{R}^{m_2}$, moreover, N does not depend on i,*

(4) *there exist positive constants β, γ such that $|\Phi(t_1) - \Phi(t_2)| \leq \beta \|t_1 - t_2\| + \gamma$ whenever $t_1, t_2 \in \mathbb{R}^{m_2}$.*

Let us now consider the expected value function

$$\phi(x) := \mathbb{E}_P[Q(x, \xi(\omega))] = \mathbb{E}_P[\Phi(\xi(\omega) - Tx)] = \mathbb{E}_\mu[\Phi(\xi - Tx)] \qquad (2.11)$$

where μ denotes the image measure $P \circ \xi^{-1}$ on $\mathbb{R}^{m_2}$. Thanks to the lower semicontinuity in part (1) of Proposition 2, the integrand $Q(x, \cdot)$ is measurable for all $x \in \mathbb{R}^{n_1}$. Moreover, lower semicontinuity and continuity are inherited according to the general principles displayed in Proposition 14 of Chapter 1 and Proposition 1 of Chapter 2. This is made precise in the following proposition.

Proposition 3. *Suppose that:* (i) W, W' *are rational*, (ii) $W(\mathbb{Z}_+^{\bar{n}_2}) + W'(\mathbb{R}_+^{n'_2}) = \mathbb{R}^{m_2}$, (iii) $\{u \in \mathbb{R}^{m_2} : W^T u \leq q, W'^T u \leq q'\} \neq \emptyset$, (iv) $\mathbb{E}_\mu[\|\xi\|] < \infty$. *Then the expected value function* ϕ *is lower semicontinuous on* $\mathbb{R}^{n_1}$. *If, moreover,* $\mu(E(x)) = 0$ *where* $E(x) := \{\xi \in \mathbb{R}^{m_2} : \Phi \text{ is discontinuous at } \xi - Tx\}$, *then* ϕ *is continuous at* x.

Proof. Let $x \in \mathbb{R}^{n_1}$ and $x_k \to x$. Denote $r := \max_{k \in \mathbb{N}} \|x_k\|$. Assumptions (ii) and (iii) in particular imply that $\Phi(0) = 0$. Part (4) of Proposition 2 then yields the estimate

$$\begin{aligned} \Phi(\xi - Tx_k) &\geq \Phi(0) - |\Phi(\xi - Tx_k) - \Phi(0)| \geq -\beta\|\xi - Tx_k\| - \gamma \\ &\geq -\beta\|\xi\| - \beta r\|T\| - \gamma \end{aligned}$$

such that, together with (iv), the function $-\beta\|\cdot\| - \beta r\|T\| - \gamma$ provides an integrable minorant for the functions $\Phi(\cdot - Tx_k)$, $k \in \mathbb{N}$. Fatou's Lemma and the lower semicontinuity of Φ now imply

$$\begin{aligned} \phi(x) &= \int_{\mathbb{R}^{m_2}} \Phi(\xi - Tx)\mu(\mathrm{d}\xi) \leq \int_{\mathbb{R}^{m_2}} \liminf_{k \to \infty} \Phi(\xi - Tx_k)\mu(\mathrm{d}\xi) \\ &\leq \liminf_{k \to \infty} \int_{\mathbb{R}^{m_2}} \Phi(\xi - Tx_k)\mu(\mathrm{d}\xi) = \liminf_{k \to \infty} \phi(x_k) \end{aligned}$$

which proves the desired lower semicontinuity. For proving continuity we again employ part (4) of Proposition 2 together with $\Phi(0) = 0$. We obtain the estimate

$$|\Phi(\xi - Tx_k)| = |\Phi(\xi - Tx_k) - \Phi(0)| \leq \beta\|\xi\| + \beta r\|T\| + \gamma$$

which, in view of (iv), provides us with an integrable majorant for the functions $|\Phi(\cdot - Tx_k)|$, $k \in \mathbb{N}$. By $\mu(E(x)) = 0$, we have

$$\lim_{k\to\infty} \Phi(\xi - Tx_k) = \Phi(\xi - Tx) \quad \text{for } \mu\text{-almost all } \xi \in \mathbb{R}^{m_2}.$$

Now Lebesgue's Dominated Convergence Theorem completes the proof:

$$\lim_{k\to\infty} \phi(x_k) = \lim_{k\to\infty} \int_{\mathbb{R}^{m_2}} \Phi(\xi - Tx_k)\mu(\mathrm{d}\xi) = \int_{\mathbb{R}^{m_2}} \Phi(\xi - Tx)\mu(\mathrm{d}\xi) = \phi(x).$$

□

Recall that, by conclusion (1.2) in front of Proposition 2, the set of discontinuity points of Φ, and hence $E(x)$, is contained in a countable union of hyperplanes, which is a set of Lebesgue measure zero. Assume that μ has a density and that (i)–(iv) are fulfilled. Then $\mu(E(x)) = 0$ for all $x \in \mathbb{R}^{n_1}$, and ϕ is continuous on $\mathbb{R}^{n_1}$. The next statement addresses Lipschitz continuity. To avoid further technicalities, we formulate the result for the case that μ has a density.

Proposition 4. *Suppose that:* (i) q, q', W, W' *all have rational entries*, (ii) $W(\mathbb{Z}_+^{\bar{m}_2}) + W'(\mathbb{R}_+^{m_2}) = \mathbb{R}^{m_2}$, (iii) $\{u \in \mathbb{R}^{m_2} : W^T u \le q,\ W'^T u \le q'\} \neq \emptyset$, (iv) $\mathbb{E}_\mu[\|\xi\|] < \infty$, (v) μ *has a density and for any nonsingular linear transformation* $B \in L(\mathbb{R}^{m_2}, \mathbb{R}^{m_2})$ *all one-dimensional marginal distributions of* $\mu \circ B$ *have bounded densities which, outside some bounded interval, are monotonically decreasing with growing absolute value of the argument. Then* ϕ *is Lipschitz continuous on any bounded subset of* $\mathbb{R}^{n_1}$.

Proof. Let x', x'' belong to some bounded subset D of $\mathbb{R}^{n_1}$. Denote

$$S(x', x'') := \bigcup_{i\in\mathbb{N}} ((Tx' + \mathcal{T}_i) \cap (Tx'' + \mathcal{T}_i))$$

where $\mathcal{T}_i$ are as in part (2) of Proposition 2. Note that for all $\xi \in S(x', x'')$ we then have the estimate

$$|\Phi(\xi - Tx') - \Phi(\xi - Tx'')| \le L \cdot \|Tx' - Tx''\|. \tag{2.12}$$

Now

$$|\phi(x') - \phi(x'')| \leq \int_{S(x',x'')} |\Phi(\xi - Tx') - \Phi(\xi - Tx'')|\mu(\mathrm{d}\xi) + \int_{\mathbb{R}^{m_2}\setminus S(x',x'')} |\Phi(\xi - Tx') - \Phi(\xi - Tx'')|\mu(\mathrm{d}\xi).$$

Employing (2.12) and part (4) of Proposition 2 gives the estimate

$$|\phi(x') - \phi(x'')| \leq (L + \beta) \cdot \|T\| \cdot \|x' - x''\| + \gamma \cdot \mu(\mathbb{R}^{m_2}\setminus S(x', x'')).$$

We complete the proof by deducing a Lipschitz estimate for the second term on the right. Let H_k, $k = 1,\ldots, K$ be the hyperplanes containing the facets of the cone $W'(\mathbb{R}^{m_2}_+)$ that arises in part (3) of Proposition 2 and is full-dimensional by (ii). The set $\mathbb{R}^{m_2}\setminus S(x', x'')$ then is contained in a finite union of sets $\mathcal{H}_k$, $k = 1,\ldots, K$ each of which is a countable union of "sandwiches" $\mathcal{H}_{k,i}$, $i \in \mathbb{N}$. Each "sandwich" $\mathcal{H}_{k,i}$ is the region in between and including the affine hyperplanes

$$t_i + Tx' + H_k \quad \text{and} \quad t_i + Tx'' + H_k$$

where t_i, $i \in \mathbb{N}$, are as in part (3) of Proposition 2. By a nonsingular linear transformation B_k we map the hyperplane H_k to the hyperplane which is orthogonal to the first coordinate vector. Then $\mu(\mathcal{H}_k)$ can be estimated using the marginal density θ_k of the first component with respect to the image measure $\mu \circ B_k^{-1}$:

$$\mu(\mathcal{H}_k) \leq c_1 \cdot \sum_{i\in\mathbb{N}} \int_{\tau_{i,k}(x')}^{\tau_{i,k}(x'')} \theta_k(\tau)\, \mathrm{d}\tau.$$

Here, $c_1 > 0$ is some constant, and $\tau_{i,k}(x')$, $\tau_{i,k}(x'')$ are the first components of $B_k(t_i + Tx')$ and $B_k(t_i + Tx'')$, respectively. Without loss of generality we assume that $\tau_{i,k}(x') < \tau_{i,k}(x'')$. Clearly, there exists a constant $c_2 > 0$ such that

$$\tau_{i,k}(x'') - \tau_{i,k}(x') \leq c_2 \cdot \|x' - x''\|. \tag{2.13}$$

Using the rationality of q, q', W, W' one can show that the sequence $(\tau_{i,k}(x'))_{i\in\mathbb{N}}$ has no accumulation points. Since x', x'' belong to the bounded set D, there exists an index $\bar{i} = \bar{i}(D)$, independent of x', x'', such that the intervals $[\tau_{i,k}(x'), \tau_{i,k}(x'')]$ (up to renumbering) meet the bounded interval from assumption (v) at most of $i \leq \bar{i}$.

According to assumption (v) we have an upper bound $\overline{\theta}_k$ for $\theta_k(\cdot)$. For $i > \bar{i}$, we denote by $\tilde{\tau}_{i,k}$ the left or right endpoint of $[\tau_{i,k}(x'), \tau_{i,k}(x'')]$ depending on whether θ_k is decreasing or increasing on that interval. This allows the estimate

$$\sum_{i\in\mathbb{N}} \int_{\tau_{i,k}(x')}^{\tau_{i,k}(x'')} \theta_k(\tau)\,\mathrm{d}\tau \le \sum_{i\le\bar{i}} \overline{\theta}_k \cdot (\tau_{i,k}(x'') - \tau_{i,k}(x')) + \sum_{i>\bar{i}} \theta_k(\tilde{\tau}_{i,k}) \cdot (\tau_{i,k}(x'') - \tau_{i,k}(x')).$$

In view of (2.13), there exists a constant $c_3 > 0$ such that the first sum on the right is estimated above by $c_3\|x'-x''\|$. For the second sum we obtain the upper estimate $c_2 \sum_{i>\bar{i}} \theta_k(\tilde{\tau}_{i,k}) \cdot \|x' - x''\|$. It remains to show that $\sum_{i>\bar{i}} \theta_k(\tilde{\tau}_{i,k})$ is finite.

Let us do so for the sum over all $i > \bar{i}$ belonging to those $\tilde{\tau}_{i,k}$ around which θ_k is decreasing. For the remaining $i > \bar{i}$ a similar argument applies. Since the $\tilde{\tau}_{i,k}$ do not accumulate, it holds with some $\varepsilon > 0$

$$1 \ge \sum_i \int_{\tilde{\tau}_{i,k}-\varepsilon}^{\tilde{\tau}_{i,k}} \theta_k(\tau)\,\mathrm{d}\tau \ge \sum_i \int_{\tilde{\tau}_{i,k}-\varepsilon}^{\tilde{\tau}_{i,k}} \theta_k(\tilde{\tau}_{i,k})\,\mathrm{d}\tau = \varepsilon \cdot \sum_i \theta_k(\tilde{\tau}_{i,k}),$$

providing the desired finiteness. Repeating the above arguments for all $\mathcal{H}_k$, $k = 1,\ldots,K$, one confirms that there exists a constant $c > 0$ such that

$$\mu(\mathbb{R}^{m_2}\backslash S(x', x'')) \le \sum_{k=1}^{K} \mu(\mathcal{H}_k) \le c \cdot \|x' - x''\|,$$

and the proof is complete. □

The following examples show that the boundedness and the monotonicity in assumption (v) of the above proposition are indispensable.

Example 5. Let $\Phi(t) = \min\{y : y \ge t, y \in \mathbb{Z}\} = \lceil t \rceil$ and μ be given by the (unbounded) density θ, for which $\theta(\tau) = \tau^{-1/2}$ if $0 < \tau \le 1/4$ and $\theta(\tau) = 0$, otherwise. Then $\phi(x) = 1$ for $-3/4 \le x \le 0$ and $\phi(x) = 1 - 2\sqrt{x}$ if $0 \le x \le 1/4$. This function is not Lipschitz continuous on neighborhoods of $x_0 = 0$.

Example 6. Let $\Phi(t) = \min\{y : y \ge t, y \in \mathbb{Z}\} = \lceil t \rceil$ and μ be given by the density θ, for which $\theta(\tau) = 1/\nu$ if $\tau \in [\nu, \nu + \frac{1}{\nu^2} \cdot c]$, $c := (\sum_{k=1}^{\infty} 1/k^3)^{-1}$, $\nu \in \mathbb{N}$, and $\theta(\tau) = 0$, otherwise. This density violates the monotonicity assumption in

Proposition 4. We show that ϕ is not Lipschitz continuous on neighborhoods of $x_0 = 0$.

Assumptions (i)–(iii) of Proposition 4 are clearly met. Moreover, θ is bounded and assumption (iv) is fulfilled, since

$$\int_{-\infty}^{+\infty} \tau\theta(\tau)\,\mathrm{d}\tau \le \sum_{\nu=1}^{\infty} (\nu+1)\cdot\frac{1}{\nu}\cdot\frac{1}{\nu^2}\cdot c = c\,\frac{\pi^2}{6}+1.$$

For arbitrary $x \in \mathbb{R}$, $0 < x < 1$, it holds

$$\phi(0) - \phi(x) = \sum_{\nu=1}^{\infty} \int_{\nu}^{\nu+x} \theta(\tau)\,\mathrm{d}\tau \ge \sum_{\nu=1}^{\overline{\nu}(x)} \frac{1}{\nu}\cdot x,$$

where $\overline{\nu}(x) := \lfloor\sqrt{\frac{c}{x}}\rfloor$. Consider $x_k := \frac{1}{k^2}\cdot c$, $k \in \mathbb{N}$. The above yields

$$\frac{1}{x_k}(\phi(0) - \phi(x_k)) \ge \sum_{\nu=1}^{k} \frac{1}{\nu}.$$

For $k \to \infty$, the left-hand side tends to infinity, proving that ϕ is not Lipschitz continuous on neighborhoods of $x_0 = 0$.

When studying the stability behaviour of the stochastic program (2.3) with respect to perturbations of the underlying probability measure, see the chapter "Stability of Stochastic Programming Problems" for an exposition of stability analysis, it is crucial to detect the continuity of the expected recourse function ϕ, jointly in the decision vector x and the probability measure μ. As a prerequisite, then a suitable convergence notion for probability measures is needed. Here, weak convergence of probability measures has proven both sufficiently general to cover relevant applications and sufficiently specific to enable substantial results. A sequence $(\mu_k)_{k\in\mathbb{N}}$ in the space $\mathcal{P}(\mathbb{R}^{m_2})$ of Borel probability measures on $\mathbb{R}^{m_2}$ is said to converge weakly to $\mu \in \mathcal{P}(\mathbb{R}^{m_2})$, written $\mu_k \xrightarrow{w} \mu$, if for any bounded continuous function $g : \mathbb{R}^{m_2} \to \mathbb{R}$ we have

$$\int_{\mathbb{R}^{m_2}} g(\xi)\mu_k(\mathrm{d}\xi) \to \int_{\mathbb{R}^{m_2}} g(\xi)\mu(\mathrm{d}\xi) \quad \text{as} \quad k \to \infty. \tag{2.14}$$

For establishing joint continuity of $\phi = \phi(x, \mu)$ a theorem on weak convergence of image measures attributed to Rubin will be very useful. This theorem says: Let g_k, g ($k \in \mathbb{N}$) be measurable functions from $\mathbb{R}^{m_2}$ to $\mathbb{R}$ and

denote $E := \{\xi \in \mathbb{R}^{m_2} : \exists \xi_k \to \xi \text{ such that } g_k(\xi_k) \not\to g(\xi)\}$. If $\mu_k \xrightarrow{w} \mu$ and $\mu(E) = 0$, then $\mu_k \circ g_k^{-1} \xrightarrow{w} \mu \circ g^{-1}$.

Proposition 7. *Fix arbitrary* $p > 1$ *and* $K > 0$, *and denote* $\Delta_{p,K}(\mathbb{R}^{m_2}) := \{\nu \in \mathcal{P}(\mathbb{R}^{m_2}) : \mathbb{E}_\nu[\|\xi\|^p] \le K\}$. *Let* $\mu \in \Delta_{p,K}(\mathbb{R}^{m_2})$ *be such that* $\mu(E(x)) = 0$. *Then the function* $\phi : \mathbb{R}^{n_1} \times \Delta_{p,K}(\mathbb{R}^{m_2}) \to \mathbb{R}$ *is continuous at* (x, μ).

Proof. Let $x_k \to x$ in $\mathbb{R}^{n_1}$ and $\mu_k \xrightarrow{w} \mu$ in $\Delta_{p,K}(\mathbb{R}^{m_2})$. Introduce measurable functions g_k, $k \in \mathbb{N}$, and g by $g_k(\xi) := \Phi(\xi - Tx_k)$ and $g(\xi) := \Phi(\xi - Tx)$. For the corresponding exceptional set E a simple continuity argument provides $E(x)^c \subseteq E^c$ or, equivalently, $E \subseteq E(x)$. Hence, $\mu(E) = 0$, and Rubin's Theorem yields

$$\mu_k \circ g_k^{-1} \xrightarrow{w} \mu \circ g^{-1}. \tag{2.15}$$

Changing variables in the assertion

$$\lim_{n\to\infty} \int_{\mathbb{R}^{m_2}} g_k(\xi)\mu_k(\mathrm{d}\xi) = \int_{\mathbb{R}^{m_2}} g(\xi)\mu(\mathrm{d}\xi)$$

yields the equivalent statement

$$\lim_{n\to\infty} \int_{\mathbb{R}} \tau\mu_k \circ g_k^{-1}(\mathrm{d}\tau) = \int_{\mathbb{R}} \tau\mu \circ g^{-1}(\mathrm{d}\tau).$$

For fixed $a \in \mathbb{R}_+$, consider the truncation $\kappa_a : \mathbb{R} \to \mathbb{R}$ with

$$\kappa_a(\tau) := \begin{cases} \tau, & |\tau| < a \\ 0, & |\tau| \ge a. \end{cases}$$

Now

$$\begin{aligned} &\Big| \int_{\mathbb{R}} \tau\mu_k \circ g_k^{-1}(\mathrm{d}\tau) - \int_{\mathbb{R}} \tau\mu \circ g^{-1}(\mathrm{d}\tau) \Big| \\ &\quad \le \Big| \int_{\mathbb{R}} (\tau - \kappa_a(\tau))\mu_k \circ g_k^{-1}(\mathrm{d}\tau) \Big| + \Big| \int_{\mathbb{R}} \kappa_a(\tau)\mu_k \circ g_k^{-1}(\mathrm{d}\tau) \\ &\qquad - \int_{\mathbb{R}} \kappa_a(\tau)\mu \circ g^{-1}(\mathrm{d}\tau) \Big| + \Big| \int_{\mathbb{R}} (\kappa_a(\tau) - \tau)\mu \circ g^{-1}(\mathrm{d}\tau) \Big|. \end{aligned} \tag{2.16}$$

The proof is completed by showing that, for a given $\varepsilon > 0$, each of the three expressions on the right becomes less than $\varepsilon/3$ provided that n and a are sufficiently large.

For the first expression we obtain

$$\left| \int_{\mathbb{R}} (\tau - \kappa_a(\tau)) \mu_k \circ g_k^{-1}(\mathrm{d}\tau) \right| \leq \int_{\{\tau : |\tau| \geq a\}} |\tau| \mu_k \circ g_k^{-1}(\mathrm{d}\tau)$$
$$= \int_{\{\xi : |g_k(\xi)| \geq a\}} |g_k(\xi)| \mu_k(\mathrm{d}\xi). \tag{2.17}$$

Since $p > 1$,

$$\int_{\mathbb{R}^{m_2}} |g_k(\xi)|^p \mu_k(\mathrm{d}\xi) \geq \int_{\{\xi : |g_k(\xi)| \geq a\}} |g_k(\xi)| \cdot |g_k(\xi)|^{p-1} \mu_k(\mathrm{d}\xi)$$
$$\geq a^{p-1} \int_{\{\xi : |g_k(\xi)| \geq a\}} |g_k(\xi)| \mu_k(\mathrm{d}\xi). \tag{2.18}$$

Therefore, the estimate in (2.17) can be continued by

$$\leq a^{1-p} \int_{\mathbb{R}^{m_2}} |g_k(\xi)|^p \mu_k(\mathrm{d}\xi). \tag{2.19}$$

Proposition 2, part (4), and $g_k(0) = 0$ imply

$$|g_k(\xi)|^p \leq (\beta \|\xi\| + \beta \|x_k\| \cdot \|T\| + \gamma)^p.$$

Since $(x_k)_{k \in \mathbb{N}}$ is bounded and all μ_k belong to $\Delta_{p,K}(\mathbb{R}^{m_2})$, there exists a positive constant c such that

$$\int_{\mathbb{R}^{m_2}} |g_k(\xi)|^p \mu_k(\mathrm{d}\xi) \leq c \quad \text{for all } k \in \mathbb{N}.$$

Hence, (2.19) can be estimated above by c/a^{p-1} which becomes less than $\varepsilon/3$ if a is sufficiently large.

We now turn to the second expression in (2.16). Since every probability measure on the real line has at most countably many atoms, we obtain that $\mu \circ g^{-1}(\{\tau : |\tau| = a\}) = 0$ for (Lebesgue-) almost all $a \in \mathbb{R}$. Therefore, κ_a is a measurable function whose set of discontinuity points D_{κ_a} has $\mu \circ g^{-1}$-measure zero for almost all $a \in \mathbb{R}$. We apply Rubin's Theorem to the weakly convergent sequence $\mu_k \circ g_k^{-1} \xrightarrow{w} \mu \circ g^{-1}$, cf. (2.15), and the identical sequence of functions κ_a. The role of the exceptional set then is taken by D_{κ_a}, and Rubin's Theorem is working due to $\mu \circ g^{-1}(D_{\kappa_a}) = 0$. This yields the conclusion

$$\mu_k \circ g_k^{-1} \circ \kappa_a^{-1} \xrightarrow{w} \mu \circ g^{-1} \circ \kappa_a^{-1} \quad \text{for almost all } a \in \mathbb{R}. \tag{2.20}$$

Consider the bounded continuos function $\eta : \mathbb{R} \to \mathbb{R}$ given by

$$\eta(\tau') := \begin{cases} -a, & \tau' \leq -a \\ \tau', & -a \leq \tau' \leq a \\ a, & \tau' \geq a. \end{cases}$$

By the weak convergence in (2.20), we obtain for $n \to \infty$

$$\int_{\mathbb{R}} \eta(\tau') \mu_k \circ g_k^{-1} \circ \kappa_a^{-1}(\mathrm{d}\tau') \to \int_{\mathbb{R}} \eta(\tau') \mu \circ g^{-1} \circ \kappa_a^{-1}(\mathrm{d}\tau'). \tag{2.21}$$

Changing variables provides

$$\begin{aligned} \int_{\mathbb{R}} \eta(\tau') \mu_k \circ g_k^{-1} \circ \kappa_a^{-1}(\mathrm{d}\tau') &= \int_{\kappa_a^{-1}(\mathbb{R})} \eta(\kappa_a(\tau)) \mu_k \circ g_k^{-1}(\mathrm{d}\tau) \\ &= \int_{\mathbb{R}} \kappa_a(\tau) \mu_k \circ g_k^{-1}(\mathrm{d}\tau). \end{aligned}$$

Analogously,

$$\int_{\mathbb{R}} \eta(\tau') \mu \circ g^{-1} \circ \kappa_a^{-1}(\mathrm{d}\tau') = \int_{\mathbb{R}} \kappa_a(\tau) \mu \circ g^{-1}(\mathrm{d}\tau).$$

The above identities together with (2.21) confirm that the second expression on the right-hand side of (2.16) becomes arbitrarily small for sufficiently large n and almost all sufficiently large a.

Let us finally turn to the third expression in (2.16). Analogously to (2.17), (2.18) and (2.19) we obtain

$$\left| \int_{\mathbb{R}} (\kappa_a(\tau) - \tau)\mu \circ g^{-1}(\mathrm{d}\tau) \right| \leq a^{1-p} \int_{\mathbb{R}^{m_2}} |g(\xi)|^p \mu(\mathrm{d}\xi).$$

The integral $\int_{\mathbb{R}^{m_2}} |g(\xi)|^p \mu(\mathrm{d}\xi)$ is finite due to part (4) of Proposition 2 and $\mathbb{E}_\mu[\|\xi\|^p] \leq K$. Hence, the third expression in (2.16) becomes less than $\varepsilon/3$ if a is large enough. □

Employing well-established arguments of parametric optimization Proposition 7, together with standard assumptions such as boundedness of the unperturbed solution set, leads to (qualitative) continuity of the (multi-) functions assigning to the underlying probability measure the optimal value and the set of optimal solutions to (2.3), respectively. Quantitative continuity of $\phi(\cdot, \cdot)$ and quantitative stability of (2.3) require the identification of suitable distances on the space $\mathcal{P}(\mathbb{R}^{m_2})$ of probability measures.

2.2 Simple integer recourse

Simple recourse models are two-stage stochastic programs where deviations of a first-stage bid Tx against the random outcome $h(\omega)$ have to be compensated at certain costs in the second-stage. In simple integer recourse this compensation must be integer. In (2.6) we had already seen a second-stage value function Φ corresponding to a simple recourse model in case $h(\omega)$ maps to $\mathbb{R}^1$.

A two-stage stochastic program with simple integer recourse is given by

$$\min_x \{c^T x + \mathbb{E}_\mu[\Phi(\xi - Tx)] : x \in X\} \tag{2.22}$$

where

$$\Phi(t) := \min\{(q^+)^T y^+ + (q^-)^T y^- : y^+ \geq t,\ y^- \geq -t,\ y^+ \in \mathbb{Z}^s_+,\ y^- \in \mathbb{Z}^s_+\}. \tag{2.23}$$

It is easy to see that the optimization problem in (2.23) is feasible for any $t \in \mathbb{R}^s$, so the model has complete recourse. If, moreover, we assume that $q^+ \geq 0$ and $q^- \geq 0$, then the LP relaxation to the second-stage problem has a feasible dual, and $\Phi(t) \in \mathbb{R}$ for all $t \in \mathbb{R}^s$. Finally, the assumption $\mathbb{E}_\mu[\|\xi\|] < \infty$ will guarantee that the expectation in (2.22) is finite such that (2.22) becomes a well-defined optimization problem.

The crucial fact, allowing for a much richer analysis in simple integer recourse than in the general situation of (2.3)–(2.4), is the expressibility of the value function $\Phi(t)$ in closed form, namely

$$\Phi(t) = \sum_{i=1}^{s} \left(q_i^+ \lceil t_i \rceil^+ + q_i^- \lceil -t_i \rceil^+ \right) \tag{2.24}$$

where $\lceil \cdot \rceil^+ := \max\{\lceil \cdot \rceil, 0\}$. For the expected value function this implies the following separable representation

$$\begin{aligned} \phi(x) &= \mathbb{E}_\mu[\Phi(\xi - Tx)] \\ &= \sum_{i=1}^{s} q_i^+ \mathbb{E}_{\mu_i}\left[\lceil \xi_i - (Tx)_i \rceil^+\right] + \sum_{i=1}^{s} q_i^- \mathbb{E}_{\mu_i}\left[\lceil (Tx)_i - \xi_i \rceil^+\right] \end{aligned}$$

where μ_i is the probability measure corresponding to the marginal distribution of the i-th component of ξ. Studying $\phi(x)$ thus is studying the functions

$$u_i(\chi_i) := \mathbb{E}_{\mu_i}\left[\lceil \xi_i - \chi_i \rceil^+\right] \quad \text{and} \quad v_i(\chi_i) := \mathbb{E}_{\mu_i}\left[\lceil \chi_i - \xi_i \rceil^+\right] \tag{2.25}$$

reflecting expected surplus and expected shortage, respectively. Both these functions in one variable are quite similar, such that we will restrict further considerations to the expected surplus function. The following proposition reveals a close relationship with the cumulative distribution function $F_i(\tau) := \mu_i(\{\xi_i : \xi_i \leq \tau\})$ of ξ_i. For notational convenience, we drop the index i from now on.

Proposition 8. *In the above setting it holds for all* $\chi \in \mathbb{R}$

$$u(\chi) = \sum_{k=0}^{\infty} (1 - F(\chi + k)). \tag{2.26}$$

Proof. We have

$$\begin{aligned}\sum_{k=0}^{\infty}(1-F(\chi+k)) &= \sum_{k=0}^{\infty}\mu(\xi-\chi>k)\\ &= \sum_{k=0}^{\infty}\sum_{j=k+1}^{\infty}\mu\big(\lceil\xi-\chi\rceil^{+}=j\big)\\ &= \sum_{j=1}^{\infty}\sum_{k=0}^{j-1}\mu\big(\lceil\xi-\chi\rceil^{+}=j\big)\\ &= \sum_{j=1}^{\infty} j\cdot\mu\big(\lceil\xi-\chi\rceil^{+}=j\big)\\ &= \mathbb{E}_{\mu}\big[\lceil\xi-\chi\rceil^{+}\big]=u(\chi),\end{aligned}$$

and the proof is complete. □

Continuity and smoothness properties of u now result from those of the distribution function F or, existence provided, from those of a suitable probability density function of ξ. In particular, it can be shown that u is Lipschitz continuous on $\mathbb{R}$ if there exists a density of ξ with bounded variation. Moreover, u is differentiable on $\mathbb{R}$ if there exists a continuous density of ξ.

The function u is convex if ξ has a probability density function that is piecewise constant on every interval $]\alpha+j,\alpha+j+1[$, $j\in\mathbb{Z}$ for some $\alpha\in[0,1[$. Hence, when relating with arbitrary probability measures piecewise constant densities depending on the shift parameter α, one obtains convex functions related with the in general nonconvex function u. Given $\mu\in\mathcal{P}(\mathbb{R})$, its cumulative distribution function F, and $\alpha\in[0,1[$ we denote $\lfloor\tau\rfloor_{\alpha}:=\lfloor\tau-\alpha\rfloor+\alpha$ and relate the following probability density function with μ

$$\theta_{\alpha}(\tau):=F(\lfloor\tau\rfloor_{\alpha}+1)-F(\lfloor\tau\rfloor_{\alpha}),\quad \tau\in\mathbb{R}.$$

In the literature, the density θ_{α} and the corresponding measure μ_{α} as well as the resulting convex function related with u are called α-approximations. We will adopt this here although it is quite clear that the "distance" of μ and μ_{α}, and thus the approximation error in terms of the functions, in general cannot be made arbitrarily small.

Proposition 9. *For any* $\alpha \in [0, 1[$ *the* α*-approximation*

$$u_\alpha(\chi) := \mathbb{E}_{\mu_\alpha}\left[\lceil \xi - \chi \rceil^+\right]$$

of the expected surplus function u is a convex function on $\mathbb{R}$.

For the expected shortage function v an analogous statement is valid. Interestingly, the resulting convex α-approximation for

$$\tilde{\phi}(\chi) := q^+ u(\chi) + q^- v(\chi), \quad q^+ + q^- > 0,$$

then, up to an additive constant, arises as an expected value function of a simple recourse model with continuous variables where the underlying probability measure has been properly modified:

Proposition 10. *Let* $\alpha \in [0, 1[$, $q^+ + q^- > 0$, *and* u_α, v_α *denote the* α*-approximations for the expected surplus and shortage function u and v, respectively. Then it holds for all* $\chi \in \mathbb{R}$*:*

$$\begin{aligned}\tilde{\phi}_\alpha(\chi) &:= q^+ u_\alpha(\chi) + q^- v_\alpha(\chi) \\ &= q^+ \mathbb{E}_{\nu_\alpha}\left[\lceil \xi - \chi \rceil^+\right] + q^- \mathbb{E}_{\nu_\alpha}\left[\lceil \chi - \xi \rceil^+\right] + \frac{q^+ q^-}{q^+ + q^-}\end{aligned}$$

where ν_α *is a discrete probability measure such that for all* $k \in \mathbb{Z}$

$$\begin{aligned}\nu_\alpha(\{\alpha + k\}) := &\frac{q^+}{q^+ + q^-}(F(\alpha + k) - F(\alpha + k - 1)) \\ &+ \frac{q^-}{q^+ + q^-}(F(\alpha + k + 1) - F(\alpha + k)).\end{aligned}$$

The results outlined above are derived in detail in van der Vlerk (1995), see also the articles mentioned in the Bibliographical Notes below.

2.3 Multi-stage models

The models discussed so far assume a two-stage setting for the gain of information. Uncertainty is unveiled at once and decisions subdivide into those before and those after the unveiling. Often, a more complex view is appropriate at this place. Multistage stochastic programs address the situation

where uncertainty is unveiled stepwise with intermediate decisions that must not anticipate future information. We refer to Chapters 1 and 2 for basic statements about modeling principles and structure in the multi-stage situation. In what follows, we will adopt a fairly general modeling perspective, similar to Section 3.3 of Chapter 1. We will go beyond the setting of that section by adding integer requirements, and we will study some first implications of such a model extension.

Consider a finite horizon sequential decision process under uncertainty where the decision $x_t \in \mathbb{R}^{n_t}$ at stage $t \in \{1,\ldots,T\}$ is based on information available up to time t only. Information is modeled as a discrete time stochastic process $\{\xi_t\}_{t=1}^{T}$ on some probability space $(\Omega, \mathcal{F}, P)$ with ξ_t taking values in $\mathbb{R}^{m_t}$. The random vector $\xi^t := (\xi_1,\ldots,\xi_t)$ then reflects the information available up to time t. Nonanticipativity, i.e., the requirement that x_t must not depend on future information, is formalized by saying that x_t is measurable with respect to the σ-algebra $\mathcal{F}_t \subseteq \mathcal{F}$ which is generated by ξ^t, $t=1,\ldots,T$. Clearly, $\mathcal{F}_t \subseteq \mathcal{F}_{t+1}$ for all $t=1,\ldots,T-1$. As in the two-stage case, the first-stage decision x_1 usually is deterministic. Therefore, $\mathcal{F}_1 = \{\emptyset, \Omega\}$. Moreover, we assume that $\mathcal{F}_T = \mathcal{F}$.

The constraints of our multi-stage models are subdivided into three groups. The first group comprises conditions on x_t arising from the individual time stages:

$$\begin{aligned} &x_t(\omega) \in X_t, \quad B_t(\xi_t(\omega))x_t(\omega) \geq d_t(\xi_t(\omega)) \\ &P\text{-almost surely}, \quad t = 1,\ldots,T. \end{aligned} \tag{2.27}$$

Here, $X_t \subseteq \mathbb{R}^{n_t}$ is a set whose convex hull is a polyhedron. In this way, integer requirements to components of x_t are allowed for. For simplicity we assume that X_t is compact. Note that, by the integrality in X_t, we go beyond the setting of all the multi-stage models analyzed in Chapters 1 and 2. As in the two-stage case, convexity will no longer be available. In particular, the arguments based on duality and conjugacy that led to Proposition 30 in Chapter 2, stating that the multi-stage stochastic program is convex, are not working anymore, and, in fact, the models become nonconvex.

The next group of constraints models linkage between different time stages:

$$\sum_{\tau=1}^{t} A_{t\tau}(\xi_t(\omega))x_\tau(\omega) \geq g_t(\xi_t(\omega)) \quad P\text{-almost surely}, \; t = 2,\ldots,T. \tag{2.28}$$

Finally, there is the nonanticipativity of x_t, i.e.,

$$x_t \text{ is measurable with respect to } \mathcal{F}_t, \quad t = 1,\ldots,T. \tag{2.29}$$

In addition to the constraints we have a linear objective function

$$\sum_{t=1}^{T} c_t(\xi_t(\omega))x_t(\omega).$$

The matrices $A_{t\tau}(\cdot)$, $B_t(\cdot)$ as well as the right-hand sides $d_t(\cdot)$, $g_t(\cdot)$ and the cost coefficients $c_t(\cdot)$ all have conformable dimensions and depend affinely linearly on the relevant components of ξ.

The decisions x_t are understood as members of the function spaces $L_\infty(\Omega, \mathcal{F}, P; \mathbb{R}^{n_t})$, $t = 1,\ldots,T$. The constraints (2.27), (2.28) then impose pointwise conditions on the x_t, whereas (2.29) imposes functional constraints, in fact, membership in a linear subspace of $\times_{t=1}^{T} L_\infty(\Omega, \mathcal{F}, P; \mathbb{R}^{n_t})$.

The multistage extension of (2.3) is the minimization of expected minimal costs subject to nonanticipativity of decisions:

$$\min\left\{\int_\Omega \min_{x(\omega)}\left\{\sum_{t=1}^{T} c_t(\xi_t(\omega))x_t(\omega) \;:\; (2.27),\ (2.28)\right\} P(\mathrm{d}\omega) \colon x \text{ fulfilling } (2.29)\right\} \tag{2.30}$$

The minimization in the integrand of (2.30) being separable with respect to $\omega \in \Omega$, it is possible to interchange integration and minimization, and the problem can be restated as

$$\min\left\{\int_\Omega \sum_{t=1}^{T} c_t(\xi_t(\omega))x_t(\omega)P(\mathrm{d}\omega) \colon x \text{ fulfilling } (2.27),\ (2.28),\ (2.29)\right\}. \tag{2.31}$$

Due to the mentioned interplay of pointwise and functional constraints it remains to check whether (2.31) is well defined, cf. Section 3.3 of Chapter 1 where this is addressed for a noninteger counterpart of (2.31).

Recall that X_t is compact and assume that $\xi_t \in L_1(\Omega, \mathcal{F}, P; \mathbb{R}^{m_t})$ for $t = 1,\ldots,T$. For each $\omega \in \Omega$ we define the subset $\mathcal{Y}(\omega)$ of $\mathcal{X} := \times_{t=1}^{T} \mathbb{R}^{n_t}$

$$\begin{aligned}\mathcal{Y}(\omega) := \Bigg\{ & y \in \mathcal{X} \colon y_t \in X_t,\ B_t(\xi_t(\omega))y_t \geq d_t(\xi_t(\omega)),\ t = 1,\ldots,T, \\ & \sum_{\tau=1}^{t} A_{t\tau}(\xi_t(\omega))y_\tau \geq g_t(\xi_t(\omega)),\ t = 2,\ldots,T\Bigg\}\end{aligned} \tag{2.32}$$

and the extended real-valued function φ

$$\varphi(y_1, \ldots, y_T, \omega) := \begin{cases} \sum_{t=1}^{T} c_t(\xi_t(\omega))y_t, & (y_1, \ldots, y_T) \in \mathcal{Y}(\omega), \\ +\infty, & \text{otherwise} \end{cases} \tag{2.33}$$

from $\mathcal{X} \times \Omega$ to $(-\infty, +\infty]$. With these notations, (2.31) is equivalent to

$$\min\{\mathbb{E}_P[\varphi(x_1, \ldots, x_T, \omega)]: x_t \text{ measurable w.r.t. } \mathcal{F}_t,\ t = 1, \ldots, T\}. \tag{2.34}$$

The real-valued function $(y, \omega) \mapsto \sum_{t=1}^{T} c_t(\xi_t(\omega))y_t$ is continuous in y for each $\omega \in \Omega$ and measurable in ω for each $y \in \mathcal{X}$, and the set-valued mapping $\mathcal{Y}$ from Ω to $\mathcal{X}$ is closed-valued and measurable (cf. Theorem 14.36 in Rockafellar and Wets, 1997). With $\mathcal{B}(\mathcal{X})$ denoting the σ-algebra of Borel sets in $\mathcal{X}$, the function φ is $\mathcal{B}(\mathcal{X}) \otimes \mathcal{F}$-measurable (cf. Example 14.32 in Rockafellar and Wets, 1997). Furthermore, the following estimate is valid for each $y \in \times_{t=1}^{T} X_t$ and $\omega \in \Omega$:

$$|\varphi(y_1, \ldots, y_T, \omega)| \le \sum_{t=1}^{T} \|c_t(\xi_t(\omega))\| \sup_{y_t \in X_t} \|y_t\| \tag{2.35}$$

Hence, $\mathbb{E}_P[\varphi(x_1, \ldots, x_T, \omega)]$ is finite for each decision $x = (x_1, \ldots, x_T)$ such that $x(\omega) \in \mathcal{Y}(\omega)$ for P-almost all $\omega \in \Omega$.

As in Evstigneev (1976), we construct recursively two sequences of functions by putting $\psi_{T+1} := \varphi$ and

$$\varphi_t(y_1, \ldots, y_t, \omega) := \mathbb{E}_P^r[\psi_{t+1}(y_1, \ldots, y_t, \cdot) \mid \mathcal{F}_t](\omega), \tag{2.36}$$

$$\psi_t(y_1, \ldots, y_{t-1}, \omega) := \inf_y \varphi_t(y_1, \ldots, y_{t-1}, y, \omega), \tag{2.37}$$

for $t = T, \ldots, 1$, and for each $\omega \in \Omega$ and $y_\tau \in X_\tau$, $\tau = 1, \ldots, T$.

Here, $\mathbb{E}_P^r[\cdot \mid \mathcal{F}_t]$ denotes the regular conditional expectation with respect to $\mathcal{F}_t$. By definition, the regular conditional expectation is a version of the conditional expectation (i.e., $\mathbb{E}_P^r[\cdot \mid \mathcal{F}_t] = \mathbb{E}_P[\cdot \mid \mathcal{F}_t]$, P-a.s.) with the property that the mapping $(z, \omega) \mapsto \Phi(z, \omega) := \mathbb{E}_P^r[\Psi(z, \cdot) \mid \mathcal{F}_t](\omega)$ from $Z_t \times \Omega$ to $(\infty, +\infty]$ is $\mathcal{B}(Z_t) \otimes \mathcal{F}_t$-measurable if Ψ is $\mathcal{B}(Z_t) \otimes \mathcal{F}$-measurable. Here, Z_t is allowed to be an arbitrary closed subset of a Euclidean space. The regular conditional expectation exists if Ψ is $\mathcal{B}(Z_t) \otimes \mathcal{F}$-measurable and uniformly integrable, i.e., there exists a (real) random variable ζ with finite first moment such that $|\Psi(z, \omega)| \le \zeta(\omega)$ for $z \in Z_t$ and $\omega \in \Omega$ (see Dynkin and Evstigneev, 1976).

Due to condition (2.35), relation (2.36) is well defined for $t = T$ and leads to a $\mathcal{B}(Z) \otimes \mathcal{F}_T$-measurable function ϕ_T, where $Z := \times_{t=1}^{T} X_t$. It is shown in Evstigneev (1976) that the relations (2.36) and (2.37) are well defined for all $t = T, \ldots, 1$. Furthermore, the following optimality criterion and existence result for (2.34) or, equivalently, for (2.31) is valid.

Proposition 11. *Adopt the above setting for* (2.31) *and assume that* (2.31) *has a feasible solution. Then* $\{\overline{x}_t\}_{t=1}^{T}$ *is an optimal solution of* (2.31) *iff*

$$\varphi_t(\overline{x}^t(\omega), \omega) = \psi_t(\overline{x}^{t-1}(\omega), \omega), \quad P - \text{a.s.}, \, t = 1, \ldots, T. \tag{2.38}$$

Moreover, there exists a solution $\overline{x}_1$ *of the first-stage optimization problem*

$$\min\{\varphi_1(x_1) = \mathbb{E}_P[\psi_2(x_1, \omega)] \colon x_1 \in X_1, \, B_1(\xi_1)x_1 \geq d_1(\xi_1)\}, \tag{2.39}$$

and given $\mathcal{F}_\tau$*-measurable functions* $\overline{x}_\tau$ *for* $\tau = 1, \ldots, t-1$, *there exists an* $\mathcal{F}_\tau$*-measurable function* $\overline{x}_t$ *such that* $\varphi_t(\overline{x}^t(\omega), \omega) = \psi_t(\overline{x}^{t-1}(\omega), \omega)$, *P-a.s.*

Relations (2.36) and (2.37) define the mixed-integer analogon to the nested formulation developed in Section 3.1 of Chapter 1 for the purely linear case, see also the general nested problem in continuous variables in Section 3.3 of Chapter 1. In particular, the optimal value $\psi_t(y_1, \ldots, y_{t-1}, \omega)$ is the cost-to-go, cf. (3.3) in Chapter 2, and (2.38) states the fact that an optimal solution to (2.31) has to fulfil the dynamic programming equation, cf. (3.4) and Proposition 30 in Chapter 2. As in (3.5) of Chapter 1, in problem (2.39) all the subsequent stages are absorbed into the function $\varphi_1(x_1)$. Hence, (2.39) is a well-defined mixed-integer extension of (3.5) in Chapter 1, with the only difference that (3.5) has staircase whereas (2.39) triangular form.

In Proposition 30 of Chapter 2 convexity of the noninteger counterpart to $\varphi_1(\cdot)$ is shown. With integer requirements, this convexity already breaks down for two-stage problems and hence cannot be expected to hold. However, lower semicontinuity of $\varphi_1(\cdot)$ still can be established.

Proposition 12. *Adopt the above setting for* (2.31) *and assume that the matrices* $A_{t\tau}(\cdot)$, $B_t(\cdot)$ *as well as the cost coefficients* $c_t(\cdot)$ *all are deterministic. Then the objective function* $\varphi_1(\cdot)$ *of the first-stage optimization problem* (2.39) *is lower semicontinuous on its domain of definition.*

Proof. (Outline) The proof is done by induction over the time stages $t = T, T-1, \ldots, 1$. According to the definitions in (2.33) and (2.37), at each stage an infimum of a parameter dependent function over a mixed-integer set constrained by linear inequalities is taken. Moreover, the latter set has parameters in the right-hand sides of the inequalities. At each stage, it can be

shown that the set-valued mapping assigning to the relevant right-hand side parameter the relevant mixed-integer constraint set is upper semicontinuous.

At stage T, the objective function of the above parametric optimization problem is linear and does not depend on a parameter. In fact, we have a mixed-integer linear program with right-hand side parameters, whose value function is lower semicontinuous according to the argument given in item (3) in front of Proposition 2. A conditional expectation of this value function, cf. (2.36), which, by Fatou's Lemma for conditional expectations, is again lower semicontinuous, enters the objective at stage $T-1$, cf. (2.37).

Therefore, at stage $T-1$, as well as in all subsequent stages $t = T-2,\ldots,2$, we have a parametric program where, with respect to the relevant parameter, the objective function is lower semicontinuous and the constraint set mapping is upper semicontinuous. Basic results from parametric optimization, such as Theorem 1.17 of Rockafellar and Wets (1997) then imply that the value function of the optimization problem at stage $T-1$ is lower semicontinuous with respect to the relevant parameters. Fatou's Lemma for conditional expectations then inherits lower semicontinuity to the objective in stage $T-2$, and the above arguments can be repeated. □

Structural properties for problem (2.31) beyond the above propositions are widely open and a field of current research.

3 Algorithms

3.1 Decomposition schemes

Consider a classical deterministic integer program

$$\text{(IP)} \quad \min\{c^T x : x \in X\}. \tag{3.40}$$

For simplicity, we consider the pure integer case, where $X \subseteq \mathbb{Z}^n$, although most of the definitions and properties presented here extend to the mixed integer case. In practice, the set X is described through a finite set of linear constraints, defining a polyhedron $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ and through integrality restrictions.

A polyhedron $P \subseteq \mathbb{R}^n$ is a formulation for $X \subseteq \mathbb{Z}^n$ iff $X = P \cap \mathbb{Z}^n$. The ideal formulation would be to replace IP by the equivalent linear program

$$\min\{c^T x : x \in \text{conv}(X)\} \tag{3.41}$$

as all extreme points of the convex hull of X belong to X. Such a formulation naturally gives an integer solution to the linear programming relaxation of (3.41).

An inequality $\pi^T x \leq \pi_0$ is a valid inequality for $X \subseteq \mathbb{R}^n$ if $\pi^T x \leq \pi_0$ for all $x \in X$. The Chvatal–Gomory procedure to construct a valid inequality for $X = P \cap \mathbb{Z}^n$ is to consider a vector $u \in \mathbb{R}^m_+$ of nonnegative weights for the rows of A, take the linear combination of the rows

$$u^T A x \leq u^T b,$$

round down the coefficients in the l.h.s,

$$\sum_{j=1}^{n} \lfloor u^T a_j \rfloor x_j \leq u^T b$$

where a_j is the jth column of A, then round-down the r.h.s.

$$\sum_{j=1}^{n} \lfloor u^T a_j \rfloor x_j \leq \lfloor u^T b \rfloor. \tag{3.42}$$

The first transformation is justified by $u \geq 0$, the second by $\sum_{j=1}^{n} \lfloor u^T a_j \rfloor x_j \leq \sum_{j=1}^{n} u^T a_j x_j$, and the third as $\sum_{j=1}^{n} \lfloor u^T a_j \rfloor x_j$ is integer. Thus, for all $u \geq 0$, (3.42) is a valid inequality for X. More surprisingly, every valid inequality for X can be obtained by applying the Chvatal–Gomory procedure a finite number of times.

The separation problem SP associated with IP is the following : given $\overline{x} \in \mathbb{R}^n$, is $\overline{x} \in \text{conv}(X)$? If not, find a valid inequality $\pi^T x \leq \pi_0$ for X such that $\pi^T \overline{x} > \pi_0$. As the Chvatal–Gomory procedure is not a constructive one, a considerable amount of research has been devoted to characterize strong valid inequalities, which define facets of P, and related separation algorithms (SA). Among the popular SA, we may cite lifted cover inequalities for knapsack constraints, mixed integer rounding inequalities (a generalization of Gomory mixed integer cuts) and disjunctive inequalities. Separation algorithms are not always efficient, as sometimes the separation problem itself is NP-hard. Even when the separation algorithm is efficient, it may not be wise to generate all possible valid inequalities as the linear program would tend to become huge. Early attempt to use Gomory cuts appeared to be disappointing as successive cuts tend to be less and less efficient and the corresponding LP harder and harder to solve. Cutting planes are thus very often combined with a branch & bound scheme. Branching on a fractional value $\overline{x}_j$ consists of defining two subregions $X \cap \{x : x_j \leq \lfloor \overline{x}_j \rfloor\}$ and $X \cap \{x : x_j \geq \lfloor \overline{x}_j \rfloor + 1\}$. These subregions can in turn be subdivided by later branchings. This generates a number of nodes N^ρ, $\rho = 1, \ldots, R$, that form a partition of $\mathbb{R}^n$, i.e., $\mathbb{R}^n = U_{\rho=1,\ldots,R} N^\rho$ and $N^\rho \cap N^\sigma = \emptyset$, $\rho \neq \sigma$.

Clearly,

$$\min\{c^T x\colon x \in X\} = \min_{\rho=1,\ldots,R}\{\min\{c^T x\colon x \in X \cap N^\rho\}\}$$

The branch & bound procedure is finite as only finitely many nodes can be generated when X is compact. Nodes are fathomed for three reasons: when they have integer solution, when they have an optimal value, which is worse than an already known integer solution, and when they are infeasible. As already said, modern codes combine cut generation and branching in so called branch & cut procedures. Another basic technique in integer programming is Lagrangian relaxation. Its main idea is to relax "complicating" constraints to end up with manageable subproblems that are coordinated by a nonsmooth dual optimization. In general, integer requirements imply the lack of duality such that, typically, heuristics have to be employed for finding promising primal solutions on the basis of the results of the dual optimzation. We now show how stochastic programming decomposition techniques could be designed using cutting planes and branch & bound in the second-stage as well as Lagrangian relaxation of nonanticipativity constraints.

Decomposition by cutting planes

Let the two stage stochastic integer problem be

$$\text{(SIP)}\quad \min\{c^T x + \mathbb{E}_\mu[\Phi(\xi - Tx)]\colon x \in X\} \tag{3.43}$$

with

$$\Phi(t) := \min\{q^T y\colon Wy = t,\ y \in Y\} \tag{3.44}$$

and $Y \subseteq \mathbb{Z}_+^{n_2}$. All notations are as in Section 2. The deterministic equivalent program is

$$\text{(DEP)}\quad \min\{c^T x + Q(x)\colon x \in X\} \tag{3.45}$$

where $Q(x) = \mathbb{E}_\mu[\Phi(\xi - Tx)]$. For all practical purposes, it can be transformed into

$$\min\{c^T x + \theta\colon x \in X,\ \theta \geq Q(x)\} \tag{3.46}$$

In view of the properties presented just above, the difficulty of solving (3.46) precisely lies in having a formulation to replace $\theta \geq Q(x)$. Any valid constraint in the $\mathbb{R}^{n_1+1}$ space of (x, θ) is called an optimality cut. It takes the form of a constraint $\theta \geq f(x)$ that holds for every $x \in X$, and $\theta \geq Q(x)$.

Definition 13. A set of s optimality cuts $\{\theta \geq f_\ell(x), \ell = 1,\ldots,s\}$ is said to be sufficient at $x \in X$ if

$$\theta \in \{\theta: \theta \geq f_\ell(x), \ell = 1 \ \cdots \ s\} \Rightarrow \theta \geq Q(x).$$

In the classical L-shaped algorithm where the second-stage only involves continuous variables and ξ has a finite support, the linear programming duality theory implies that a sufficient set of optimality cuts can be found for all $x \in X$.

We now show how optimality cuts could be generated through a cutting plane solution of the second-stage program. For simplicity of presentation, we assume complete recourse for the second-stage so that feasibility cuts can be left aside. Take the case when the assumptions of Proposition 4 are satisfied. We also assume $Y \subseteq \mathbb{Z}^{n_2}$. For each outcome $\xi^k \in \Xi$, the second-stage problem is

$$\Phi(\xi^k - Tx) = \min\{q^T y: Wy \geq \xi^k - Tx, y \in \mathbb{Z}^{n_2}\} \tag{3.47}$$

Let $\mathbb{F}$ be the set of dual price functions. Although more general classes can be considered, take $\mathbb{F}$ to be the class of functions $F: \mathbb{R}^{m_2} \to \mathbb{R}$ that are nondecreasing, subadditive and satisfy $F(0) = 0$. The dual problem of (3.47) is

$$\max_F\{F(\xi^k - Tx): F(w_j) = q_j, j = 1,\ldots,n_2, F \in \mathbb{F}\} \tag{3.48}$$

where w_j is the jth column of W.

Proposition 14. *Suppose* $\hat{F}^k$, $k = 1,\ldots,K$, *are optimal dual price functions obtained by solving* (3.48) *with* $x = x^\nu$ *for each* $\xi^k \in \Xi$. *Then, an optimality cut at* x^ν *is given by*

$$\theta \geq \sum_{k=1}^{K} p^k \hat{F}^k(\xi^k - Tx) \tag{3.49}$$

Proof. For each $k = 1,\ldots,K$, let $\overline{y}^k$ and $\hat{F}^k$ be optimal solutions of (3.47) and (3.48), respectively, corresponding to $x = x^\nu$. By integer duality, they satisfy $q^T\overline{y}^k = \hat{F}^k(\xi^k - Tx^\nu)$. For each feasible $F \in \mathbb{F}$ and each feasible $y \in \mathbb{Z}^{n_2}$, we have $F(\xi^k - Tx^\nu) \leq \hat{F}^k(\xi^k - Tx^\nu) = q^T\overline{y}^k \leq q^T y$. Now, for each $x \in X$ and corresponding optimal dual price $\overline{F}^k$, we have

$$Q(x) = \sum_{k=1}^{K} p^k \overline{F}^k(\xi^k - Tx) \geq \sum_{k=1}^{K} p^k \hat{F}^k(\xi^k - Tx),$$

with

$$Q(x^v) = \sum_{k=1}^{K} p^k \hat{F}^k(\xi^k - Tx^v). \qquad \square$$

In a cutting plane procedure, the LP-relaxation of the integer program is considered. Valid inequalities are successively generated and added to the formulation, until the LP-relaxation optimal solution spontaneously meets the integrality restrictions. The valid inequalities can be written as

$$\sum_{j=1}^{n_2} F^{(\ell)}(w_j) y_j \geq F^{(\ell)}(q), \quad \ell = 1, \ldots, s, \tag{3.50}$$

where $F^{(\ell)} \in \mathbb{F}$, $\ell = 1, \ldots, s$. Let $(u_1, \ldots, u_{m_2}, u_{m_2+1}, \ldots, u_{m_2+s})$ be the dual variables associated to the optimum of the final LP-relaxation. We construct the function $F: \mathbb{R}^{m_2} \to \mathbb{R}$ as

$$F(t) := \sum_{i=1}^{m_2} u_i t_i + \sum_{i=1}^{s} u_{m_2+i} F^{(i)} t \tag{3.51}$$

By construction, $F \in \mathbb{F}$ and is an optimal solution to (3.48).

As an example, consider the case where valid inequalities are generated by the Gomory's fractional cutting plane algorithm.

Let $h = \xi^k - Tx^v$ in (3.47).

Let

$$y_{B_i} + \sum_{j \in NB} \overline{w}_{ij} y_j = \overline{h}_i \quad \text{for } i = 1, \ldots, m_2 \tag{3.52}$$

be the optimal basis for second-stage problem (3.47) associated with a given k and given x^v. This optimal basis can be rewritten in such a way that $\overline{w}_{ij} \leq 0$, $j \in \text{NB}$, $\overline{h}_i \geq 0$, $i = 1, \ldots, m_2$. NB is the set of nonbasic variables, while B_i is the variable basic in row i. Let y^* be the corresponding optimal solution. If y^* does not meet the integrality requirements, there is at least one row i with $\overline{h}_i \notin \mathbb{Z}^1$. Choosing such a row, the Chvatal–Gomory cut for row i is

$$y_{B_i} + \sum_{j \in \text{NB}} \lfloor \overline{w}_{ij} \rfloor y_j \leq \lfloor \overline{h}_i \rfloor$$

which, by eliminating y_{B_i}, can be rewritten as

$$\sum_{j \in \mathrm{NB}} (\overline{w}_{ij} - \lfloor \overline{w}_{ij} \rfloor) y_j \geq \overline{h}_i - \lfloor \overline{h}_i \rfloor$$

or

$$\sum_{j \in \mathrm{NB}} f_{ij} y_j \geq f_i \tag{3.53}$$

where $f_{ij} = \overline{w}_{ij} - \lfloor \overline{w}_{ij} \rfloor$, $f_i = \overline{h}_i - \lfloor \overline{h}_i \rfloor$. By construction, $0 \leq f_{ij} < 1$ and $0 < f_i < 1$. As $y_i^* = 0$, $i \in \mathrm{NB}$ at the optimal LP solution, this inequality cuts off y^*.

Now let β be the row in the basis inverse corresponding to the row which generated the cut (3.53). Let $g = \beta - \lfloor \beta \rfloor$. Then, the Gomory cut (3.53) can be rewritten in terms of the original variables as the Chvatal–Gomory inequality

$$\sum_{j=1}^{n} \lfloor g w_j \rfloor y_j \leq \lfloor g h \rfloor,$$

namely

$$\sum_{j=1}^{n} \lfloor g w_j \rfloor y_j \leq \lfloor g(\xi^k - T x^\nu) \rfloor. \tag{3.54}$$

Thus, expressed in terms of the first-stage variable, the optimality cuts involve $\lfloor g(\xi^k - Tx) \rfloor$ terms, i.e., rounding-down of combinations of x. Each one of these rounding down operations requires an additional constraint and an additional integer variable in the first-stage. The left-hand-side of (3.54) does not depend on x, so that similar cuts could be generated for various realizations of $\xi \in \Xi$. At the moment, it is not known how to use this property efficiently. The approach remains impracticable as the number of auxiliary variables and constraints will equal the total number of round down operations needed to generate the cuts.

Decomposition by branch & bound

We now indicate how optimality cuts can be obtained through a branch & bound solution procedure in the second-stage. Consider a given first-stage iterate point x^ν and a given realization ξ^k of ξ. Based on a full branching on the second-stage problem (3.47) for $x = x^\nu$, one obtains a partition of $\mathbb{R}^{n_2}$ into R terminal nodes $Y^\rho = \{y : a^\rho \leq y \leq b^\rho\}$, $\rho = 1, \ldots, R$. The optimal objective value of the second-stage program over Y^ρ is

$$Q^\rho(x^\nu, \xi^k) = \min\{q^T y : Wy = \xi^k - Tx^\nu,\ a^\rho \leq y \leq b^\rho\}$$

By linear programming duality, it is also

$$Q^\rho(x^\nu, \xi^k) = (\pi^\rho)^T(\xi^k - Tx^\nu) + (\underline{\pi}^\rho)^T a^\rho + (\overline{\pi}^\rho)^T b^\rho$$

where π^ρ, $\underline{\pi}^\rho$ and $\overline{\pi}^\rho$ are the optimal dual variables associated with the original constraints, lower and upper bounds on $y \in Y^\rho$, respectively.

To simplify notations, we represent this expression as

$$Q^\rho(x^\nu, \xi^k) = (\sigma_k^\rho)^T x^\nu + \tau_k^\rho$$

with $(\sigma_k^\rho)^T = -(\pi^\rho)^T \cdot T$ and $\tau_k^\rho = (\pi^\rho)^T \xi^k + (\underline{\pi}^\rho)^T a^\rho + (\overline{\pi}^\rho)^T b^\rho$.

Duality theory also implies that $Q^\rho(x, \xi^k) \geq (\sigma_k^\rho)^T x^\nu + \tau_k^\rho$.

Also, by construction of the branch & bound,

$$Q(x, \xi^k) = \min_{\rho=1,\ldots,R} Q^\rho(x, \xi^k).$$

Thus,

$$\theta^k \geq p_k \min_{\rho=1,\ldots,R} \{(\sigma_k^\rho)^T x + \tau_k^\rho\} \tag{3.55}$$

is a valid optimality cut for $Q(x, \xi^k)$. It can thus be embedded in a multicut representation $\theta = \sum_{k=1,\ldots,K} \theta^k$. When SIP has not complete recourse, some of the terminal nodes may be infeasible, in which case their dual solutions contain unbounded rays with dual objective values going to $-\infty$, so that the minimum is restricted to the feasible terminal nodes.

The optimality cut (3.55) is a sufficient set of optimality cuts at x^ν for θ^k. Unfortunately, as it is well known, (3.55) is a nonlinear expression. R auxiliary binary variables and $R+1$ constraints are required to describe (3.55) in a polyhedral representation with mixed integer variables. This, plus solving a full branch & bound for each x^ν and each k, makes a decomposition by branch & bound of little practical value.

Scenario decomposition

In a general setting, scenario decomposition can be understood as a solution method for the multi-stage stochastic integer program (2.31) where, out of the constraints (2.27), (2.28), (2.29), the nonanticipativity condition (2.29) is subjected to Lagrangian relaxation. We will demonstrate scenario decomposition at the following two-stage model:

$$\min\{c^T x + \mathbb{E}_\mu[\Phi(\xi - Tx)] : x \in X\} \tag{3.56}$$

with

$$\Phi(t) := \min\{q^T y \colon Wy = t,\ y \in Y\}$$

where $X \subseteq \mathbb{R}^{n_1}$, $Y \subseteq \mathbb{R}^{n_2}$ are polyhedra, possibly involving integer requirements to components of x and y. We assume that all problem data have conformable dimensions, that W is rational, and that μ follows a discrete distribution with realizations (or scenarios) $\xi_1, \ldots, \xi_N$ and probabilities $\pi_1, \ldots, \pi_N$. Then problem (3.56) can be written as the following mixed-integer linear program

$$\min_{x, y_j} \left\{ c^T x + \sum_{j=1}^{N} \pi_j q^T y_j \colon Tx + Wy_j = \xi_j,\ y_j \in Y,\ x \in X \right\}. \tag{3.57}$$

Due to the sheer size, general purpose mixed-integer linear programming solvers quickly fail at these problems. We reformulate (3.57) by introducing copies x_j, $j = 1, \ldots, N$, and adding the explicit nonanticipativity constraints $x_1 = \cdots = x_N$, or an equivalent system. For notational convenience, the latter is written as $\sum_{j=1}^{N} H_j x_j = 0$ with proper (l, n_1)-matrices H_j, $j = 1, \ldots, N$. Problem (3.57) then becomes

$$\min \left\{ \sum_{j=1}^{N} \pi_j (c^T x_j + q^T y_j) \colon Tx_j + Wy_j = \xi_j,\ x_j \in X,\ y_j \in Y,\ \sum_{j=1}^{N} H_j x_j = 0 \right\}. \tag{3.58}$$

For $\lambda \in \mathbb{R}^l$ we consider the functions

$$L_j(x_j, y_j, \lambda) := \pi_j (c^T x_j + q^T y_j) + \pi_j \lambda^T H_j x_j, \quad j = 1, \ldots, N, \tag{3.59}$$

and form the Lagrangian

$$L(x, y, \lambda) := \sum_{j=1}^{N} L_j(x_j, y_j, \lambda).$$

Later on, the Lagrangian will have to be minimized over a mixed-integer polyhedral set, which will be accomplished by mixed-integer linear programming solvers. To avoid nonlinearities, our Lagrangian is very much standard in that it is based on linear expressions and does not involve nonlinear augmentation terms.

Given that the objective in (3.58) is an expectation, it is natural to base the Lagrangian at a probabilistic inner product which, in (3.59), leads to the factors π_j in front of the terms $\lambda^T H_j x_j$. Conceptually, this follows the lines of dualization as developed in Rockafellar and Wets (1978). Furthermore, ill-conditioning in the Lagrangian dual when disregarding the probabilities in the second term of (3.59) is avoided this way, see Bacaud et al. (2001) and Gröwe-Kuska et al. (2002) for respective observations.

The Lagrangian dual of (3.58) is the optimization problem

$$\max\{D(\lambda)\colon \lambda \in \mathbb{R}^l\} \tag{3.60}$$

where

$$D(\lambda) = \min\left\{\sum_{j=1}^{N} L_j(x_j, y_j, \lambda)\colon Tx_j + Wy_j = \xi_j,\ x_j \in X,\ y_j \in Y\right\}. \tag{3.61}$$

The above minimization is separable, and we have

$$D(\lambda) = \sum_{j=1}^{N} D_j(\lambda) \tag{3.62}$$

where

$$D_j(\lambda) = \min\{L_j(x_j, y_j, \lambda)\colon Tx_j + Wy_j = \xi_j,\ x_j \in X,\ y_j \in Y\}. \tag{3.63}$$

$D_j(\lambda)$ is the pointwise minimum of affine functions in λ, and hence piecewise affine and concave. Therefore, (3.60) is a nonsmooth concave maximization (or convex minimization) problem that can be solved by bundle methods from nondifferentiable optimization, see Hiriart-Urruty and Lemaréchal (1993) and Kiwiel (1990). At each iteration, these methods require the objective value and one subgradient of D. These are obtained by solving the optimization problem in (3.61) which, thanks to the separability in (3.62), reduces to solving N problems of single-scenario size. The latter are mixed-integer linear programs and very often within the reach of advanced general purpose solvers. Altogether, the optimal value φ_{LD} of (3.60) provides a lower bound to the optimal value φ of problem (3.57). Specifying a well-known result in Lagrangian relaxation of mixed-integer linear programs, see e.g., Nemhauser and Wolsey (1988), leads to the following proposition.

Proposition 15. *It holds* $\varphi \geq \varphi_{\mathrm{LD}}$. *If for some multiplier* $\lambda \in \mathbb{R}^l$ *the optimal solutions* (x_j, y_j), $j = 1, \ldots, N$, *to the optimization problem in* (3.61) *fulfil* $\sum_{j=1}^{N} H_j x_j = 0$, *then* $\varphi = \varphi_{\mathrm{LD}}$ *and* (x_j, y_j), $j = 1, \ldots, N$, *are optimal for* (3.58).

With φ_{LP} denoting the optimal value of the linear programming relaxation to (3.58) it holds $\varphi_{\mathrm{LD}} \geq \varphi_{\mathrm{LP}}$.

Equality of φ and φ_{LD} in Proposition 15 being a rare exception, "promising" feasible points for the original primal problem (3.58) are derived by heuristics using the results of the dual optimization. Since, in our situation, the relaxed constraints ($x_1 = \cdots = x_N$) are particularly simple, ideas for such heuristics arise quite naturally. For example, examine the x_j-components, $j = 1, \ldots, N$, of solutions to (3.63) for optimal or nearly optimal λ, and decide for the most frequent value arising or average and round if necessary. If the heuristic provides a feasible solution to (3.58), then the objective value of the latter yields an upper bound $\overline{\varphi}$ for φ.

The difference $\overline{\varphi} - \varphi_{\mathrm{LD}}$ then indicates the quality of the feasible solution found. If desired, this quality certificate can be improved by embedding the procedure described so far into a branch & bound scheme for (3.56) seen as a nonconvex global optimization problem. Recall from (2.11) the notation $\phi(x) := \mathbb{E}_\mu[\Phi(\xi - Tx)]$. Let $\mathbf{P}$ denote the list of current problems and $\varphi_{\mathrm{LD}} = \varphi_{\mathrm{LD}}(P)$ the Lagrangian lower bound for $P \in \mathbf{P}$. The scheme then consists of the following steps.

Scenario decomposition algorithm

Step 1 *Initialization:* Set $\overline{\varphi} = +\infty$ and let $\mathbf{P}$ consist of problem (3.58).

Step 2 *Termination:* If $\mathbf{P} = \emptyset$ then the solution $\hat{x}$ that yielded $\overline{\varphi} = c^T\hat{x} + \phi(\hat{x})$ is optimal.

Step 3 *Node selection:* Select and delete a problem P from $\mathbf{P}$ and solve its Lagrangian dual. If the optimal value $\varphi_{\mathrm{LD}}(P)$ hereof equals $+\infty$ (infeasibility of a subproblem) then go to step 2.

Step 4 *Bounding:* If $\varphi_{\mathrm{LD}}(P) \geq \overline{\varphi}$ go to step 2 (this step can be carried out as soon as the value of the Lagrangian dual rises above $\overline{\varphi}$).

(i) The scenario solutions x_j, $j = 1, \ldots, N$, are identical: if $c^T x_j + \phi(x_j) < \overline{\varphi}$ then let $\overline{\varphi} = c^T x_j + \phi(x_j)$ and delete from $\mathbf{P}$ all problems P' with $\varphi_{\mathrm{LD}}(P') \geq \overline{\varphi}$. Go to step 2.

(ii) The scenario solutions x_j, $j = 1, \ldots, N$ differ: compute the average $\overline{x} = \sum_{j=1}^{N} \pi_j x_j$ and round it by some heuristic to obtain $\overline{x}^R$. If $c^T\overline{x}^R + \phi(\overline{x}^R) < \overline{\varphi}$ then let $\overline{\varphi} = c^T\overline{x}^R + \phi(\overline{x}^R)$ and delete from $\mathbf{P}$ all problems P' with $\varphi_{\mathrm{LD}}(P') \geq \overline{\varphi}$. Go to step 5.

Step 5 *Branching:* Select a component $x_{(k)}$ of x and add two new problems to $\mathbf{P}$ obtained from P by adding the constraints $x_{(k)} \leq \lfloor \overline{x}_{(k)} \rfloor$ and $x_{(k)} \geq \lfloor \overline{x}_{(k)} \rfloor + 1$, respectively (if $x_{(k)}$ is an integer component), or $x_{(k)} \leq \overline{x}_{(k)} - \varepsilon$ and $x_{(k)} \geq \overline{x}_{(k)} + \varepsilon$, respectively, where $\varepsilon > 0$ is a tolerance parameter to have disjoint subdomains.

This scheme is obviously finite if X is bounded and all x-components have to be integers. If x is mixed-integer some stopping criterion to avoid endless branching on the continuous components has to be employed.

As neither of the classical approaches for IP can provide a comprehensible decomposition for SIP, research has been devoted to a number of cases which receive a nice treatment. Subsequently, some of these cases will be addressed.

3.2 Simple integer recourse

A two-stage stochastic program with simple integer recourse was defined in Section 2.2 as

$$\min_x\{c^T x + \mathbb{E}_\mu[\Phi(\xi - Tx)]\colon x \in X\}$$

where

$$\Phi(t) = \min\{(q^+)^T y^+ + (q^-)^T y^-\colon y^+ \geq t,\ y^- \geq -t,\ y^+ \in \mathbb{Z}_+^s,\ y^- \in \mathbb{Z}_+^s\}.$$

We again use the notation

$$u_i(\chi_i) := \mathbb{E}_{\mu_i}\left[\lceil \xi_i - \chi_i \rceil^+\right] \quad \text{and} \quad v_i(\chi_i) := \mathbb{E}_{\mu_i}\left[\lfloor \chi_i - \xi_i \rfloor^+\right]$$

to represent the expected surplus and the expected shortage, respectively. Letting

$$\Psi_i(\chi_i) = q_i^+ u_i(\chi_i) + q_i^- v_i(\chi_i) \tag{3.64}$$

the two-stage stochastic program with simple integer recourse can be transformed into

$$\min_x \left\{ c^T x + \sum_{i=1}^s \Psi_i(\chi_i)\colon \chi = Tx,\ x \in X \right\}. \tag{3.65}$$

To make the presentation simpler, we now consider the case with expected shortage only. It is defined as

$$\min_x \left\{ c^T x + \sum_{i=1}^s q_i^+ u_i(\chi_i)\colon \chi = Tx,\ x \in X \right\}. \tag{3.66}$$

Note that all results available for the expected shortage easily translate to the expected surplus, and therefore also to the functions $\Psi_i(\chi_i)$, $i = 1, \ldots, s$. For notational convenience, we will drop the index i whenever it is not required.

Proposition 16. *$u(\chi+1)-u(\chi)$ is a nondecreasing function of χ.*

Proof. From Proposition 8, we have

$$u(\chi) = \sum_{k=0}^{\infty} (1 - F(\chi + k)).$$

It follows that, for all $n \in \mathbb{Z}_+$, we have

$$u(\chi + n) = u(\chi) - \sum_{k=0}^{n-1} (1 - F(\chi + k)). \tag{3.67}$$

Taking $n = 1$, we get

$$u(\chi + 1) - u(\chi) = F(\chi) - 1.$$

The proposition holds as $F(\cdot)$ is a cumulative distribution function. □

If we consider χ values which are integer apart, we may draw a piecewise linear function through successive points $(\chi \pm k, u(\chi \pm k))$, k integer. This piecewise linear function is convex by Proposition 16. It may sometimes be the convex hull of $u(\chi)$. A sufficient condition for that is that the support of ξ is a subset of $\mathbb{Z}$. But in general, it is not. Take the simple example where $\xi = 1/2$ or $3/2$ with probability $1/2$ each, and observe that $u(1/4) = 3/2 > \frac{1}{2}(u(0) + u(1/2))$, as $u(0) = 3/2$ and $u(1/2) = 1/2$. In any case, as this piecewise linear function is convex, we can derive valid inequalities in the $(\chi, u(\chi))$ space which are supporting half-lines of this function.

Proposition 17. *Let $\eta \in \mathbb{Z}$. Define $\pi = u(\eta)-u(\eta+1)$ and $e = (\eta+1)u(\eta)-\eta u(\eta+1)$. Then, for all $\chi \in \mathbb{Z}$,*

$$u(\chi) \geq e - \pi\chi. \tag{3.68}$$

Moreover

$$u(\eta) = e - \pi \cdot \eta. \tag{3.69}$$

Proof. Consider the case where $\chi \geq \eta$. Then

$$u(\chi) - u(\eta) = \sum_{k=0}^{\chi-\eta-1} [u(\chi - k) - u(\chi - k - 1)].$$

By Proposition 16, each term in the sum is bounded below by $u(\eta+1)-u(\eta)$. Hence $u(\chi)-u(\eta) \geq (\chi-\eta)(u(\eta+1)-u(\eta))$.

The inequality (3.68) follows. The case where $\chi \leq \eta$ is similar. Finally, (3.69) is obtained by straightforward computation of $e-\pi\eta$. □

We now propose an exact algorithm for the case where χ is integer.

Algorithm SPSIR

Step 1 *Initialization:* $\nu := 0$, $r_i := 0$, $i = 1,\ldots,s$.

Step 2 *Current problem:* Let $\nu := \nu + 1$. Solve the program

$$\min_x \left\{ c^T x + \sum_{i=1}^{s} q_i^+ \theta_i : \chi = Tx,\ x \in X,\ \pi_{\ell(i)}\chi_i + \theta_i \geq e_{\ell(i)},\ \ell(i) = 1,\ldots,r_i \right\}. \tag{3.70}$$

Let $(x^\nu, \theta_1^\nu, \ldots, \theta_s^\nu)$ be an optimal solution to (3.70). If $r_i = 0$ for some i, θ_i^ν is set to $-\infty$ and is not considered in the computation of x^ν.

Step 3 *Termination:* Let $\chi^\nu = Tx^\nu$. If $\theta_i^\nu = u_i(\chi_i^\nu)$ for $i = 1,\ldots,s$, then x^ν is an optimal solution. Stop.

Step 4 *Cut generation:* For any i, $i = 1,\ldots,s$, such that $\theta_i^\nu < u_i(\chi_i^\nu)$, compute

$$\pi_{r_i+1} = u_i(\chi_i^\nu) - u_i(\chi_i^\nu + 1)$$

and

$$e_{r_i+1} = (\chi_i^\nu + 1)u_i(\chi_i^\nu) - \chi_i^\nu u_i(\chi_i^\nu + 1).$$

Set $r_i := r_i + 1$. Go to step 2.

Proposition 18. *Assume X is bounded. Also assume $\chi \in \mathbb{Z}^s$ for all $x \in X$. Then, the SPSIR algorithm finds an optimal solution to (3.66) in a finite number of steps, provided $u_i(\chi_i)$ can be obtained through a finite computation, $i = 1,\ldots,s$.*

Proof. X being bounded, there are only finitely many different values of χ_i, $i = 1,\ldots,s$. For each χ_i, only one cut can be generated. Thus, r_i is finite. By (3.69), the same cut can only be generated once in step 4. □

Example 19. *Newsboy Problem Revisited.* Assume now that the newsboy can purchase a number of publications (newspaper, magazines, ...). Each publication i, $i = 1,\ldots,s$, has a purchase cost c_i, a selling cost s_i and a demand which is a random variable ξ_i with cumulative distribution function F_i. For simplicity, we assume unsold publications cannot be returned to the vendor.

Each publication has a weight (or volume) a_i. A total weight (or volume) b is available to the newsboy.

The problem is naturally an integer program as publications sell by the unit. It reads as follows

$$\min\left\{\sum_{i=1}^{s} c_i x_i + \mathbb{E}_\mu\left[-\sum_{i=1}^{s} s_i y_i(\xi_i)\right] : \sum_{i=1}^{s} a_i x_i \le b, \right.$$
$$\left. 0 \le y_i(\xi_i) \le \xi_i,\ y_i(\xi_i) \le x_i,\ x \in \mathbb{Z}_+^s,\ y(\xi) \in \mathbb{Z}_+^s \right\}.$$

Letting $y_i^+(\xi_i) = \xi_i - y_i(\xi_i)$, one obtains an equivalent formulation

$$\min\left\{\sum_{i=1}^{s} c_i x_i + \mathbb{E}_\mu\left[\sum_{i=1}^{s} s_i y_i^+(\xi_i)\right] - \mathbb{E}_\mu\left[\sum_{i=1}^{s} \xi_i\right] : \right.$$
$$\left. \sum_{i=1}^{s} a_i x_i \le b,\ y_i^+(\xi_i) \ge \xi_i - x_i,\ x \in \mathbb{Z}_+^s,\ y^+(\xi) \in \mathbb{Z}_+^s \right\}.$$

Omitting the constant term and using the notation $u_i(x_i)$ previously defined for the expected surplus, the newsboy problem becomes

$$\min\left\{\sum_{i=1}^{s} c_i x_i + \sum_{i=1}^{s} s_i u_i(x_i) : \sum_{i=1}^{s} a_i x_i \le b,\ x \in \mathbb{Z}_+^s \right\}.$$

This problem obviously satisfies the assumptions in Proposition 18: X is bounded through the upper limit on the total weight and $\chi = x$ is integer by definition.

It now remains to study a number of cases where the computation of $u(\chi)$ is finite.

(i) ξ has a finite range.
This case includes random variables with finite support, but also continuous random variables on a finite range. For instance, if ξ has a uniform density on $[0, a]$, then for $0 \le x \le a$,

$$u(x) = \sum_{k=0}^{\lceil a-x \rceil - 1} [1 - F(x+k)].$$

(ii) Closed form expressions can be found.

Let ξ have a negative exponential density with parameter $\lambda > 0$. Then, for $x \geq 0$,

$$u(x) = \sum_{k=0}^{\infty} (1 - F(x+k)) = \sum_{k=0}^{\infty} e^{-\lambda(x+k)}$$
$$= \frac{e^{-\lambda x}}{1 - e^{-\lambda}}$$

(iii) The support of ξ is a subset of $\mathbb{Z}$.
By (3.67), we have

$$u(\chi + n) = u(\chi) - \sum_{k=0}^{n-1} (1 - F(\chi + k)).$$

Observe that $F(t) = F(\lfloor t \rfloor)$ for all $t \in \mathbb{R}$, as the support of $\xi \in \mathbb{Z}$. Thus, $u(t) = u(\lfloor t \rfloor)$ for all $t \in \mathbb{R}$. Consider $x \geq 0$. Apply (3.67) with $\chi = 0$ and $n = \lfloor x \rfloor$. It follows that

$$u(x) = u(0) - n + \sum_{k=0}^{n-1} F(k).$$

Now, as the support of ξ is a subset of $\mathbb{Z}$, $u(0) = \mathbb{E}[\lceil \xi \rceil^+] = \mathbb{E}[\xi^+] = \mathbb{E}[\max(\xi, 0)]$. In particular, for such distributions with $\xi \geq 0$, $u(0) = \mathbb{E}[\xi]$. Such is the case for a Poisson distribution, for instance.

Example 20. *Newsboy Problem Continued.* Take the newsboy problem with $s = 2$, $c^T = (1, 2)$, $s_1 = 3$, $s_2 = 7$, $a^T = (2, 3)$, $b = 12$. Assume the demand for both publications follows a Poisson distribution with parameter 3. Assume a starting point of $x^T = (0, 0)$. The initial cuts are found as follows.

By definition, $u(0) = 3$. Compute $u(1) = u(0) + F(0) - 1 = 2.0498$. Then $\pi_1 = 0.9502$ and $e_1 = 3$.

The next iterate will be $x^T = (0, 4)$, with one extra cut generated, then $x^T = (3, 2)$ with two new cuts, then $x^T = (3, 2)$ again, which is optimal. Observe that the mean value optimum is $x^T = (1, 3)$.

Note, finally, that even for such continuous densities as the normal distribution, it is possible to compute $u(x)$ within a fixed tolerance in a finite number of steps (see Theorem 3.30 in Birge and Louveaux, 1997). This tolerance can be chosen equal to the machine tolerance so that the computation can be considered exact.

3.3 Binary first-stage variables

When the first-stage variables are binary, it is possible to obtain a sufficient set of optimality cuts at each $x \in X \cap \{0,1\}^{n_1}$. A finite algorithm, called the integer L-shaped, has been designed. It can be made efficient for hard problems when lower bounding functionals are available.

Assumption 21. There exists a finite lower bound L satisfying

$$L \leq \min_x \{Q(x) : x \in X\}.$$

Assumption 22. For $x \in X$, $Q(x)$ is computable in a finite number of steps.

At a given stage of the algorithm, we consider the current problem (CP)

$$(\mathrm{CP})_\rho \quad \min_{x,\theta} \{c^T x + \theta : x \in P \cap N^\rho, \theta \geq L, \theta \geq f_\ell(x), \ell = 1, \ldots, s\}$$

where P is a formulation of X, N^ρ is node ρ of the first-stage branching scheme, $\mathbb{R}^{n_1} = \cup_{\rho=1,\ldots,R} N^\rho$. L is a finite lower bound as in Assumption 21 and $\theta \geq f_\rho(x), \ell = 1,\ldots, s$, are the optimality cuts. We now present a general scheme.

Integer L-shaped algorithm

Step 1 *Initialization:* Let $s = 0$, $\nu = 0$, $\rho = 1$, $\overline{z} = \infty$, $\theta := L$. A list is created that contains one single node corresponding to the initial problem, i.e., $N^1 := \mathbb{R}^{n_1}$.

Step 2 *Selection:* Select one node ρ in the list, if none exists, stop.

Step 3 *Solution:* Set $\nu := \nu + 1$. Solve $(\mathrm{CP})_\rho$. If it is infeasible, fathom node N^ρ and go to step 2. Otherwise, let (x^ν, θ^ν) be an optimal solution. If $c^T x^\nu + \theta^\nu \geq \overline{z}$, fathom node N^ρ and go to step 2.

Step 4 *Branch & cut:* Check for integrality restrictions. If some restriction is violated, apply a separation algorithm to find a valid inequality. If some is found, adapt P, then return to step 3. If not, create two new nodes following the usual branching. Append the nodes to the list, fathom node N^ρ and go to step 2.

Step 5 *Second-stage value*. Compute $Q(x^\nu)$ and $z^\nu = c^T x^\nu + Q(x^\nu)$. If $z^\nu < \overline{z}$, update $\overline{z} := z^\nu$.

Step 6 *Optimality cuts:* If $\theta^\nu \geq Q(x^\nu)$, fathom node N^ρ and go to step 2. Otherwise, find some optimality cut, set $s := s + 1$ and go to step 3.

The integer L-shaped method yields an optimal solution (when one exists) in a finite number of steps when a sufficient set of optimality cuts exists for each $x \in X$. As $X \subseteq \{0,1\}^{n_1}$, step 6 can only be performed a finite number

of times. All other steps of the method are finite, as the branch & cut procedure is finite.

When needed, the integer L-shaped can be used in a multicut version. Then, θ is replaced by $\sum_{k=1,\dots,K} \theta^k$, and the restriction $\theta \geq Q(x)$ becomes $\theta^k \geq Q(x, \xi^k)$, $k = 1,\dots, K$. In this case, step 6 becomes

Step 6 *Optimality cuts in the multicut version:* If $(\theta^k)^\nu \geq Q(x^\nu, \xi^k)$, for all $k = 1,\dots, K$, fathom node N^ρ and go to step 2. Otherwise, for each k such that $(\theta^k)^\nu < Q(x^\nu, \xi^k)$, find some optimality cut, adapt s_k and go to step 3.

Based on the fact that the first-stage variables are binary, we easily obtain a sufficient set of optimality cuts.

Proposition 23. *Let $x_i = 1$, $i \in S$, and $x_i = 0$, $i \notin S$, be some first-stage feasible solution. Let $q_S = Q(x)$ be the corresponding recourse function value. Define the optimality cut as*

$$\theta \geq (q_S - L)\left(\sum_{i \in S} x_i - \sum_{i \notin S} x_i\right) - (q_S - L)(|S| - 1) + L. \tag{3.71}$$

Then, the optimality cut (3.71) is sufficient at x.

Proof. Let $\delta(x, S) = \sum_{i \in S} x_i - \sum_{i \notin S} x_i$. We have $\delta(S) \leq |S|$. $\delta(S) = |S|$ only if $x_i = 1$, $i \in S$, and $x_i = 0$, $i \notin S$. In that case, the right-hand side of (3.71) takes the value q_S and $\theta \geq q_S$ is valid as q_S is precisely $Q(x)$. In all other cases, $\delta(S) \leq |S| - 1$. Then the right hand side of (3.71) takes a value smaller than or equal to L and $\theta \geq L$ is valid by Assumption 21. This single cut is sufficient at x since $Q(x) = q_S$. □

An alternative is to consider the integer L-shaped as a particular case of the cutting plane decomposition of Section 3.1.1, with

$$F(t) = \begin{cases} \phi(t) & \text{for } t \geq \xi - Tx^\nu, \\ L & \text{otherwise.} \end{cases}$$

To avoid sending too many optimality cuts, it is helpful to add a number of lower bounding functionals on $Q(x)$. One general possibility is to add cuts from the continuous L-shaped. Let

$$\Gamma(x, \xi) = \min\{q^T y \colon Wy = \xi - Tx,\ y \in \mathbb{R}^{n_2}_+\}$$

and $\Gamma(x) = \mathbb{E}_\mu[\Gamma(x, \xi)]$. Then,

$$\theta \geq \Gamma(x^\nu) + \partial\Gamma(x^\nu)^T(x - x^\nu) \tag{3.72}$$

is a valid lower bounding functional. We now present two different situations where the optimality cuts can be improved and the integer L-shaped becomes more efficient.

s-Neighbors

When more information is available on $Q(x)$, improvements on (3.71) can be obtained. Let

$$N(s, S) = \{x: \delta(x, S) = |S| - s, x \in X\}$$

be the s-neighbors of S. Assume we can find

$$\lambda(s, S) \leq \min\{Q(x): x \in N(s, S)\}, s = 0, \ldots, |S|,$$

or at least a series of $\lambda(s, S)$, $s \leq t$. Observe that $\lambda(0, S) = q_S$.

Proposition 24. *Let $x_i = 1$, $i \in S$, $x_i = 0$, $i \notin S$, be a feasible solution to SIP, with $q_S = Q(x)$. Define $a = \max\{q_S - \lambda(1, S), (q_S - L)/2\}$. Then*

$$\theta \geq a\delta(x, S) + q_S - a|S| \tag{3.73}$$

is a sufficient optimality cut at x.

Proof. For $x \in N(s, S)$, the right-hand side of (3.73) is equal to $q_S - as$. We show that this value is a valid lower bound on $Q(x)$. This is obvious for $s = 0$. When $s = 1$, the r.h.s. is $q_S - a$. By definition of a, $q_S - a$ is bounded above by $q_S - (q_S - \lambda(1, S)) = \lambda(1, S)$, which is by definition a lower bound on one neighbors of S. When $s = 2$, $q - 2a \leq q_S - 2(q_S - L)/2 = L$. Finally, for $s \geq 3$, $q_S - as \leq q_S - 2a \leq L$. (3.73) is sufficient at x as $\theta \geq q_S$ for $s = 0$. □

Geometrically, (3.73) defines a half-space in the (δ, θ) space, above a line passing through the two points $(|S|, q_S)$ and $(|S|-1, \lambda(1, S))$ when $a = q_S - \lambda(1, S)$, or the two points $(|S|, q_S)$ and $(|S|-2, L)$ where $a = (q_S - L)/2$.

Proposition 25. *Let $x_i = 1$, $i \in S$, $x_i = 0$, $i \notin S$, be a feasible solution, with $q_S = Q(x)$. Let $1 \leq t \leq |S|$ be some integer. Then (3.73) holds with*

$$a = \max\{\max_{s \leq t}(q_S - \lambda(s, S)/s;\ (q_S - L)/(t + 1)\}. \tag{3.74}$$

Proof. If $x \in N(s, S)$, the right-hand-side of (3.73) is $q_S - as$. By (3.74), for all $s \le t$, $q_S - as \le q_S - (q_S - \lambda(s, S)) = \lambda(s, S)$ which is a lower bound on $Q(x)$ by definition. For $s > t$, $q_S - as \le q_S - a(t+1) \le q_S - (q_S - L) = L$ which is also valid. □

Proposition 26. *Assume* $q_S > \lambda(1, S)$. *If* $\lambda(s-1, S) - \lambda(s, S)$ *is nonincreasing in* s *for* $1 \le s \le \lfloor (q_S - L)/(q_S - \lambda(1, S)) \rfloor$, *then* (3.73) *holds with* $a = q_S - \lambda(1, S)$.

Proof. It suffices to show that in (3.74), the maximum in the right-hand side is obtained when $s = 1$. Let $t = \lfloor (q_S - L)/q_S - \lambda(1, S) \rfloor$. For $s \le t$, we have $q_S - \lambda(s, S) = \sum_{i=1}^{s} (\lambda(i-1, S) - \lambda(i, S))$. By assumption, each term of the sum is smaller than the first term of the sum, so the total is less than s times $q_S - \lambda(1, S)$. By definition of t, we have $t + 1 \ge (q_S - L)/(q_S - \lambda(1, S))$ or $q_S - \lambda(1, S) \ge (q_S - L)/(t+1)$. □

Example 27. Let $Q(x) = \sum_{j=1}^{m_2} Q_j(x)$ with

$$Q_j(x, \xi) = \min \left\{ r_j y_j : d_j y_j \ge d_j - \sum_{i \in T(j)} \xi_{ij} x_i,\ y_j \in \{0, 1\} \right\}$$

and

$$Q_j(x) = \mathbb{E}_\mu [Q_j(x, \xi)].$$

Assume $x_i \in \{0, 1\}$, $i \in T$, with $T = \cup_{j=1,\dots,m_2} T(j)$ and $T(j) \cap T(k) = \phi$, $j \ne k$. This can be seen as a number of investments x_i, $i \in T$, which are made in a first stage. They have a random yield ξ_{ij} in product j. Any deficiency in attaining the target d_j for product j results in a penalty r_j. The second-stage value is then simply

$$Q(x) = \sum_{j=1}^{m_2} r_j P\left(\sum_{i \in S(j)} \xi_{ij} < d_j \right) \tag{3.75}$$

where

$$S(j) = \{ i \in T(j) : x_i = 1 \}, \quad j = 1, \dots, m_2, \text{ at the current solution } x.$$

To apply Proposition 26, take $S = \cup_{j=1,\dots,m_2} S(j)$. By definition, $q_S = Q(x)$ and is easily computed from (3.75) when the sum of the random variables ξ_{ij}, $i \in S(j)$, has a known distribution, $j = 1, \dots, m_2$. Such is case for standard distributions as the Poisson or normal distribution. There are two ways 1-neighbors can be obtained. First, for one j, $1 \le j \le m_2$, one x_i, $i \in S(j)$, goes

from one to zero (and all other x_i's are unchanged). In that case, $Q(x)$ is increased and any value smaller than or equal to q_S is a valid lower bound. Second, we may have one single x_i, $i \notin S$, going from zero to one, and again all other x_i's unchanged. As the $T(j)$, $j=1,\ldots,m_2$, form a partition, one single term is modified in (3.75). We easily obtain a lower bound by simply assuming this term vanishes. Thus

$$q_S - \lambda(1, S) \leq \max_{j=1,\ldots,m_2} \left\{ r_j \, P\left(\sum_{i \in S(j)} \xi_{ij} < d_j \right) \right\}$$

Without loss of generality, order the j's in decreasing order of $r_j \, P(\sum_{i \in S(j)} \xi_{ij} < d_j)$. Repeating the argument above, we get

$$q_S - \lambda(s, S) \leq \sum_{j=1}^{s} r_j \, P\left(\sum_{i \in S(j)} \xi_{ij} < d_j \right)$$

and Proposition 26 applies.

Vehicle routing problems

The integer L-shaped method can be used to solve hard problems for specific applications. We now illustrate this on an example of a routing problem. The capacitated vehicle routing problem, is defined on an undirected graph $G=(V, E)$ where $V=\{v_1,\ldots,v_n\}$ is a vertex set and $E=\{(v_i, v_j)$: $v_i, v_j \in V, i<j\}$ is an edge set. Vertex v_1 is a depot at which are based m identical vehicles of capacity D, while the remaining vertices are customers. A symmetric travel cost matrix $C=(c_{ij})$ is defined on E. With each customer v_i is associated a nonnegative demand to be collected or delivered, but not both. Without loss of generality, we consider the first case here. In the classical VRP, each v_i has a known demand d_i. The problem then consists of designing m vehicle routes: (i) each starting and ending at the depot, (ii) such that every customer is visited only once by one vehicle, (iii) the total demand of any route does not exceed D, and (iv) the total routing cost is minimized. The VRP is known to be NP-hard.

In the stochastic case, each customer v_i has a stochastic demand ξ_i. The consequence of having stochastic demands is that a planned vehicle route may fail at a given customer location whenever the accumulated demand exceeds D. In such a case, a failure is said to occur and a recourse action must be implemented. The stochastic VRP can easily be represented as

$$\min\{cx + Q(x): x \in X\}$$

where X defines the usual restrictions on the routes (degree constraints at the nodes, no subtour, and expected demand of any route does not exceed the vehicle capacity). $x=(x_{ij})$ is defined as the arc variables, with $x_{ij}=1$ if (v_i, v_j) belongs to a route and $x_{ij}=0$ otherwise, and $Q(x)$ is the expected cost of recourse actions in case of failure. This expected cost is separable in the routes and must be computed for each of its two orientations

$$Q(x) = \sum_{k=1}^{m} \min\{Q^{k,1}, Q^{k,2}\}$$

where $Q^{k,\delta}$ denotes the expected cost of recourse corresponding to route k and orientation $\delta=1$ or 2. For a given route k defined by $V_k=(v_{i_1}=v_1, v_{i_2}, \ldots, v_{i_t+1}=v_1)$, one orientation corresponds to following the route with customers visited in the natural order. The other orientation corresponds to visiting them in backward order. Observe that orientation does matter in a stochastic setting, while it does not in a deterministic one. Assume the recourse action simply consists of the following steps: return to the depot, unload the vehicle then resume the route at the customer where failure occurs. This recourse action is called a return trip to the depot. Then, the expected recourse cost for the first orientation is

$$Q^{k,1} = 2\sum_{j=2}^{t}\sum_{\ell=1}^{\infty} P\left(\sum_{s=2}^{j-1}\xi_{i_s} \le \ell D \le \sum_{s=2}^{j}\xi_{i_s}\right) c_{1i_j}. \tag{3.76}$$

It can be rewritten as

$$Q^{k,1} = 2\sum_{j=2}^{t}\sum_{\ell=1}^{\infty} [F^{j-1}(\ell D) - F^{j}(\ell D)] c_{1i_j} \tag{3.77}$$

where $F^j(\ell D) = P(\sum_{s=2}^{j}\xi_{i_s} \le \ell D)$. If $\xi_i \le D$ a.s. for all i, then the upper limit in the second summation in (3.76) and (3.77) can be brought down to $j-1$.

(i) Lower bounding functionals on $Q(x)$ can be obtained at fractional first-stage solutions. They are based on the concept of "partial routes". A partial route h is specified by two ordered vertex sets $S_h=\{v_1, \ldots, v_{s_h}\}$ and $T_h=\{v_1, \ldots, v_{t_h}\}$ satisfying $S_h \cap T_h = \{v_1\}$, and a third set U_h satisfying $S_h \cap U_h = \{v_{s_h}\}$ and $T_h \cap U_h = \{v_{t_h}\}$.

For simplicity, we write $(v_i, v_j) \in S_h$ or T_h if v_i and v_j are consecutive in S_h or T_h. The partial route h induced by these sets is made up of the two chains $(v_1, \ldots, v_{s_h})$, $(v_1, \ldots, v_{t_h})$ and of the unstructured set U_h. Let $R_h = S_h \cup T_h \cup U_h$. Define

$$W_h(x) = \sum_{(v_i, v_j)\in R_h} x_{ij} - |R_h| + 1.$$

Let P be a lower bound on any solution containing the partial routes and, as usual, L a lower bound on $Q(x)$.

Proposition 28. *The constraint*

$$\theta \geq L + (P - L)\left(\sum_{h=1}^{r} W_h(x) - r + 1\right) \tag{3.78}$$

is a valid inequality for SVRP.

Proof. By construction $W_h(x) \leq 1$. Hence, $\sum_{h=1}^{r} W_h(x) - r + 1 \leq 1$. It is only 1 when $W_h(x) = 1$ for all h. In this case, (3.78) becomes $\theta \geq P$. Otherwise (3.78) is redundant. □

A greedy heuristic to find out partial routes proves to be an efficient separation algorithm to detect violated inequalities (3.78). We construct $P = \sum_{h=1}^{r+1} P_h$ as follows. For $h \leq r$, we create an artificial customer v_a with demand $\xi_a = \sum_{v_i \in U_h \setminus \{v_{s_h}, v_{t_h}\}} \xi_i$ and $c_{1a} = \min_{v_i \in U_h \setminus \{v_{s_h}, v_{t_h}\}} \{c_{1i}\}$. Then construct route k equal to $\{v_1, \ldots, v_{s_h}, v_a, v_{t_h}, \ldots, v_1\}$ and compute $P_h = \min\{Q^{k,1}, Q^{k,2}\}$ as before for this artificial route. P_{r+1} is a lower bound on the expected recourse restricted to the customer set $V \setminus U_{h=1}^{r} R_h$ and $m-r$ vehicles. It is similar to the computation of L for $Q(x)$, that is now described in the next paragraph.

(ii) Lower bound on $Q(x)$. Relabel all customers in nondecreasing order of their distance to the depot. Denote by X_k the random demand on route k, $k = 1, \ldots, m$, and let $F_k(\cdot)$ be its distribution function. Let X_T be the total random demand and $F_T(\cdot)$ its distribution function.

Proposition 29. *Let $\pi(F_k, D)$ be a lower bound on the probability of having at least one failure on a route whose demand is defined by F_k. A valid lower bound on $Q(x)$ is given by*

$$\begin{aligned} L = \inf_{(F_1, \ldots, F_m)} \Bigg\{ & 2 \sum_{k=1}^{m} c_{1,k+1} \pi(F_k, D) : \\ & \int \cdots \int_{x_1 + \cdots + x_m \leq x} \mathrm{d}F_1(x_1) \cdots \mathrm{d}F_m(x_m) = F_T(x) \quad \text{for all } x, \\ & F_k \in \mathbb{F}_k, \, k = 1, \ldots, m \Bigg\} \end{aligned} \tag{3.79}$$

where $\mathbb{F}_k$ is a family of distribution functions to be specified. In particular, it must be such that $\mathbb{E}_\mu[X_k] \leq D$.

Proof. From (3.77), the expected cost $Q^{k,\delta}$ of route k with orientation δ is obtained by computing the cost of having the ℓth failure at the jth customer, then summing up over all ℓ and j. Each of these terms contribute to $Q^{k,\delta}$ by a nonnegative amount. A valid lower bound is obtained by only considering the first failure, $\ell = 1$. By definition of $\pi(F_k, D)$, we obtain $Q(x) \geq 2 \sum_{k=1}^{m} \gamma_k \cdot \pi(F_k, D)$ for any lower bound γ_k on the distance of a customer in route k to the depot. We may then replace $\gamma_1, \ldots, \gamma_m$ by the m least distances to the depot, to obtain the objective function in (3.79). □

Apply this proposition to the case where $\mathbb{F}_k$ is the set of normal distributions. Let $x_k = \mathbb{E}[X_k]$ and $y_k = \mathrm{Var}[X_k]$. Let also $\mu_T = \mathbb{E}[X_T]$ and $\sigma_T^2 = \mathrm{Var}[X_T]$. Let $x = (x_1, \ldots, x_m)$, $y = (y_1, \ldots, y_m)$.

Then (3.79) reduces to

$$L = \min_{x,y} \left\{ 2 \sum_{k=1}^{m} c_{1,k+1} \left[1 - G\left(\frac{D - x_k}{\sqrt{y_k}} \right) \right] : \sum_{k=1}^{m} x_k = \mu_T, \right.$$
$$\left. \sum_{k=1}^{m} y_k = \sigma_T^2, \, y \geq 0, \, x \geq 0 \right\}. \tag{3.80}$$

where $G(t) = P(Z \leq t)$, $Z \sim N(0, 1)$.

As the objective function is neither convex nor concave, we obtain a more workable problem by replacing y_k by some lower bound y_0. We know that for each route $x_k \geq x_0 := \max\{\mu_T - (m-1)D; \, \min_{i=2,\ldots,n} \mathbb{E}[\xi_i]\}$. Then, $y_0 = \min_Z \{\sum_{i=2}^{n} \mathrm{Var}[\xi_i] z_i : \sum_{i=2}^{n} \mathbb{E}[\xi_i] z_i \geq x_0, \, z_i \in \{0, 1\}\}$ is a lower bound on y_k. We may thus replace (3.80) by

$$L = \min_{x} \left\{ 2 \sum_{k=1}^{m} c_{1,k+1} \left[1 - G\left(\frac{D - x_k}{\sqrt{y_0}} \right) \right] : \sum_{k=1}^{m} x_k = \mu_T, \, x \geq 0 \right\}. \tag{3.81}$$

The objective function in (3.81) is convex in x_k for $x_k \leq D$, so that (3.81) can be solved by applying the Karush–Kuhn–Tucker conditions. Let γ be the multiplier on $\sum_{k=1}^{m} x_k = \mu_T$. Define $b = -\gamma \sqrt{\pi y_0 / 2}$. Then

$$x_k = \begin{cases} D & \text{if } c_{1,k+1} \leq b \\ D - \sqrt{2 y_0 \ln(c_{1,k+1}/b)} & \text{if } c_{1,k+1} \geq b. \end{cases}$$

As x_k, $k = 1, \ldots, m$, is a nondecreasing function of b, they can be determined recursively up to the moment where $\sum_{k=1}^{m} x_k = \mu_T$ within some tolerance. Similar results apply for other cases, such as the Poisson distribution.

The combination of lower bounding functionals on partial routes and the calculation of a lower bound L based on (3.81) for the normal case prove to be very efficient.

3.4 Second-stage integer variables

We now consider the case where $Y \subseteq \mathbb{Z}_+^{m_2}$. Assume again T and W are fixed. Suppose also that the second-stage program for a given ξ can be represented as

$$\Psi(\chi, \xi(\omega)) = \min\{q^T y \colon Wy \geq \xi - \chi,\ y \in Y\}$$

where $Y \subseteq \mathbb{Z}_+^{m_2}$. Assume also that Ξ has a finite support.

Proposition 30. *Let W be integer. Assume $\{u \in \mathbb{R}^{m_2} \colon W^T u \leq q\} \neq \emptyset$. For every $\xi \in \Xi$ and $j = 1, \ldots, m_2$, $\Psi(\chi_j, \xi)$ is lower semicontinuous and nondecreasing in χ_j. Moreover, for any $h \in \mathbb{Z}$, $\Psi(\chi_j, \xi)$ is constant over $\chi_j \in (h - \xi_j - 1, h - \xi_j]$, $\xi \in \Xi$, $j = 1, \ldots, m_2$.*

Proof. The first part of the proposition comes from Proposition 2. Any solution $y \in \arg\min\{q^T y \colon Wy \geq \xi - \chi',\ y \in Y\}$ belongs to $\{y \colon Wy \geq \xi - \chi,\ y \in Y\}$ when $\chi' \leq \chi$. So $\Psi(\chi, \xi(w)) \leq \Psi(\chi', \xi(w))$. Define $h = \lceil \xi_j - \chi_j' \rceil$. With integral W, the jth constraint $(Wy)_j \geq \xi_j - \chi_j$ implies $(Wy)_j \geq h_j$. Hence $\Psi(\chi_j, \xi(\omega))$ is constant over $(h - \xi_j - 1, h - \xi_j]$. □

It is thus possible to partition the χ space in an orthogonal complex $\mathcal{C}$, where each cell $C \in \mathcal{C}$ is of the form $\prod_{j=1,\ldots,m_2} (a_j, b_j]$. On each of these cells, $\Psi(\chi)$ is constant. The closure of a nonempty cell forms a full-dimensional hyper-rectangle. The nonempty cells form a partition of $\mathbb{R}^{m_2}$. A branch & bound algorithm can be constructed, where the branching consists of considering other cells of this orthogonal complex than those already considered.

Define, as usual,

$$\Psi(\chi) = \mathbb{E}_\mu[\Psi(\chi, \xi(\omega))]$$

Define the current problem associated to a set S as

$$\mathrm{CP}(S) \quad z = \inf_{x, \chi, \theta} \{cx + \theta \colon x \in X,\ \chi \in S,\ Tx = \chi,\ \theta \geq f_\ell(x),\ \ell = 1, \ldots, s\}$$

where, as before, $\theta \geq f_\ell(x)$ are optimality cuts or lower bounding functionals. Typically, (3.72) can be used here.

Algorithm for second-stage integer variables

Step 1 *Initialization:* Set $\nu = 0$, $\overline{z} = \infty$. Choose s and find s valid inequalities $\theta \geq f_\ell(x)$, $\ell = 1, \ldots, s$. A list is created that contains the single node $N^1 := \chi$. Set $\rho = 1$.

Step 2 *Selection:* Select a node N^ρ in the list; if none exists, stop.

Step 3 *Current solution:* Set $\nu := \nu + 1$. Solve CP(N^ρ). If N^ρ is infeasible, fathom node N^ρ, go to step 2. Otherwise, let $(x^\nu, \chi^\nu, \theta^\nu)$ be an optimal solution. If $cx^\nu + \theta^\nu \geq \overline{z}$, fathom the current node and go to step 2.

Step 4 *Cell:* Find the cell $C = \prod_{j=1,\ldots,m_2} (a_j, b_j]$ s.t. $\chi^\nu \in C$. Let $b = (b_1, \ldots, b_{m_2})$. Compute $\Psi(b, \xi)$ for all $\xi \in \Xi$ and let $\Psi^\nu = \mathbb{E}_\mu[\Psi(b, \xi)]$.

Step 5 *Solution value:* Solve

$$v = \inf_{x,\chi} \{c^T x : x \in X,\ \chi \in C,\ Tx = \chi\}.$$

Let $\overline{x}$, $\overline{\chi}$ be an optimal solution. Compute $z^\nu = c\overline{x} + \Psi^\nu$. If $\overline{z} < z^\nu$, update $\overline{z} := z^\nu$.

Step 6 *Branch:* Select a component j, $1 \leq j \leq m_2$. Create two new nodes $N^\rho \cap \{\chi : \chi_j \leq a_j\}$ and $N^\rho \cap \{\chi : \chi_j > b_j\}$. Fathom node N^ρ and go to step 2.

The algorithm is finite as the number of cells of the complex is finite since X is compact. Additional valid inequalities can be found along the way in step 3. In practice, open sets and inf would be replaced by closed sets and min, through a well-chosen tolerance on open sets. In the following case, the definition of the cells is easier. When the discrete random variable is such that all realizations of component ξ_j, $j = 1, \ldots, m_2$, are of the form $h_j + k$, $k \in \mathbb{Z}$ for some h_j. Then, all cells can be defined as $\prod_{j=1,\ldots,m_2} (a_j, a_{j+1}]$.

Sampling

In many cases, the second-stage value function may be hard or impossible to compute, even for a fixed value of the first-stage solution. It is natural then to design methods based on some form of sampling. Such methods are currently used in continuous stochastic programming and are described elsewhere in this handbook. We provide here an introduction to some specific aspects of stochastic discrete programs.

Stochastic branch & bound

At a given stage of the algorithm, one considers a current problem

$$(\mathrm{CP})_\rho \quad z^*(N^\rho) = \min\{c^T x + \mathbb{E}_\mu[\Psi(x, \xi)] : x \in P \cap N^\rho\} \tag{3.82}$$

where P is a formulation of X, N^ρ, is a node of the first-stage branching scheme, $\mathbb{R}^n = \cup_{\rho=1,\ldots,R} N^\rho$ and

$$\Psi(x, \xi) = \min\{q^T(\xi) y : W(\xi) y \leq h(\xi) - T(\xi) x,\ y \in Y\}. \tag{3.83}$$

As $z^*(N^\rho)$ is too difficult to compute, the stochastic branch & bound does try to find a lower bound function $L(N^\rho)$ and an upper bound function $U(N^\rho)$ which satisfy the following two conditions

(i) for every $N^\rho \neq \emptyset$, $L(N^\rho) \leq z^*(N^\rho) \leq U(N^\rho)$
(ii) for every singleton $x \in X$, $L(\{x\}) = z^*(\{x\}) = U(\{x\})$.As usual, $z^* = \min_{\rho=1,\ldots,R} z^*(N^\rho)$. Also, $z^* \leq \overline{z} = \min_{\rho=1,\ldots,R} U(N^\rho)$. Thus, a node N^ρ can only be fathomed when $L(N^\rho) \geq \overline{z}$. Alternatively, a node $N_\ell = \emptyset$ can also be fathomed.

There are various ways to obtain lower and upper bound functions in stochastic discrete programs. One such way is to use the classical wait-and-see (WS) and expected result of using the EV solution (EEV) values. Let $\overline{x}(\xi) \in \arg\min\{c^T x + \Psi(x, \xi) : x \in P \cap N^\rho\}$ be the WS solution for a given $\xi \in \Xi$ and node N^ρ. Similarly, let $x^* \in \arg\min\{c^T x + \mathbb{E}_\mu[\Psi(x, \xi)] : x \in P \cap N^\rho\}$ be a solution of the current problem $(\mathrm{CP})_\rho$. For any $\xi \in \Xi$, $c^T \overline{x}(\xi) + \Psi(\overline{x}(\xi), \xi) \leq c^T x^* + \Psi(x^*, \xi)$. The lower bound function is obtained by taking the expectation of the left-hand-side of this inequality:

$$L(N^\rho) := \mathbb{E}_\mu[\min\{c^T x + \Psi(x, \xi) : x \in P \cap N^\rho\}].$$

By definition, the expectation of the right-hand of the same inequality is precisely $z^*(N^\rho)$, which proves the condition $L(N^\rho) \leq z^*(N^\rho)$ holds. Moreover, for every singleton x, we have $L(\{x\}) = z^*(\{x\}) = c^T x + \mathbb{E}_\mu[\Psi(x, \xi)]$.

Similarly, let $\overline{x}_\rho \in \arg\min\{c^T x + \Psi(x, \mathbb{E}_\mu[\xi]) : x \in P \cap N^\rho\}$ be the solution of the EV problem on node N^ρ. We define

$$U(N^\rho) := \mathbb{E}_\mu[c^T \overline{x}_\rho + \Psi(\overline{x}_\rho, \xi)]$$

which is the EEV value for node N^ρ. As $\overline{x}_\rho$ is a feasible solution to $(\mathrm{CP})_\rho$, it follows that $z^*(N^\rho) \leq U(N^\rho)$ holds. Clearly, for every singleton, $U(\{x\}) = z^*(\{x\})$. We now illustrate why finding $U(N^\rho)$ and, especially, $L(N^\rho)$ is still a hard problem.

Example 31. *Project Financing*. Assume we can invest in n projects. Project j can be started at a cost c_j. In the second stage, projects which have been started can be continued or not. If project j is continued, it provides a revenue $-q_j(\xi)$. Project j requires t_{ij} units of resource i in the first stage and an additional $w_{ij}(\xi)$ units to be continued. A total of $h_i(\xi)$ units of resource i will be available over the two stages. The two-stage stochastic program reads as $\min\{c^T x + \mathbb{E}_\mu[\Psi(x, \xi)] : x \in P\}$ where P is a formulation of some first-stage constraints and

$$\Psi(x, \xi) = \min\{q^T(\xi) y : W(\xi) y \leq h(\xi) - Tx, y \leq x, y \in \{0, 1\}^n\}. \quad (3.84)$$

The current problem is as (3.82). To obtain an upper bound on N^ρ, we solve the expected value problem

$$\min\{c^T x + \mathbb{E}[q^T(\xi)]y : Tx + \mathbb{E}[W(\xi)]y \leq \mathbb{E}[h(\xi)], \\ y \leq x, y \in \{0, 1\}^n, x \in P \cap N^\rho\}. \tag{3.85}$$

This is a multiknapsack problem. For simplicity, assume $P \subseteq \{0,1\}^n$. Now, N^ρ will include additional restrictions on x. If $x \in N^\rho$ implies $x_j = 0$ for some j, then x_j and y_j can be removed from the computation in (3.84). If $x \in N^\rho$ implies $x_j = 1$ for some j, then x_j is also removed from the computation in (3.84). Finally, if $x \in N^\rho$ implies $x_j \in \{0,1\}$, then it follows that (3.84) can be simplified as $y_j = x_j$ in any deterministic solution. Thus, finding $\overline{x}_\rho$ is relatively easy. Yet, finding $U(N^\rho)$ still requires computing $\mathbb{E}_\mu[\Psi(\overline{x}_\rho, \xi)]$. This involves solving a multiknapsack problem for each $\xi \in \Xi$. Unless the support of Ξ is discrete with low cardinality, $\mathbb{E}_\mu[\Psi(\overline{x}_\rho, \xi)]$ cannot be computed exactly, but only estimated through sampling. Similarly, computing $L(N^\rho)$ amounts to solving a deterministic multiknapsack for each $\xi \in \Xi$. The same simplifications as above are available, yet $L(N^\rho)$ cannot be computed exactly but only estimated through sampling.

A stochastic lower bound function can be obtained by Monte Carlo simulation. For i.i.d. observations ξ^k, $k = 1, \ldots, S$, one obtains

$$\kappa(n^\rho) = \frac{1}{S} \sum_{k=1}^{S} \min\{c^T x + \Psi(x, \xi^k) : x \in P \cap N^\rho\}.$$

For the same sample, one obtains a stochastic upper bound

$$\eta(N^\rho) = \frac{1}{S} \sum_{k=1}^{S} (c^T \overline{x}_\rho + \Psi(\overline{x}_\rho, \xi^k))$$

where, as before, $\overline{x}_\rho$ is the solution of the EV problem on node N^ρ.

When the lower and upper bounds are random variables, fathoming of the nodes becomes more problematic. Deletion of a node on the basis of the statistical estimates may lead to the loss of the optimal solution. On the other hand, the stochastic lower and upper bounds $\kappa(N^\rho)$ and $\eta(N^\rho)$ can be updated each time a new sample is drawn, in a manner similar to the one which updates cuts in a stochastic decomposition method. As $\lim_{S\to\infty} \kappa(N^\rho) \to L(N^\rho)$, a.s. and $\lim_{S\to\infty} \eta(N^\rho) = U(N^\rho)$ a.s., the errors $|\kappa(N^\rho) - L(N^\rho)|$ and $|\eta(N^\rho) - U(N^\rho)|$ can be bounded in probability. Then, deletion can only occur after a sufficiently larger number of iterations where the estimations of the bounds are improved. Convergence of the stochastic

branch & bound relies on two arguments. As just mentioned, repeated sampling lets the stochastic bounds tend to the lower and upper bounds $L(N^\rho)$ and $U(N^\rho)$. On the other hand, repeated partitioning lets the nodes tend to singletons for which, by construction, $L(\{x\}) = z^*(\{x\}) = U(\{x\})$. A detailed description can be found in Norkin et al. (1998).

Sample average approximation

Consider a sample ξ^k, $k = 1, \ldots, S$, of sample scenarios generated from ω according to the probability distribution P. The Sample Average Approximation problem is the following

$$\text{(SAA)} \quad z_S = \min\left\{ c^T x + \frac{1}{S} \sum_{k=1}^{S} \Psi(x, \xi^k) : x \in P \right\}$$

where P is a formulation of X. It is a stochastic program with discrete distribution. It can be solved by the classical L-shaped method if the second-stage is continuous, combined with a branch & cut scheme to recover integrality in the first-stage. The SAA method proceeds by solving problem (SAA) repeatedly. Assume M independent samples, each of size S, are generated and, for each, the corresponding (SAA) problem is solved. We obtain optimal values $z_S^1, z_S^2, \ldots, z_S^M$ and associated candidate solutions $\hat{x}^1, \hat{x}^2, \ldots, \hat{x}^M$. It is then natural to pick-up one of these as the (estimated) optimal solution. This would require finding the one which minimizes $c^T x + \mathbb{E}_\mu[\Psi(x, \xi)]$. As this calculation is again impossible to carry out, this value is estimated by

$$\hat{z}_N(x) = c^T x + \frac{1}{N} \sum_{n=1}^{N} \Psi(x, \xi^n)$$

where $\{\xi^1, \ldots, \xi^N\}$ is a sample of size N. Typically, N is chosen to be quite large and must be independent of the M samples generated above.

The estimated optimal solution is

$$\hat{x}^* \in \arg\min \{\hat{z}_N(x) : x \in \{\hat{x}^1, \hat{x}^2, \ldots, \hat{x}^M\}\}.$$

Now, for any x, $\hat{z}_N(x)$ is an unbiased estimator of $c^T x + \mathbb{E}_\mu[\Psi(x, \xi)]$, and therefore, for any feasible x, we have $\mathbb{E}_\mu[\hat{z}_N(x)] \geq z^*$. On the other hand, we may also consider the average of the M optimal values of the SAA problems

$$\overline{z}_S = \frac{1}{M} \sum_{m=1}^{M} z_S^m.$$

As $\mathbb{E}[\overline{z}_S] = WS$, it follows that $\mathbb{E}[\overline{z}_S] \leq z^*$. Thus, the quality of the solution $\hat{x}^*$ can be evaluated by computing the optimality gap estimate $\hat{z}_N(\hat{x}^*) - \overline{z}_S$. As this quantity is random, it is helpful to estimate its variance. As the samples used to compute those quantities are independent, the variance of the gap is simply the sum of the variances of the two terms. These are

$$\hat{\sigma}^2_{\overline{z}_S} = \frac{1}{(M-1)M} \sum_{m=1}^{M} (z_S^m - \overline{z}_S)^2$$

and

$$\hat{\sigma}^2_{\hat{z}_N(\hat{x}^*)} = \frac{1}{(N-1)N} \sum_{n=1}^{N} (c^T \hat{x}^* + \Psi(\hat{x}, \xi^n) - \hat{z}_N(\hat{x}^*))^2.$$

The above procedure for statistical evaluation of a candidate solution was suggested in Mak et al. (1999). Convergence properties of the SAA method were studied in Kleywegt et al. (2001) and Ahmed and Shapiro (2002). An example of application to routing problems can be found in Verweij et al. (2003).

3.5 *Bibliographical notes*

The systematic investigation of stochastic integer programs started during the 1990s, only. Survey and introductory articles were published by Klein Haneveld and van der Vlerk (1999) and Römisch and Schultz (2001). An annotated bibliography was compiled by Stougie and van der Vlerk (1997). The textbook Birge and Louveaux (1997) has a chapter on stochastic integer programs. Integer programming basics are laid out in the textbooks Nemhauser and Wolsey (1988) and Wolsey (1998), for instance.

The first continuity result for the expected recourse function of a stochastic linear program with integer recourse is due to Stougie (1985). The sufficient conditions for lower semicontinuity, continuity, and Lipschitz continuity of Propositions 3 and 4, with extensions to random technology matrix T in (2.4), were derived in Schultz (1993, 1995). Joint qualitative (semi-) continuity of the expected recourse as a function of the first-stage decision and the integrating probability measure together with conclusions towards stability were addressed in Artstein and Wets (1994), Schultz (1992, 1995). Quantitative joint continuity of the expected recourse function and quantitative stability of the problem (2.3) were studied in Rachev and Römisch (2002) and in Schultz (1996).

For the first time, stochastic programs with simple integer recourse were investigated in Louveaux and van der Vlerk (1993). The analysis was pushed ahead in Klein Haneveld and van der Vlerk (1994), Klein Haneveld et al. (1996, 1997), see also the dissertation van der Vlerk (1995).

So far, very little is known about the structure of multi-stage stochastic integer programs. An introduction into this problem class was given in Römisch and Schultz (2001). Our exposition in Section 2.3 essentially follows parts of that paper. Power problems leading to multi-stage stochastic integer programs and specialized methods for their solution were discussed in Dentcheva and Römisch (1998), Nowak (2000), Nowak and Römisch (2000) and Takriti et al. (1996).

More details on the definitions and properties presented in Section 3.1 can be found in Wolsey (1998).

The first use of decomposition methods in stochastic programs having discrete decisions in the second-stage is the integer L-shaped method of Laporte and Louveaux (1993) for the case of first-stage binary variables. This method has been applied to solve a variety of location and routing problems. The example considered in Section 3.3.2 is taken from Laporte et al. (2002). The generalization of the integer L-shaped to mixed integer first-stage was done by Carøe (1998) in his doctoral dissertation and presented in Carøe and Tind (1997, 1998).

The first attempt to design a method based on first-stage integer variables branching is due to Ahmed et al. (2000), see the exposition in Section 3.4. In Schultz et al. (1998) a solution procedure based on enumeration and bounding in the first-stage while handling the second-stage by algebraic methods exploiting problem similarities has been proposed. Another application of algebraic methods to stochastic integer programming has been carried out in Hemmecke and Schultz (2003). The paper deals with test set decomposition. It identifies building blocks that enable generation of improving vectors for the stochastic program and that can be computed by a completion procedure from computational algebra.

The scenario decomposition method displayed in Section 3.1.3 stems from Carøe and Schultz (1999). In Løkketangen and Woodruff (1996) a scenario decomposition method employing Augmented Lagrangians for the dualization and tabu search for the resulting quadratic mixed-integer subproblems has been developed.

Recently, Alonso-Ayuso et al. (2003) have proposed a branch-and-fix coordination approach to solving multi-stage stochastic integer programs. Their idea is to formulate the stochastic program in a splitting variable representation, to perform scenario-wise an LP-based branch & bound, and to establish the nonanticipativity in the course of the branching. The latter is achieved by a coordinated fixing of variables that puts variables for different scenarios at identical values provided a nonanticipativity condition requires the identity.

References

Ahmed, S., A. Shapiro (2002). The sample average approximation method for stochastic programs with integer recourse, Preprint, Georgia Institute of Technology, Atlanta.

Ahmed, S., M. Tawarmalani, N.V. Sahinidis (2000). A finite branch and bound algorithm for two-stage stochastic integer programs. Stochastic Programming E-Print Series 16—2000, http://dochost.rz.hu-berlin.de/speps/.

Alonso-Ayuso, A., L.F. Escudero, A. Garín, M.T. Ortuño, G. Pérez (2003). An approach for strategic supply chain planning under uncertainty based on stochastic 0-1 programming. *Journal of Global Optimization* 26, 97–124.

Artstein, Z., R.J.-B. Wets (1994). Stability results for stochastic programs and sensors, allowing for discontinuous objective functions. *SIAM Journal on Optimization* 4, 537–550.

Bacaud, L., C. Lemaréchal, A. Renaud, C. Sagastizábal (2001). Bundle methods in stochastic optimal power management: A disaggregated approach using preconditioners. *Computational Optimization and Applications* 20, 227–244.

Bank, B., J. Guddat, D. Klatte, B. Kummer, K. Tammer (1982). *Non-linear Parametric Optimization*, Akademie-Verlag, Berlin.

Bank, B., R. Mandel (1988). *Parametric Integer Optimization*, Akademie-Verlag, Berlin.

Birge, J.R., F.V. Louveaux (1997). *Introduction to Stochastic Programming*, Springer-Verlag, New York.

Blair, C.E., R.G. Jeroslow (1977). The value function of a mixed integer program: I. *Discrete Mathematics* 19, 121–138.

Carøe, C.C. (1998). Decomposition in stochastic integer programming, Ph.D. Thesis, Institute of Mathematical Sciences, Department of Operations Research, University of Copenhagen, Denmark.

Carøe, C.C., R. Schultz (1999). Dual decomposition in stochastic integer programming. *Operations Research Letters* 24, 37–45.

Carøe, C.C., J. Tind (1997). A cutting plane approach to mixed 0–1 stochastic integer programs. *European Journal of Operations Research* 101, 306–316.

Carøe, C.C., J. Tind (1998). *L*-shaped decomposition of two-stage stochastic programs with integer recourse. *Mathematical Programming* 83, 139–152.

Dentcheva, D., W. Römisch (1998). Optimal power generation under uncertainty via stochastic programming, in: K. Marti, P. Kall (eds.), *Stochastic Programming Methods and Technical Applications. Lecture Notes in Economics and Mathematical Systems,* Vol. 458*,* Springer-Verlag, Berlin, pp. 22–56.

Dynkin, E.B., I.V. Evstigneev (1976). Regular conditional expectation of correspondences. *Theory of Probability and Applications* 21, 325–338.

Evstigneev, I. (1976). Measurable selection and dynamic programming. *Mathematics of Operations Research* 1, 267–272.

Gröwe-Kuska, N., K.C. Kiwiel, M.P. Nowak, W. Römisch, I. Wegner (2002). Power management in a hydro-thermal system under uncertainty by Lagrangian relaxation, in: C. Greengard, A. Ruszczyński (eds.), *Decision Making under Uncertainty: Energy and Power IMA Volumes in Mathematics and its Applications*, Vol. 128*,* Springer-Verlag, New York.

Hemmecke, R., R. Schultz (2003). Decomposition of test sets in stochastic integer programming. *Mathematical Programming* 94, 323–341.

Hiriart-Urruty, J.B., C. Lemaréchal (1993). *Convex Analysis and Minimization Algorithms*, Springer-Verlag, Berlin.

Kiwiel, K.C. (1990). Proximity control in bundle methods for convex nondifferentiable optimization. *Mathematical Programming* 46, 105–122.

Klein Haneveld, W.K., M.H. van der Vlerk (1994). On the expected value function of a simple integer recourse problem with random technology matrix. *Journal of Computational and Applied Mathematics* 56, 45–53.

Klein Haneveld, W.K., L. Stougie, M.H. van der Vlerk (1996). An algorithm for the construction of convex hulls in simple integer recourse programming. *Annals of Operations Research* 64, 67–81.

Klein Haneveld, W.K., L. Stougie, M.H. van der Vlerk (1997). Convex approximations for simple integer recourse models by perturbing the underlying distribution. Research Report 97A19, SOM, University of Groningen.

Klein Haneveld, W.K., M.H. van der Vlerk (1999). Stochastic integer programming: General models and algorithms. *Annals of Operations Research* 85, 39–57.

Kleywegt, A.J., A. Shapiro, T. Homem-de-Mello (2001). The sample average approximation method for stochastic discrete optimization. *SIAM Journal on Optimization* 12, 479–502.

Laporte, G., F.V. Louveaux (1993). The integer *L*-shaped method for stochastic integer programs with complete recourse. *Operations Research Letters* 13, 133–142.

Laporte, G., F.V. Louveaux, L. Van Hamme (2002). An integer *L*-shaped algorithm for the capacitated vehicle routing problem with stochastic demands. *Operations Research* 50, 415–423.

Løkketangen, A., D.L. Woodruff (1996). Progressive hedging and tabu search applied to mixed integer (0,1) multi-stage stochastic programming. *Journal of Heuristics* 2, 111–128.

Louveaux, F.V., M.H. van der Vlerk (1993). Stochastic programming with simple integer recourse. *Mathematical Programming* 61, 301–325.

Mak, W.K., D.P. Morton, R.K. Wood (1999). Monte-Carlo bounding techniques for determining solution quality in stochastic programs. *Operations Research Letters* 24, 47–56.

Nemhauser, G.L., L.A. Wolsey (1988). *Integer and Combinatorial Optimization*, Wiley, New York.

Norkin, V.I., Y.M. Ermoliev, A. Ruszczyński (1998). On optimal allocation of indivisibles under uncertainty. *Operations Research* 46, 381–395.

Nowak, M.P. (2000). Stochastic Lagrangian relaxation in power scheduling of a hydro-thermal system under uncertainty. Ph.D. Thesis, Institute of Mathematics, Humboldt-University Berlin.

Nowak, M.P., W. Römisch (2000). Stochastic Lagrangian relaxation applied to power scheduling in a hydro-thermal system under uncertainty. *Annals of Operations Research* 100, 251–272.

Rachev, S.T., W. Römisch (2002). Quantitative stability in stochastic programming: The method of probability metrics. *Mathematics of Operations Research* 27, 792–818.

Rockafellar, R.T., R.J.-B. Wets (1978). The optimal recourse problem in discrete time: L^1-multipliers for inequality constraints. *SIAM Journal on Control and Optimization* 16, 16–36.

Rockafellar, R.T., R.J.-B. Wets (1997). *Variational Analysis*, Springer-Verlag, Berlin.

Römisch, W., R. Schultz (2001). Multistage stochastic integer programs: an introduction, in: M. Grötschel, S.O. Krumke, J. Rambau (eds.), *Online Optimization of Large Scale Systems*, Springer-Verlag, Berlin, pp. 581–600.

Schultz, R. (1992). Continuity and stability in two-stage stochastic integer programming, in: Marti, K. (ed.), *Stochastic Optimization, Numerical Methods and Technical Applications. Lecture Notes in Economics and Mathematical Systems*, Vol. 379, Springer, Berlin.

Schultz, R. (1993). Continuity properties of expectation functions in stochastic integer programming. *Mathematics of Operations Research* 18, 578–589.

Schultz, R. (1995). On structure and stability in stochastic programs with random technology matrix and complete integer recourse. *Mathematical Programming* 70, 73–89.

Schultz, R. (1996). Rates of convergence in stochastic programs with complete integer recourse. *SIAM Journal on Optimization* 6, 1138–1152.

Schultz, R., L. Stougie, M.H. van der Vlerk (1998). Solving stochastic programs with integer recourse by enumeration: A framework using Gröbner basis reductions. *Mathematical Programming* 83, 229–252.

Stougie, L. (1985). Design and analysis of algorithms for stochastic integer programming. Ph.D. Thesis, Center for Mathematics and Computer Science, Amsterdam.

Stougie, L., M.H. van der Vlerk (1997). Stochastic integer programming: *Annotated Bibliographies in Combinatorial Optimization*, in: M. Dell'Amico, F. Maffioli, S. Martello (eds.), Wiley, Chichester, pp. 127–141.

Takriti, S., J.R. Birge, E. Long (1996). A stochastic model for the unit commitment problem. *IEEE Transactions on Power Systems* 11, 1497–1508.

van der Vlerk, M.H. (1995). Stochastic programming with integer recourse. Ph.D. Thesis, University of Groningen.

Verweij, B., S. Ahmed, A.J. Kleywegt, G.L. Nemhauser, A. Shapiro (2003). The sample average approximation method applied to stochastic routing problems: a computational study. *Computational Optimization and Applications* 24, 289–333.

Wolsey, L.A. (1998). *Integer Programming*, Wiley, New York.

A. Ruszczyński and A. Shapiro, Eds., *Handbooks in OR & MS, Vol. 10*

Chapter 5

Probabilistic Programming

András Prékopa

RUTCOR, Rutgers Center for Operations Research, 640 Bartholomew Road, Piscataway, NJ 08854-8003, USA

Abstract

Probabilistic programming means two strongly connected models as well as the study of their mathematical properties, solutions of the relevant optimization problems and their applications. The two models are: maximizing (or minimizing) a probability under constraints and programming under probabilistic constraints. There are a number of variants and special cases of these models and we present them in Section 1. In Section 2 we summarize those mathematical theories which can be used to prove the convexity of large classes of our problems and we also show how they can be applied in this context. In Section 3 we present solution algorithms of our stochastic programming problems. Since we are handling probabilities of sets in higher dimensional spaces, it is necessary to use bounding and other approximation algorithms to find these probabilities with satisfactory precision. This is the subject of Section 5. In Section 4 we present two-stage and multi-stage problems which are combined with probabilistic constraints. Some duality and stability theorems are presented in Section 6. Finally, in Section 7, we present applications of our model constructions.

1 Model constructions

1.1 Statistical decisions

Stochastic programming is a science that solves problems in connection with stochastic systems, where the mathematical form of the problem is of optimization type. It follows from this that the main ingredients of this science are: statistical decision principles, optimization methods and computer science. If decision is taken only once in time, then the model is static.

If decisions are taken subsequently in such a way that between two subsequent decisions an observation of a random variable takes place, then the model is dynamic.

Looking at the problem from another angle, a stochastic programming model can be of a type where the functioning of the system is paramount and we maximize the probability of the system functioning or optimize another objective function subject to a probabilistic constraint. Another case is, where we allow violations of the constraints that describe the system but penalize them in such a way that the system cost plus the expected penalty of violations should be as small as possible. Both of these principles can be used in static and dynamic model constructions.

One of the simplest and most classical examples of statistical decisions is the testing of statistical hypotheses. Assume, for the sake of simplicity, that we have two probability distributions $P = \{p_k\}$ and $Q = \{q_k\}$ on the set of integers and we want to test the hypothesis that P is the true distribution, against the alternative that it is Q. The way the test is constructed is as follows: find a set of integers, the critical set, such that its probability with respect to P is smaller than or equal to $\varepsilon > 0$ (where ε is a previously given fixed number) and the probability of this set with respect to Q should be maximal. The problem can be formalized in the following way:

$$\begin{aligned} &\max \sum_k q_k x_k \\ &\text{subject to} \\ &\sum_k p_k x_k \leq \varepsilon, \end{aligned} \tag{1.1}$$

where $x_k \in \{0, 1\}$ for every k. The critical set C is then given by

$$C = \{k \mid x_k = 1\}.$$

The probability $P(C)$ is called the first kind error, while the probability $Q(\overline{C})$ is the second kind error.

After the construction of the test we work with it in such a way that we take a sample, a random integer, and reject the hypothesis P if the integer is in C otherwise we accept it (or at least say that the sample does not contradict the hypothesis.) Problem (1.1) can be termed the Neyman-Pearson problem. Problem (1.1) can also be described in such a way that the first kind error (rejecting the true hypothesis) should be smaller than or equal to the small number ε, and, given this, the second kind error (accepting a false hypothesis) should be as small as possible.

The above statistical decision is static one. An example for dynamic type statistical decision situation is provided by Wald's sequential analysis, where

we want to decide in favor or against the hypothesis by given probabilities, respectively, by the use of a sequential decision process.

In the above examples we see that optimization under probabilistic constraints and maximizing a probability under constraints are classical statistical decision principles.

1.2 Static stochastic programming models: programming under probabilistic constraints and maximizing a probability under constraints

A general formulation of the first problem mentioned in the title is the following:

$$\begin{aligned} &\min h(x) \\ &\text{subject to} \\ &h_0(x) = P(g_1(x, \xi) \geq 0, \ldots, g_r(x, \xi) \geq 0) \geq p_0 \\ &h_1(x) \geq p_1, \ldots, h_m(x) \geq p_m, \end{aligned} \tag{1.2}$$

where x is the decision vector, ξ is a random vector, $h(x)$, $h_1(x),\ldots,h_m(x)$ are given functions, $0 < p_0 \leq 1$, $p_1,\ldots,p_m$ are given numbers.

A special case of problem (1.2) is the following:

$$\begin{aligned} &\min c^T x \\ &\text{subject to} \\ &P(Tx \geq \xi) \geq p \\ &Ax \geq b,\ x \geq 0. \end{aligned} \tag{1.3}$$

Sometimes we have no special objective function that expresses some cost and we only want to maximize a probability. The general form of this problem is

$$\begin{aligned} &\max P(g_1(x, \xi) \geq 0, \ldots, g_r(x, \xi) \geq 0) \\ &\text{subject to} \\ &h_1(x) \geq p_1, \ldots, h_m(x) \geq p_m. \end{aligned} \tag{1.4}$$

A stochastic programming problem is frequently formulated in such a way that we have a "deterministic underlying problem" also called "base problem". Then we observe that some of the parameters in it are not constants but random variables and therefore the problem is meaningless in its original form. The use of a statistical decision principle takes us to a stochastic programming problem which can be problem (1.1) or (1.2).

Note that the use of a probabilistic constraint does not exclude the use of another principle: penalizing constraint violations. In fact, both principles can be used simultaneously, thereby arriving at a hybrid model. For example, a hybrid model constructed from problem (1.3) is the following:

$$\begin{aligned} &\min\left\{c^T x + \sum_{i=1}^{r} q_i E([\xi_i - T_i x]_+)\right\} \\ &\text{subject to} \\ &P(Tx \geq \xi) \geq p \\ &Ax \geq b,\ x \geq 0, \end{aligned} \tag{1.5}$$

where $T_1, \ldots, T_r$ are the rows of the matrix T, $\xi_1, \ldots, \xi_r$ are the components of the random vector ξ and $q_1, \ldots, q_r$ are nonnegative constants, penalizing violations of $T_1 x \geq \xi_1, \ldots, T_r x \geq \xi_r$, respectively.

1.3 *Related measures of violation*

In connection with the stochastic constraints $T_i x \geq \xi_i$, $i = 1, \ldots, r$, some measure of violation has been incorporated into the objective function in problem (1.5). There are, however, other measures of violation that can be used in practice. We present two of them.

The first one, introduced in Prékopa (1973a), is the collection of the conditional expectations $E(\xi i - T_i x \mid \xi_i - T_i x > 0)$, $i = 1, \ldots, r$. We can incorporate them into problem (1.5) but we prefer to use them among the constraints, rather than in the objective function. The new constraints that we supplement to problem (1.5) are:

$$E(\xi_i - T_i x \mid \xi_i - T_i x > 0) \leq d_i \quad i = 1, \ldots, r, \tag{1.6}$$

where d_i, $i = 1, \ldots, r$ are some given constants. We call them conditional expectation constraints. Problem (1.5), together with the constraints (1.6) prescribes that if we decide on x and after that observe the random vector ξ, then in at least $p100\%$ of the cases no violation occurs in $Tx \geq \xi$, further, if we single out the cases where $\xi_i - T_i x > 0$, then the average magnitude of the violation should be less than or equal to d_i, $i = 1, \ldots, r$. Subject to these constraints we minimize the system cost $c^T x$ plus the expected penalized sum of the violations.

Another measure of violation is the integrated chance or probabilistic constraint. For the case of $r = 1$ it is defined as

$$E([\xi - Tx]_+) \leq d. \tag{1.7}$$

For the case of $r \geq 1$ the integrated chance constraint is

$$E\left(\max_i [\xi_i - T_i x]_+\right) \leq d. \tag{1.8}$$

In (1.7) and (1.8) d is a constant.

A more general form of this kind of constraint can be formulated as follows. Suppose that in the underlying problem we have the stochastic constraints

$$g_i(x, \xi) \leq 0, \quad i = 1, \ldots, r,$$

where $g_1, \ldots, g_r$ are some functions. Then in the stochastic programming problem we may include the constraint

$$E\left(\max_i [g_i(x, \xi)]_+\right) \leq d. \tag{1.9}$$

A practical interpretation of the integrated chance constraint is the following. If ξ designates the demand for power in an area on a given day and Tx is the total generating capacity ($r = 1$ now), then $E([\xi - Tx]_+)$ is the expected unserved energy. In power system engineering this measure is considered equally important to loss of load probability (LOLP), accounted for in probabilistic constraints.

1.4 Bibliographical notes

Historically the first paper that used the programming under probabilistic constraint principle was the one by Charnes et al. (1958), where, however, probabilistic constraints are imposed individually on each constraint involving random variables. This formulation called "chance constrained programming" by the authors may be correct in some cases, especially when the random variables, appearing in different stochastic constraints, are independent. In general, however, it has the serious defect of ignoring joint probability distribution, i.e., the type of stochastic dependence of the random variables involved. A paper by Miller and Wagner (1965) takes the probabilistic constraint jointly on the stochastic constraints but handles only independent random variables appearing on the right hand sides of the stochastic constraints. Prékopa (1970, 1973a) initiated the model and its research, where the probabilistic constraint is taken jointly for the stochastic constraints and the random variables involved are stochastically dependent, in general.

Constraints of type (1.6) were introduced in Prékopa (1973a) and those in (1.8) by Klein Haneveld (1986).

2 Convexity theory

2.1 Basic theory of logconcave and α-concave measures

Logconcave measures have been introduced in the stochastic programming framework but they became widely used also in statistics, convex geometry, mathematical analysis, economics, etc.

Definition 2.1. A function $f(z) \geq 0$, $z \in \mathbb{R}^n$ is said to be logarithmically concave (logconcave), if for any z_1, z_2 and $0 < \lambda < 1$ we have the inequality

$$f(\lambda z_1 + (1-\lambda)z_2) \geq [f(z_1)]^{\lambda}[f(z_2)]^{(1-\lambda)}. \tag{2.1}$$

If $f(z) > 0$ for $z \in \mathbb{R}^n$, then this means that $\log f(z)$ is a concave function in $\mathbb{R}^n$.

Definition 2.2. A probability measure defined on the Borel sets of $\mathbb{R}^n$ is said to be logarithmically concave (logconcave) if for any convex subsets of $\mathbb{R}^n$: A, B and $0 < \lambda < 1$ we have the inequality

$$P(\lambda A + (1-\lambda)B) \geq [P(A)]^{\lambda}[P(B)]^{1-\lambda}, \tag{2.2}$$

where $\lambda A + (1-\lambda)B = \{z = \lambda x + (1-\lambda)y \mid x \in A,\ y \in B\}$. The basic theorem of logconcave measures is the following.

Theorem 2.1. *If the probability measure P is absolutely continuous with respect to the Lebesgue measure and is generated by a logconcave probability density function then the measure P is logconcave.*

Remark 2. The proof of Theorem 2.1 provides us with a more general assertion: if P is an absolutely continuous probability measure (that can be extended in a trivial way to all measurable subsets of $\mathbb{R}$) and A, $B \subset \mathbb{R}^n$ are two Borel sets then we have the inequality (2.2) for any $0 < \lambda < 1$. In this assertion we used the fact that Borel measurability of A and B implies the (Lebesgue) measurability of $\lambda A + (1-\lambda)B$. Even though this more general assertion holds true, Theorem 2.1 provides us with enough basis to derive our convexity theory of stochastic programming.

Theorem 2.2. *If P is a logconcave probability distribution and $A \subset \mathbb{R}^n$ is a convex set, then $P(A + x)$, $x \in \mathbb{R}^n$ is a logconcave function.*

Theorem 2.3. *If $\xi \in \mathbb{R}^n$ is a random variable, the probability distribution of which is logconcave, then the probability distribution function $F(x) = P(\xi \leq x)$ is a logconcave function in $\mathbb{R}^n$.*

Theorem 2.4. *If $n=1$ in Theorem 2.3 then also $1-F(x)=P(\xi>x)$ is a logconcave function in $\mathbb{R}^1$.*

Theorems 2.2, 2.3 and 2.4 are easy consequences of the notion of a logconcave measure. Less obvious are the following

Theorem 2.5. *If $g_1(x,y),\ldots,g_r(x,y)$ are quasi-concave functions of the variables $x\in\mathbb{R}^n$, $y\in\mathbb{R}^m$ and $\xi\in\mathbb{R}^m$ is a random variable that has logconcave probability distribution, then the function $G(x)=P(g_1(x,\xi)\geq 0,\ldots,g_r(x,\xi)\geq 0)$, $x\in\mathbb{R}^n$ is logconcave.*

Theorem 2.6. *If $f(x,y)$, $x\in\mathbb{R}^n$, $y\in\mathbb{R}^m$ is a logconcave function, then*

$$\int_{\mathbb{R}^m} f(x,y)\,dy, \quad x\in\mathbb{R}^n$$

is also a logconcave function. The above theorem implies

Theorem 2.7. *If $f(x)$, $g(x)$, $x\in\mathbb{R}^n$, are logconcave functions then their convolution*

$$\int_{\mathbb{R}^m} f(x-y)g(y)\,dy, \quad x\in\mathbb{R}^n$$

is also logconcave.

Theorem 2.8. *If the random vector ξ has logconcave probability distribution and A is a constant matrix of appropriate size, then $\eta=A\xi$ also has a logconcave distribution.*

Definition 2.3. A function $f(z)$, $z\in\mathbb{R}^n$ is said to be α-concave if for any $x,y\in\mathbb{R}^n$, such that $f(x)>0$, $f(y)>0$ we have the inequality

$$f(\lambda x+(1-\lambda)y)\geq[\lambda f^{\alpha}(x)+(1-\lambda)f^{\alpha}(y)]^{\frac{1}{\alpha}}, \tag{2.3}$$

where $-\infty\leq\alpha<\infty$. The expression on the right hand side is defined by continuity for the cases of $\alpha=-\infty$, $\alpha=0$.

Inequality (2.3) for different α values means the following:

(a) $\alpha=-\infty$. In this case we have $f(\lambda x+(1-\lambda)y)\geq\min(f(x),f(y))$, i.e., the function is quasi-concave.

(b) $-\infty<\alpha<0$. Then we have $f^{\alpha}(\lambda x+(1-\lambda)y)\leq\lambda f^{a}(x)+(1-\lambda)f^{a}(y)$, i.e., $f^{a}(x)$ is convex on the set $\{x\,|\,f(x)>0\}$.

(c) $\alpha = 0$. We have the inequality $f(\lambda x + (1-\lambda)y) \geq [f(x)]^{\lambda}[f(y)]^{1-\lambda}$, and the function is logconcave.
(d) $0 < \alpha < \infty$. This case is similar to case (b) but now $f^{\alpha}(x)$ is concave on the set $\{x \mid f(x) > 0\}$.

Definition 2.4. The probability measure $\boldsymbol{P}$ defined on the Borel sets of $\mathbb{R}^n$ is said to be α-concave, if for any convex subsets of $\mathbb{R}^n$: A, B for which $\boldsymbol{P}(A)$, $\boldsymbol{P}(B) > 0$ and $< \lambda < 1$ we have the inequality

$$\boldsymbol{P}(\lambda A + (1-\lambda)B) \geq (\lambda[\boldsymbol{P}(A)]^{\alpha} + (1-\lambda)[\boldsymbol{P}(B)]^{\alpha})^{\frac{1}{\alpha}}, \tag{2.4}$$

where $-\infty \leq \alpha < \infty$. The expression on the right hand side is defined by continuity for the cases $\alpha = -\infty$, $\alpha = 0$. The inequality (2.4) has similar interpretation in the special cases as the inequality (2.3). If $\alpha = -\infty$, then the measure $\boldsymbol{P}$ is said to be quasi-concave. If $\alpha = 0$, then $\boldsymbol{P}$ is a logconcave measure.

Theorem 2.9. *If the probability measure $\boldsymbol{P}$ is generated by an α-concave probability density function, then $\boldsymbol{P}$ is γ-concave, where $\gamma = \alpha/(1+n\alpha)$.*

Theorem 2.10. *If the probability measure $\boldsymbol{P}$ is generated by the probability density function $f(z)$, $z \in \mathbb{R}^n$ which has the property that $f^{-1/n}(z)$, $z \in \mathbb{R}^n$ is convex, then the probability measure $\boldsymbol{P}$ is quasi-concave, i.e., for any convex subsets A, B of $\mathbb{R}^n$ and $0 < \lambda < 1$, we have the inequality*

$$\boldsymbol{P}(\lambda A + (1-\lambda)B) \geq \min(\boldsymbol{P}(A), \boldsymbol{P}(B)). \tag{2.5}$$

Note that in this case there is no need to assume that $\boldsymbol{P}(A) > 0$, $\boldsymbol{P}(B) > 0$ because if at least one of them is 0 then (2.5) holds trivially.

Theorem 2.9 holds in a stronger form, too. That form, however, requires the probability measure P to be defined on all (Lebesgue) measurable sets of $\mathbb{R}^n$. If P is defined on the σ-algebra of all Borel sets of $\mathbb{R}^n$, then there is a trivial extension of P to all measurable sets provided that P is generated by a probability density function. In fact, any measurable set can be obtained from a Borel set by adding to it or removing from it a measurable set of measure 0. Thus if we assign $P(C) = 0$ to all measurable set C with measure 0, then the required extension can be obtained.

It is well-known, that if A, B are Borel sets in $\mathbb{R}^n$ and $0 < \lambda < 1$, then $\lambda A + (1-\lambda)B$ is a Lebesgue measurable set.

The extension of Theorem 2.9 asserts that if P is generated by an α-convex probability density function and A, B are Borel sets in $\mathbb{R}^n$, $0 < \lambda < 1$ then (2.4) holds with α replaced by $\gamma = \alpha/(1+n\alpha)$. In the applications, however, we use

our Theorem 2.1 and 2.9 in their original forms, rather than their more general forms.

Theorem 2.5 has a counterpart for α-concave measures.

Theorem 2.11. *If $g_1(x,y),\dots,g_r(x,y)$ are quasi-concave functions of the variables $x\in\mathbb{R}^n$, $y\in\mathbb{R}^m$ and $\xi\in\mathbb{R}^m$ is a random variable that has continuous probability distribution and α-concave probability density with $1+m\alpha\geq 0$, then the function $G(x)=P(g_1(x,\xi),\dots,g_r(x,\xi)\geq 0)$ satisfies (2.3) with α replaced by $\gamma=\alpha/(1+m\alpha)$, where the inequalities for $\gamma=-\infty$, $\gamma=0$ are interpreted by continuity.*

It follows that if $\alpha=0$, then $G(x)$ is logconcave (as asserted in Theorem 2.5) and if $\alpha=-(1/m)$, then $G(x)$ is quasi-concave.

2.2 Examples of multivariate probability distributions

(1) Uniform distribution. Let D be a convex subset of $\mathbb{R}^n$ with finite, positive measure $|D|$ Then

$$f(x)=\begin{cases}\frac{1}{|D|} & \text{if } x\in D\\ 0 & \text{if } x\notin D\end{cases}$$

is a lonconcave probability density function.

(2) Normal distribution. Its probability density function is defined by

$$f(x)=\frac{1}{\sqrt{|C|}(2\pi)^{\frac{n}{2}}}\,e^{-\frac{1}{2}(x-\mu)^T C^{-1}(x-\mu)},\quad x\in\mathbb{R}^n,$$

where μ is the expectation vector and C the covariance matrix of the distribution; $|C|$ designates the determinant of C. The matrix C is supposed to be positive definite. It follows that C^{-1} is also positive definite, hence the quadratic form $(x-\mu)^T C^{-1}(x-\mu)$ is a convex function (as it is well-known and easy to prove). This implies that $f(x)$ is a logconcave function.

(3) Wishart distribution. The probability density function of it is defined by

$$f(X)=\frac{|X|^{\frac{N-p-2}{2}}\,e^{-\frac{1}{2}SpC^{-1}X}}{2^{\frac{N-1}{2}p}\,\pi^{\frac{p(p-1)}{4}}\,|C|^{\frac{N-1}{2}}\prod_{i=1}^{p}\Gamma\left(\frac{N-i}{2}\right)}$$

if X is positive definite, and $f(X)=0$ otherwise. Here C and X are $p\times p$ matrices, C is fixed and positive definite while X contains the variables.

Since X is assumed to be symmetrical, there are $n = \frac{1}{2}p(p+1)$ independent variables. We also assume that $N \geq p+2$. Since the $p \times p$ matrices form a convex set in the n-dimensional space and it is well-known that for any two $p \times p$ positive definite matrices we have the inequality

$$|\lambda X_1 + (1-\lambda)X_2| \geq |X_1|^{\lambda}|X_2|^{1-\lambda},$$

where $0 < \lambda < 1$, the function f is logconcave.

(4) Beta distribution. Its probability density function is defined by

$$f(X) = \frac{c(n_1, p)c(n_2, p)}{c(n_1+n_2, p)} |X|^{\frac{1}{2}(n_1-p-1)} |I-X|^{\frac{1}{2}(n_2-p-1)},$$

if $X, 1-X$ are positive definite $p \times p$ matrices and $f(X) = 0$ otherwise, where

$$\frac{1}{c(k,p)} = 2^{\frac{pk}{2}} \pi^{\frac{p(p-1)}{2}} \prod_{i=1}^{p} \Gamma\left(\frac{k-i+1}{2}\right).$$

It is supposed that $n_1 \geq p+1$, $n_2 \geq p+1$. The number of independent variables in X is $n = \frac{1}{2}p(p+1)$. The logconcavity of the function f follows the same way as that of f in example (3).

(5) Dirichlet distribution. Its probability density function is defined by

$$f(x) = \frac{\Gamma(p_1 + \cdots + p_{n+1})}{\Gamma(p_1) \cdots \Gamma(p_{n+1})} x_1^{p_1-1} \cdots x_n^{p_n-1}(1 - x_1 - \cdots - x_n)^{p_{n+1}-1}$$

for $x_1 > 0, \ldots, x_n > 0$, $x_1 + \cdots + x_n < 1$ and $f(x) = 0$ otherwise, where $p_1, \ldots, p_{n+1}$ are positive constants. If $p_1 \geq 1, \ldots, p_{n+1} \geq 1$, then f is a logconcave function in $\mathbb{R}^n$. If, however, $p_1 < 1, \ldots, p_{n+1} < 1$, then f is logconvex, i.e., the inequality (2.1) holds always in reversed form. The logconvexity of f does not hold in the entire space $\mathbb{R}^n$ but it holds in the open simplex $x_1 > 0, \ldots, x_n > 0$, $x_1 + \cdots + x_n < 1$.

(6) Cauchy distribution. It is the joint distribution of the random variables

$$\zeta_i = \frac{\sqrt{\nu}\xi_i}{\eta}, \quad i = 1, \ldots, n,$$

where $(\xi_1, \ldots, \xi_n)$ has standard normal distribution (each component is $N(0,1)$-distributed), $(\xi_1, \ldots, \xi_n)$ is independent of η which has

χ-distribution with ν degrees of freedom. The probability density function is

$$f(x) = \frac{\Gamma\left(\frac{1}{2}(\nu+n)\right)}{(\pi\nu)^{\frac{n}{2}}\Gamma\left(\frac{1}{2}\nu\right)|R|^{\frac{1}{2}}}\left(1+\frac{1}{\nu}\,x^T R^{-1}x\right)^{-\frac{1}{2}(\nu+n)}$$

for $x \in \mathbb{R}^n$. If $n=1$ and $\nu=1$ this reduces to the well-known univariate Cauchy density

$$f(x) = \frac{1}{\pi}\,\frac{1}{1+x^2}, \quad -\infty < x < \infty.$$

The n-variate Cauchy density has the property that $f^{-\frac{1}{n}}$ is convex in $\mathbb{R}^n$, hence Theorem 2.10 applies and the distribution is quasi-concave.

(7) Pareto distribution. Its probability density function is

$$f(x) = a(a+1)\cdots(a+n-1)\left(\prod_{j=1}^{n}\Theta_j\right)^{-1}\left(\sum_{j=1}^{n}\Theta_j^{-1}x_j - n+1\right)^{-(a+n)}$$

for $x_i > \Theta_i$, $i=1,\ldots,n$ and $f(x)=0$ otherwise; $\Theta_1,\ldots,\Theta_n$ are positive constants. Since $f^{-\frac{1}{n}}$ is convex in $\mathbb{R}^n$, Theorem 2.10 applies and the probability distribution is quasi-concave.

(8) Gamma distribution. A univariate probability distribution is said to be gamma distribution if its probability density has the form

$$f(z) = \frac{\lambda^{\vartheta} z^{\vartheta-1} e^{-\lambda z}}{\Gamma(\vartheta)} \quad \text{if } z > 0$$

and $f(z)=0$ for $z \le 0$; $\lambda > 0$, $\vartheta > 0$ are constants. If $\lambda = 1$, then the distribution is said to be standard. If ξ has gamma distribution, then $\vartheta\xi$ has standard gamma distribution. Both the expectation and the variance of a standard gamma distribution are equal to ϑ.

An r-variate gamma distribution can be defined in the following way. Let A be the $r \times (2^r-1)$ matrix the columns of which are all 0–1 component vectors of size r except for the 0 vector. Let $\eta_1,\ldots,\eta_s$, $s=2^r-1$ be independent standard gamma distributed random variables and designate by η the vector of these components. Then we say that the random vector

$$\xi = A\eta$$

has an r-variate standard gamma distribution.

Let $\vartheta_1,\ldots,\vartheta_s$ be the expectations of $\eta_1,\ldots,\eta_s$, respectively and designate by ϑ the vector of these components. Then $E(\xi)=AE(\eta)$ and the components of $E(\xi)$ are, simultaneously, the variances of the components of ξ. A practical method for fitting this distribution to empirical data is the following.

Suppose we have the random variables $\zeta_i=\frac{1}{\lambda_i}\xi_i$, $\lambda_i>0$, $i=1,\ldots,r$, where we assume that ξ can be written as $\xi=A\eta$. Suppose that we estimated $E(\xi)$ and the covariances c_{ik} of the pairs ξ_i, ξ_k, $i<k$. Let $a_1,\ldots,a_r$ be the rows of A and a_ia_k the componentwise product of a_i and a_k. Then we write up the problem:

$$\begin{aligned}&\text{find } \vartheta\\ &\text{such that}\\ &c_{ik}=(a_ia_k)\vartheta, \quad i\le k\\ &\vartheta\ge 0.\end{aligned}$$

Note that $c_{ii}=Var(\xi_i)=E(\xi_i)$, $i=1,\ldots,r$. The problem can be solved by the first phase of the simplex method. The joint probability distribution function of the components of ξ is logconcave. If $\vartheta_1\ge 1,\ldots,\vartheta_s\ge 1$, then the assertion follows from the fact that any linear combination of independent random variables having logconcave density is logconcave. If $\vartheta_i<1$ for some i's then the logconcavity of the joint distribution still holds but it needs a separate proof.

(9) The posynomial distribution. This distribution has support $\{z\mid 0\le z_i\le 1,\ i=1,\ldots,r\}$ and its distribution function is defined by

$$F(z_1,\ldots,z_r)=\frac{1}{\sum_{i=1}^{N} c_i z_1^{\alpha_{i1}}\ \cdots\ z_r^{\alpha_{ir}}},\quad 0<z_i\le 1,\ i=1,\ldots,r,$$

where $\alpha_{i1}\le 0,\ldots,\alpha_{ir}\le 0$, $\alpha_{i1}+,\ldots,+\alpha_{ir}<0$, and $c_i>0$, $i=1,\ldots,N$ are constants. Since the denominator is the sum of logconvex functions and logconvexity carries over for sums, it follows that $F(z_1,\ldots,z_r)$ is logconcave in the support of the distribution. There is no general rule, however, to decide that under what parameter values is F in fact a probability distribution function. It is easy to see that F is nondecreasing in each variable and is 0 if at least one variable is $-\infty$, and is 1, if all variables are $+\infty$. Thus, we only need to know that under what condition do we have

$$\frac{\partial^r F(z_1,\ldots,z_r)}{\partial z_1,\ \ldots,\ \partial z_r}\ge 0.$$

It is proved that if $r=2$ and $\alpha_{11} \le \alpha_{12} \le, \ldots, \le \alpha_{1r}$, $\alpha_{21} \le \alpha_{22} \le, \ldots, \le \alpha_{2r}$, then the above inequality is satisfied and $F(z_1, z_2)$ is a probability distribution function.

2.3 *Discrete distributions*

There is no discrete counterpart of the theory of logconcave and α-concave measures described in Section 2.1. Already the logconcavity and α-concavity has many forms in the literature. There is no problem, however, in the univariate case.

Definition 2.5. The discrete distribution $\{p_n\}$ defined on the integer lattice points of the real line is said to be logconcave if

$$p_n^2 \ge p_{n-1}p_{n+1}, \quad n = 0, \pm 1, \ldots.$$

It follows from this that if n, i, j are three integers such that $n = \lambda i + (1-\lambda)j$, where $0 < \lambda < 1$, then

$$p_n \ge p_i^{\lambda} p_j^{1-\lambda}. \tag{2.6}$$

The convolution theorem (Theorem 2.7) has a counterpart for univariate discrete distributions and is the following.

Theorem 2.12. *If $\{p_n\}$ and $\{q_n\}$ are two logconcave distributions on the lattice points of $\mathbb{R}$, then their convolution*

$$r_n = \sum_{k=-\infty}^{\infty} p_{n-k}q_k, \quad n = 0, \pm 1, \ldots$$

is also a logconcave distribution on the same lattice.

The above theorem implies that if $\{p_n\}$ is logconcave, then for any fixed k the sequence of nonnegative numbers

$$\sum_{i=n-k}^{n} p_i$$

is also logconcave (satisfies the inequality in Definition 2.5). This, in turn, implies that both sequences

$$F(n) = \sum_{i=-\infty}^{n} p_i, \quad 1 - F(n) = \sum_{i=n+1}^{\infty} p_i$$

are logconcave.

Examples for univariate logconcave distributions are the binomial, Poisson, geometric, hypergeometric, uniform and other known distributions.

Simple examples show that Theorem 2.12 does not carry over to the multivariate case. This is one indication why logconcavity and α-concavity properties of discrete distributions remain largely unexplored.

There are a few definitions to call a multivariate discrete distribution, defined on the lattice points, logconcave. One way is to write up (2.6) for vectors with integer components i, j, n. Similar definition is possible for α-concavity (see, e.g., Dentcheva et al. (2000)). However, there is no general theorem that would infer from logconcavity or α-concavity of the probability function to the same property of the distribution function, say.

Still, there is one definition of discrete logconcavity which enables us to enunciate some results.

Definition 2.6. The multivariate discrete distribution $p(k)$ defined on the lattice points of $\mathbb{R}^n$ is said to be logconcave if there exists a logconcave function $f(x)$, $x \in \mathbb{R}^n$ such that $p(k) = f(k)$ for any lattice point k of $\mathbb{R}^n$.

Our assertion concerns trinomial distributions on the nonnegative lattice points of $\mathbb{R}^2$. A triangular distribution is the distribution of a sum of independent bivariate random variables $\xi_1, \ldots, \xi_r$ where the support of each ξ_i is $\{(0,0), (0,1), (1,0)\}$ but their distributions may be different. The following theorem holds true.

Theorem 2.13. *Any triangular distribution is logconcave and the convolution of two triangular distributions is logconcave.*

The second assertion follows from the first one because the sum of two triangularly distributed random vectors is also triangularly distributed.

2.4 Applications to stochastic programming and other convexity statements

Theorem 2.5 and 2.11 give direct answers to the convexity questions that arise in connection with problems (1.2) and (1.4). Recall that a nonlinear programming problem is said to be convex if the set of feasible solutions is convex and the objective function to be minimized is convex (or to be maximized is concave).

Any logconcave function is quasi-concave, hence if $\xi \in \mathbb{R}^q$ has a continuous distribution and logconcave density or a density which is $-\frac{1}{q}$-concave then $h_0(x)$ in problem (1.2) is quasi-concave. Hence, $h_0(x)$ allows for the convex programming property of problem (1.2). If h is convex and we assume that $h_1, \ldots, h_m$ are quasi-concave, then the problem is in fact convex.

As regards problem (1.4), if $\xi \in \mathbb{R}^q$ has logconcave probability density function, then the objective function is logconcave. The function can be replaced by its logarithm, without changing the problem, and then it allows for a convex programming problem. To have a convex problem, the functions $h_1, \ldots, h_m$ have to be quasi-concave.

Problem (1.3) and the problem

$$\begin{aligned} &\max P(Tx \geq \xi) \\ &\text{subject to} \\ &Ax \geq b,\ x \geq 0 \end{aligned} \tag{2.7}$$

are special cases of problems (1.2) and (1.4), respectively, so the convexity of these problems can be derived from the above discussion.

It is noteworthy, however, that if $F(z) = P(\xi \leq z)$ is the distribution function of the random vector ξ which is assumed to have continuous distribution and logconcave density, then, by Theorem 2.3, $F(z)$ is a logconcave function. This implies that $F(Tx)$ is a logconcave function of $x \in \mathbb{R}^n$, and thus, the constraint

$$P(Tx \geq \xi) = F(Tx) \geq p$$

determines a convex set of x vectors for any fixed p. Also, the objective function of problem (2.7) is logconcave.

There are practical problems where the random vector ξ in problem (1.3) has a probability distribution which is a mixture of logconcave distributions. Since logconcavity does not carry over from terms to sums, we may need the stronger concavity property of the probability distributions involved, in order to obtain convex nonlinear programming problems. For the case of the normal distribution we have the following

Theorem 2.14. *Let $\Phi(z, R)$ be the n-variate standard normal distribution function. This function is concave on the set $\{z \mid z_i \geq \sqrt{n-1},\ i = 1, \ldots, n\}$.*

Some further results can be mentioned in connection with bivariate normal distributions. First we introduce a notion and mention a general theorem.

Definition 2.7. Let $F(z)$, $z \in \mathbb{R}^n$ be a probability distribution function, where the variables are partitioned as $z = (x^T, y^T)^T$. Suppose that x, y have k_1 and k_2 components, respectively, $1 \leq k_1,\ k_2 \leq n-1$. We say that $F(x, y)$ is

concave with respect to x in the positive direction on the convex set $E \subset \mathbb{R}^n$, if for any pair

$$z_1 = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} \in E, \quad z_2 = \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} \in E$$

for which $x_1 \leq x_2$, the function F is concave on the line segment connecting z_1 and z_2.

Theorem 2.15. *If $F(z)$ is concave in the positive direction in a closed n-dimensional interval E, with respect to any subset of its variables having at most $n-1$ elements, then F is quasi-concave in all variables in the same interval E.*

Let $\Phi(z_1, z_2; \varrho)$ be the bivariate standard normal probability distribution function. It is easy to see that

$$\frac{\partial^2 \Phi(z_1, z_2;\ \varrho)}{\partial z_1^2} = \Phi\left(\frac{z_2 - \varrho z_1}{\sqrt{1-\varrho^2}}\right) \frac{-\varrho}{\sqrt{1-\varrho^2}}\ \varphi(z_1) - \Phi\left(\frac{z_2 - \varrho z_1}{\sqrt{1-\varrho^2}}\right) z_1 \varphi(z_1)$$

and a similar formula holds for the second derivative with respect to z_2, where $\varphi(z)$ is the univariate standard normal probability density function. If $\varrho \geq 0$ and $E = \{z \mid z_1 \geq 0, z_2 \geq 0\}$, then $\Phi(z_1, z_2; \varrho)$ is concave on E with respect to any of the variables z_1, z_2. Hence, by Theorem 2.15, we have

Theorem 2.16. *If $\varrho \geq 0$, then $\Phi(z_1, z_2; \varrho)$ is concave in the positive direction and it is quasi-concave in E.*

For the case of the non-positive correlation we have

Theorem 2.17. *If $\varrho \leq 0$, then the function $\Phi(z_1, z_2; \varrho)$ is concave in each variable on*

$$E = \left\{ z \mid z_i \geq \sqrt{\frac{\varphi(1)}{2\Phi(1) + \varphi(1)}} = 0.346, \quad i = 1, 2 \right\}.$$

Also, the function is quasi-concave on E.

If $\varrho = 0$ or $\varrho = -1$, then $\Phi(z_1, z_2; \varrho)$ is easily seen to be a concave function on the nonnegative orthant.

The conditional expectation constraints (1.6) can easily be converted into equivalent linear constraints if the random variables $\xi_1,\ldots,\xi_r$ have (individually) continuous distributions and logconcave density functions. Let $r=1$ for the sake of simplicity. This case already captures the result in this respect. First we mention

Theorem 2.18. *If ξ is a univariate random variable that has continuous distribution and logconcave density function, then*

$$g(z) = E(\xi - z \mid \xi - z > 0) \tag{2.8}$$

is a decreasing function of z.

Proof. If $P(\xi - z > 0) = 0$, then the conditional expectation (2.8) is 0, by definition. Let $P(\xi - z > 0) > 0$. If $F(z) = P(\xi \le z)$ is the probability distribution function of ξ, then

$$g(z) = E(\xi - z \mid \xi - z > 0) = \frac{\displaystyle\int_z^\infty (1 - F(t))\,\mathrm{d}t}{1 - F(z)} = \frac{-1}{\frac{d}{dz}\log \displaystyle\int_z^\infty (1 - F(t))\,\mathrm{d}t}. \tag{2.9}$$

By Theorem 2.4, $1-F(t)$ is a logconcave function and the same theorem implies that

$$\int_z^\infty (1 - F(t))\,\mathrm{d}t$$

is also logconcave. This already implies that (2.9) is a decreasing function. □

The constraint

$$g(Tx) = E(\xi - Tx \mid \xi - Tx > 0) \le d$$

can be written in the equivalent form:

$$Tx \ge g^{-1}(d)$$

which is a linear one.

So far we have looked at stochastic programming problems where random variables appear only on the right hand sides of the constraints. Now we turn our attention to stochastic constraints where there are random variables also in the technology matrix. We state some results for the case where the random variables have joint normal or related distribution.

One of the most interesting case is connected with Kataoka's model that has important applications in finance, among others:

$$\begin{aligned}&\max d\\&\text{subject to}\\&P\left(\sum_{i=1}^{n}\xi_i x_i \geq d\right) \geq p\\&\sum_{i=1}^{n} x_i = M,\ x \geq 0\end{aligned} \tag{2.10}$$

We assume that $\xi = (\xi_1, \ldots, \xi_n)^T$ has an n-variate (nondegenerate or degenerate) normal distribution with

$$\mu_i = E(\xi_i), \quad i = 1, \ldots, n,\ \mu = (\mu_1, \ldots, \mu_n)^T$$
$$C = E[(\xi - \mu)(\xi - \mu)^T].$$

Problem (2.10) can be converted into a problem that turns out to be convex provided that $p \geq 1/2$. Note that p and M are constants and the decision variables are $x_1, \ldots, x_n, d$. Since $E(\xi x) = \mu^T x$, $Var(\xi^T x) = x^T C x$, if for an x we have $x^T C x > 0$, then

$$\begin{aligned}P(\xi^T x \geq d) &= P\left(\frac{(\xi - \mu)^T x}{\sqrt{x^T C x}} \geq \frac{d - \mu^T x}{\sqrt{x^T C x}}\right)\\&= 1 - \Phi\left(\frac{d - \mu^T x}{\sqrt{x^T C x}}\right),\end{aligned}$$

hence the probabilistic constraint in (2.10) is equivalent to

$$\mu^T x + \Phi^{-1}(1 - p)\sqrt{x^T C x} \geq d.$$

If, on the other hand, $x^T Cx = 0$, then the above constraint is also equivalent to the probabilistic constraint. This implies that (2.10) can be written in the equivalent form:

$$\max\left\{\mu^T x + \Phi^{-1}(1-p)\sqrt{x^T Cx}\right\}$$
$$\text{subject to}$$
$$\sum_{i=1}^{n} x_i = M, \quad x \geq 0. \tag{2.11}$$

Since C is a positive semidefinite matrix, the function $\sqrt{x^T Cx}$ is convex. On the other hand, $p \geq \frac{1}{2}$, hence $\Phi^{-1}(1-p) \leq 0$ and (2.11) turns out to be a convex programming problem.

Few convexity results are known for the case where randomness is in the technology matrix and the number of constraints is more than one. We take the stochastic constraints in the form $Tx \leq 0$, where T is an $r \times n$ random matrix and consider the set of $x \in \mathbb{R}^n$ vectors that satisfy the probabilistic constraint

$$P(Tx \leq 0) \geq p. \tag{2.12}$$

If the original stochastic constraint is of the form $Tx \leq \xi$, where ξ may be constant or random, then we introduce the new matrix $(T, -\xi)$ and the new decision vector $(x^T, x_{n+1})^T$. The set of x vectors satisfying

$$P(Tx \leq \xi) \geq p \tag{2.13}$$

is the same as those, satisfying

$$P\left((T, -\xi)\begin{pmatrix} x \\ x_{n+1} \end{pmatrix} \leq 0\right) \geq p$$
$$x_{n+1} = 1. \tag{2.14}$$

The constraint in the first line of (2.14) is already of the form (2.12) and the second constraint determines a convex set of the decision vectors. Hence, a statement for the constraint (2.12) can easily be carried over to the constraint (2.13). Let T_{i*} and T_{*j} designate the ith row and the jth column of T, respectively. Let further μ_{i*} and μ_{*j} designate the corresponding expectation vectors.

Theorem 2.19. *Suppose that the entries of T have a joint normal distribution and for the cross-covariance matrices of the columns of T we have that*

$$E[(T_{*j} - \mu_{*j})(T_{*k} - \mu_{*k})^T] = s_{jk}C,$$

where C is a fixed covariance matrix and the $s_{jk} = s_{kj}$, $j,k = 1,\ldots,n$ are constants, then the set of x vectors satisfying (2.12) is convex, provided that $p \geq \frac{1}{2}$.

If there is just one random column in T, then the above condition is clearly satisfied.

Theorem 2.20. *Suppose that the entries of T have a joint normal distribution and for the cross-covariance matrices of the rows of T we have that*

$$E[(T_{i*} - \mu_{i*})^T(T_{k*} - \mu_{k*})] = s_{ik}C,$$

where C is a fixed covariance matrix and the $s_{ik} = s_{ki}$, $i,k = 1,\ldots,r$, then the set of x vectors satisfying (2.12) is convex.

In our last convexity theorem we assume that the random entries in T are nonnegative and are lognormally distributed. While in Theorems 2.19 and 2.20 it is unimportant if we write up the stochastic constraint in the form $Tx \leq 0$ or $Tx \geq 0$, because the entries of T have joint normal distribution iff the same holds for their negatives, in the next theorem it is otherwise. The result is stated for the probabilistic constraint

$$G(x) = P(Tx \leq d) \geq p, \tag{2.15}$$

where d is a constant vector.

Theorem 2.21. *Suppose that the random entries $t_{i,k}$, $(i,k) \in I$, of the matrix T are positive valued and the joint probability distribution of $\alpha_{ik} = \log t_{ik}$, $(i,k) \in I$ is logconcave. Assume, for the sake of simplicity, that the random entries of T are in the first s columns and that all non-random entries of T in these columns are nonnegative.*

Under these conditions the function

$$G(e^{x_1},\ldots,e^{x_s},x_{s+1},\ldots,x_n)$$

is logconcave in $x \in \mathbb{R}^n$.

The above theorem tells us that if we replace e^{x_i} for x_i, $i = 1,\ldots,s$ in the probabilistic constraint (2.15), then the set of $x = (x_1,\ldots,x_s,x_{s+1},\ldots,x_n)^T$ vectors that satisfy this new form of the constraint is convex.

2.5 Bibliographical notes

The notion of a logconcave measure was introduced in Prékopa (1971), where also Theorem 2.1 and its simple consequences (Theorems 2.2–2.4) were also proved. Theorems 2.5 was proved in Prékopa (1972a, 1973c) for the case of concave $g_i(x, y)$, $i = 1, \ldots, r$. Tamm (1977) has observed that in the proof it is enough to assume that these functions are quasi-concave. Theorem 2.6 is from Davidovich et al. (1969) and Prékopa (1973b). Borell (1975) and Brascamp and Lieb (1976) introduced the α-concave measures (as generalizations of logconcave measures) and proved Theorem 2.9 and its special case Theorem 2.10. Theorem 2.11 can be proved in the same way as Theorem 2.5. It is mentioned in Prékopa (1995). Theorem 2.8 is a simple consequence of the logconcavity inequality written up for the probability measure associated with the random vector ξ. The logconcavity of the distributions in Examples (1)–(5) in Section 2.2 were shown in Prékopa (1971). Borell (1975) has shown the quasi-concavity of the probability distributions in Examples (6)–(7) of the same section. The gamma distribution in Example (8) is from Prékopa and Szántai (1978a). The posynomial distribution in Example (9) is from Prékopa (1988).

Theorem 2.12 is due to Fekete (see Fekete and Pólya (1912)). The notion of a discrete multivariate logconcave distribution (Definition 2.6) is due to Barndorff–Nielsen (1973). Theorem 2.13 is due to Pedersen (1975).

Theorem 2.14 was proved in Prékopa (2001). Theorems 2.15–2.17 are taken from Prékopa (1970). Theorem 2.18 has been known in reliability theory and actuarial science, where the value $g(z)$ is called expected residual lifetime and remaining life, respectively. In the stochastic programming context Prékopa (1973a) has mentioned it first.

The model (2.10) and its equivalent (2.11) was introduced, independently, by Kataoka (1963) and van de Panne and Popp (1963). The last mentioned paper applies it to an animal feed problem while the first one to finance. Theorems 2.19 and 2.21 are from Prékopa (1974). A special case of Theorem 2.20, assuming the rows of T to be stochastically independent, was proved in the same paper. Its present generality is due to Burkauskas (1986).

3 Numerical solution of probabilistic constrained stochastic programming problems

3.1 The case of continuously distributed random variable

We will consider the following special cases of problems (1.2) and (1.4):

$$\begin{aligned} &\min h(x) \\ &\text{subject to} \\ &G(x) = P(Tx \geq \xi) \geq p \\ &Ax = b,\ x \geq 0, \end{aligned} \tag{3.1}$$

$$\begin{aligned}
&\max G(x) = P(Tx \geq \xi)\\
&\text{subject to}\\
&Ax = b\\
&x \geq 0.
\end{aligned} \tag{3.2}$$

The methods that we describe in this section apply to more general problems too. However, we restrict ourselves to problems (3.1) and (3.2) because these are the problems which can be solved by existing codes.

There are two separate issues to solve these problems: (1) to adapt or create a nonlinear programming technique and (2) to adapt or develop a suitable method to compute the values and gradients (if necessary) of the function $G(x)$. First we concentrate on issue (1). However, we remark that if ξ has logconcave probability density function then both problems (3.1) and (3.2) are convex.

The method of feasible directions

Historically it was the first technique to solve problem (3.1). The method can be described as follows. Assuming that an initial feasible solution x has been found, the method works as follows. Let $x^0 = x$.

Step 1. Solve the following direction finding problem:

$$\begin{aligned}
&\text{Minimize } z\\
&\text{subject to}\\
&\nabla h(x^k)(x - x^k) - z \leq 0\\
&\nabla G(x^k)(x - x^k) + \Theta z \geq 0, \quad \text{if } G(x^k) = p\\
&Ax = b\\
&x \geq 0,
\end{aligned} \tag{3.3}$$

where Θ is a positive constant, fixed throughout the procedure. Let (z_{opt}, x^*) be an optimal solution of problem (3.3). If $z_{opt} = 0$ then x^* is an optimal solution of problem (3.1). If $z_{opt} > 0$ then go to Step 2.

Step 2. Solve the steplength determining problem:

$$\begin{aligned}
&\max \lambda\\
&\text{subject to}\\
&\lambda \geq 0 \quad \text{and}\\
&x^k + \lambda(x^* - x^k) \quad \text{is feasible.}
\end{aligned} \tag{3.4}$$

Go to Step 1.

The convergence of this procedure was proved by Prékopa (1970) under the following conditions: h, G are quasi-concave and have continuous gradients;

there exists an x^0 such that $G(x^0) > p$ (Slater's condition); the set $\{x \mid Ax = b, x \geq 0\}$ is bounded.

To find an initial feasible solution we may use a simple gradient method to maximize $G(x)$ subject to the constraints $Ax = b$, $x \geq 0$. It is needless to carry out the whole procedure, we may stop when an x^0 is encountered that satisfies $G(x^0) > p$.

The advantage of this method, when solving problem (3.1), is that it provides us with a possibility to handle the determination and use of $G(x)$ as well as $\nabla G(x)$ in a stable manner. The effect of noise when these values are approximated by simulation or bounding methods, can be controlled and, if the result is not satisfactory, the sample size can be increased or the bounds can be improved. This remark applies to all other solution methods too, that we present here concerning problem (3.1).

The Logarithmic Barrier Function Method (SUMT)

If $G(x)$ is a logconcave function in $\mathbb{R}^n$ then so is $G(x)-p$ on the set $\{x \mid G(x) \geq p\}$. This fact suggests the application of the Sequential Unconstrained Minimization Technique (SUMT) to solve problem (3.1).

The method works in such a way that we take a sequence of positive numbers $\{s^k\}$ such that $s^k > s^{k+1}$, $k = 0, 1, \ldots$, $\lim_{k\to\infty} s^k = 0$ and solve the problem

$$\begin{aligned} &\min \{h(x) - s^k \log(G(x) - p)\} \\ &\text{subject to} \\ &Ax = b \\ &x \geq 0, \end{aligned} \tag{3.5}$$

in principle for each k. If x^k is an optimal solution of (3.5) then, under some conditions, we have that

$$\lim_{k\to\infty} h(x^k) = \min_{\substack{G(x) \geq p \\ Ax = b,\ x \geq 0}} h(x). \tag{3.6}$$

The conditions are satisfied if h is a continuous, convex and G is a continuous logconcave function; $G(x^0) > p$ for some $x^0 \in \{x \mid Ax = b, \ x \geq 0\}$ and $\{x \mid Ax = b, \ x \geq 0\}$ is a bounded set. Under these conditions the objective function in problem (3.5) is convex for any s. For a more general convergence

theory of the SUMT method we refer to the book by Fiacco and McCormick (1968).

The supporting hyperplane method

It is assumed that h, G are quasi-concave and have continuous gradients; there exists a vector x^0 such that $G(x^0) > p$, $x^0 \in \{x \mid Ax = b, x \geq 0\}$; there exists a bounded convex polyhedron K^1 containing the set of feasible solutions of problem (3.1).

In a first phase we find a feasible x^0 satisfying Slater's condition. This can be done by maximizing $G(x)$ subject to $Ax \geq b$, $x \geq 0$, by use of the method described below. The second phase consists of the following steps.

Step 1. Solve the problem:

$$\begin{aligned} &\min h(x) \\ &\text{subject to} \\ &x \in K^s. \end{aligned} \tag{3.7}$$

Let x^s be an optimal solution. If x^s is feasible, then Stop, x^s is an optimal solution of problem (3.1). Otherwise, go to Step 2.

Step 2. Let λ^s be the largest $\lambda \geq 0$ such that $x^0 + \lambda(x^s - x^0)$ is feasible and

$$y^s = x^0 + \lambda^s(x^s - x^0).$$

Choose any constraint that is satisfied with equality sign for y^s. If it is $G(y^s) = p$, then we define

$$K^{s+1} = \{x \mid x \in K^s, \nabla G(y^s)(x - y^s) \geq 0\}.$$

If it is a linear constraint then K^{s+1} is defined as the intersection of K^s and the set determined by this linear constraint. Go to Step 1.

The reduced gradient method

Let $h(x) = c^T x$ in problem (3.1) that we intend to solve. We assume that $G(x)$ is logconcave, $\nabla G(x)$ is Lipschitz-continuous, Slater's condition holds and the set of feasible solutions is bounded.

In the kth iteration we are given a feasible x^k, a tolerance ε^k, the partitioning $x^k = (y^k, z^k)$, $A = (B, R)$, $c = (c_B, c_R)$, where B is a nonsingular square matrix and $(y^k)_j \geq 1$ for all j. We perform the following steps.

Step 1. Solve the direction finding problem

$$\begin{aligned}
&\min\ t\\
&\text{subject to}\\
&\boldsymbol{c}_B^T\boldsymbol{u} + \boldsymbol{c}_R^T v \le t\\
&\nabla G(x^k)\boldsymbol{u} + \nabla G(x^k)v \ge \Theta t, \quad \text{if } G(x^k) \le p + \varepsilon^k\\
&B\boldsymbol{u} + Rv = 0\\
&v_j \ge 0 \quad \text{if } z_j \le \varepsilon^k\\
&v_j \le 1 \quad \text{all } j,
\end{aligned} \tag{3.8}$$

where $\boldsymbol{u}$, v and t are the decision variables and Θ is a positive constant. An equivalent form of this problem is

$$\begin{aligned}
&\min\ t\\
&\text{subject to}\\
&\boldsymbol{r}^T v \le t\\
&\boldsymbol{s}^T v \ge \Theta t, \quad \text{if } G(x^k) \le p + \varepsilon^k\\
&v_j \ge 0 \quad \text{if } z_j \le \varepsilon^k\\
&v_j \le 1 \quad \text{all } j,
\end{aligned} \tag{3.9}$$

where $\boldsymbol{r} = c_R - c_B^T B^{-1} R$ and $\boldsymbol{s} = \nabla_z G(x^k) - \nabla_y G(x^k) B^{-1} R$ are the reduced gradients of the objective function and the probabilistic constraint, respectively. Let $(v^*, \boldsymbol{t}^*)$ be an optimal solution of problem (3.9). There are two cases.

Case 1. We have $t^* > \varepsilon^k$. Then we compute $\boldsymbol{u}^* = -B^{-1}R\boldsymbol{v}^*$ and go to Step 2.

Case 2. We have $t^* \le \varepsilon^k$. Then ε^k is halved. If the new ε^k is smaller than a zero tolerance then we accept x^k as optimal solution. Otherwise we solve problem (3.8) with the new ε^k and go to Step 2.

Step 2. Let λ_1 be the largest $\lambda \ge 0$ such that $x^k + \lambda \boldsymbol{w}^*$ satisfies the linear constraints, where $\boldsymbol{w}^* = (\boldsymbol{u}^*, v^*)$. Let λ_2 be any λ satisfying

$$p \le G(x^k + \lambda \boldsymbol{w}^*) \le p + \varepsilon^k$$

and $\lambda^* = \min(\lambda_1, \lambda_2)$. Define $x^{k+1} = x^k + \lambda^* \boldsymbol{w}^*$ and go to Step 3.

Step 3. If $(y^{k+1})_j \le \varepsilon^k$ for some j then the nondegeneracy assumption is violated. Find a new partition for which the nondegeneracy assumption holds. If necessary, reduce ε^k to meet this assumption. Go to Step 1.

A primal–dual method

A primal–dual method has been developed to solve the following problem

$$\begin{aligned} &\min \boldsymbol{c}^T\boldsymbol{x} \\ &\text{subject to} \\ &F(\boldsymbol{y}) \geq p \\ &T\boldsymbol{x} \geq \boldsymbol{y} \\ &D\boldsymbol{x} \geq \boldsymbol{d}, \end{aligned} \tag{3.10}$$

where F is the probability distribution function of the random vector $\boldsymbol{\xi}$: $F(\boldsymbol{y}) = P(\boldsymbol{\xi} \leq \boldsymbol{y})$. Problem (3.10) comes from the problem

$$\begin{aligned} &\min \boldsymbol{c}^T\boldsymbol{x} \\ &\text{subject to} \\ &P(T\boldsymbol{x} \geq \boldsymbol{\xi}) \geq \boldsymbol{p} \\ &D\boldsymbol{x} \geq \boldsymbol{d}, \end{aligned} \tag{3.11}$$

to which it is equivalent. Let $\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{y} \in \mathbb{R}^r$ and suppose F is a strictly logconcave probability distribution function, i.e., for every pair $\boldsymbol{y}_1, \boldsymbol{y}_2 \in \mathbb{R}^r$, $\boldsymbol{y}_1 \neq \boldsymbol{y}_2$, and $0 < \lambda < 1$, we have

$$F(\lambda\boldsymbol{y}_1 + (1-\lambda)\boldsymbol{y}_2) > [F(\boldsymbol{y}_1)]^{\lambda}[F(\boldsymbol{y}_2)]^{1-\lambda}.$$

The advantage of problem (3.10) over problem (3.11) is that the probabilistic constraint involves only the probability distribution function of $\boldsymbol{\xi}$ and not the composite function $F(T\boldsymbol{x})$.

Starting to solve problem (3.10) we associate with it a dual problem

$$\begin{aligned} &\max \left[\min_{F(\boldsymbol{y}) \geq p} \boldsymbol{u}^T\boldsymbol{y} + v^T\boldsymbol{d} \right] \\ &\text{subject to} \\ &T^T\boldsymbol{u} + D^T v = \boldsymbol{c} \\ &\boldsymbol{u} \geq 0,\ v \geq 0. \end{aligned} \tag{3.12}$$

The procedure works in the following manner. First we assume that a pair of vectors $(\boldsymbol{u}^{(1)}, v^{(1)})$ is available, for which

$$\left(\boldsymbol{u}^{(1)}, v^{(1)}\right) \in V = \left\{(\boldsymbol{u}, v) \mid T^T\boldsymbol{u} + D^T v = \boldsymbol{c},\ v \geq 0\right\}.$$

Suppose that $(\boldsymbol{u}^{(k)}, \boldsymbol{v}^{(k)})$ has already been chosen where $\boldsymbol{u}^{(k)} \geq 0$. Then we perform the following steps.

Step 1. Solve the problem

$$\min(\boldsymbol{u}^{(k)})^T \boldsymbol{y}$$

subject to

$$F(\boldsymbol{y}) \geq p. \tag{3.13}$$

Let $\boldsymbol{y}(\boldsymbol{u}^{(k)})$ designate the optimal solution. Then we solve the direction finding problem

$$\max\{\boldsymbol{u}^T \boldsymbol{y}(\boldsymbol{u}^{(k)}) + \boldsymbol{d}^T v\}$$

subject to

$$(u, v) \in V. \tag{3.14}$$

Let $(\boldsymbol{u}_k^*, \boldsymbol{v}_k^*)$ be an optimal solution to this problem. If $\boldsymbol{u}_k^* = \varrho \boldsymbol{u}^{(k)}$ then $(\boldsymbol{u}_k^*, \boldsymbol{v}_k^*)$ is an optimal solution of the dual problem (3.12) is an optimal solution to the primal problem (3.10), where $\hat{\boldsymbol{x}}$ is an optimal solution of the LP:

$$\min \boldsymbol{c}^T \boldsymbol{x}$$

subject to

$$T\boldsymbol{x} \geq \boldsymbol{y}(\boldsymbol{u}^{(k)})$$

$$D\boldsymbol{x} \geq \boldsymbol{d}.$$

Otherwise, go to Step 2.

Step 2. Find $\lambda^{(k)}$ $(0 < \lambda < 1)$ satisfying

$$\left(\boldsymbol{u}_k^*\right)^T \boldsymbol{y}\left(\frac{\lambda^{(k)}}{1-\lambda^{(k)}}\, \boldsymbol{u}^{(k)} + \boldsymbol{u}_k^*\right) > \left(\boldsymbol{u}^{(k)}\right)^T \boldsymbol{y}\left(\boldsymbol{u}^{(k)}\right) + \left(\boldsymbol{v}^{(k)}\right)^T \boldsymbol{d}.$$

Then we define

$$\boldsymbol{u}^{(k+1)} = \lambda^{(k)} \boldsymbol{u}^{(k)} + (1 - \lambda^{(k)}) \boldsymbol{u}_k^*$$

$$v^{(k+1)} = \lambda^{(k)} v^{(k)} + (1 - \lambda^{(k)}) v_k^*.$$

If the procedure is infinite, then the sequence $(\boldsymbol{u}^{(k)}, v^{(k)})$ converges and the limiting pair has the same property as (u_k^*, v_k^*) in Step 1.

A primal–dual interior point algorithm

If we look at problem (1.2), where $h(x)$ is assumed to be convex, $h_0(x)$ logconcave, and $h_1(x), \ldots, h_m(x)$ concave or logconcave, then the barrier function

$$h(x) - \mu \sum_{i=0}^{m} \log h_i(x) \tag{3.15}$$

is a convex function for any fixed $\mu > 0$.

Function (3.15) is the classical Fiacco McCormick logarithmic barrier function in connection with which usually three problems arise: the convexity of the function, finding an initial feasible solution, and the ill-conditioning of the Hessian matrix. In our case the barrier function is convex, under general assumptions, an initial feasible solution can frequently be found by using the probabilistic nature of the problem. If this is not the case, we would maximize $h_0(x)$ subject to all remaining constraints, until a point x, with $h_0(x) > 0$ is found.

We introduce slack variables and rewrite (1.2) as

$$\begin{aligned} &\min h(x) \\ &h_i(x) - w_i = 0 \\ &w_i \geq 0, \quad i = 0, 1, \ldots, m. \end{aligned}$$

The next step is to eliminate the inequalities $w_i \geq 0$ by the introduction of logarithmic barrier terms to the objective function. This yields the problem

$$\min \left[h(x) - \mu \sum_{i=0}^{m} \log w_i \right]$$

subject to

$$h_i(x) - w_i = 0, \quad w_i \geqq 0, \quad i = 0, 1, \ldots, m.$$

Then we take the Lagrangian

$$L(x, w, y, \mu) = h(x) - \mu \sum_{i=0}^{m} \log w_i - \sum_{i=0}^{m} y_i (h_i(x) - w_i),$$

write up the first order KKT conditions and the method is essentially an iterative solution method for these equations, based on Newton's method. The search directions Δx, Δy, Δz are determined by the system of linear equations

$$\begin{pmatrix} -H(x, y) & 0 & A^T(x) \\ 0 & -W^{-1}Y & -I \\ A(x) & -I & 0 \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta w \\ \Delta y \end{pmatrix} = \begin{pmatrix} \sigma \\ -\gamma \\ \varrho \end{pmatrix},$$

where

$$H(x, y) = \nabla^2 h(x) - \sum_{i=0}^{m} y_i \nabla^2 h_i(x)$$

$$A(x) = \begin{pmatrix} \nabla h_0(x) \\ \vdots \\ \nabla h_m(x) \end{pmatrix}$$

$$\sigma = \nabla h(x) - A^T(x)y,$$

W is the diagonal matrix with $w_0, \ldots, w_m$ in the main diagonal, $\gamma = \mu W^{-1}e - y$, $\varrho = (\varrho_0, \ldots, \varrho_m)^T$, $\varrho_i = w_i - h_i(x)$, $i = 0, \ldots, m$, and e is the vector with all components equal to one.

Starting from an initial x^0, w^0, y^0, we proceed through a sequence of points:

$$x^{k+1} = x^k + \alpha^k \Delta x^k,$$

$$w^{k+1} = w^k + \alpha^k \Delta w^k,$$

$$y^{k+1} = y^k + \alpha^k \Delta y^k,$$

where Δx^k, Δw^k and Δy^k are the subsequent search directions.

The above algorithm proved to be very efficient on a large number of problems.

In the next two methods we present not only problem solving algorithms but, simultaneously, the estimation of the probabilities involved.

Solution of the probabilistic constrained problem by the use of nonparametric estimates of distribution functions

We look at the problem

$$\begin{aligned} &\min c^T x \\ &\text{subject to} \\ &P(Tx \le \xi) \ge p \\ &Ax = b \\ &x \ge 0, \end{aligned} \tag{3.16}$$

where we assume that the random variables $\xi_1, \ldots, \xi_r$ are independent and each has continuous distribution. If we introduce the notations $F_i(z) = P(\xi_i \le z)$, $i = 1, \ldots, r$, $Tx = y$, then problem (3.16) has the equivalent form:

$$\begin{aligned} &\min c^T x \\ &\text{subject to} \\ &\prod_{i=1}^{r} (1 - F_i(y_i)) \ge p \\ &T_i x = y_i, \quad i = 1, \ldots, r \\ &Ax = b \\ &x \ge 0. \end{aligned} \tag{3.17}$$

The method works with the hazard rate functions defined by

$$g_i(t) = \frac{f_i(t)}{1 - F_i(t)} \tag{3.18}$$

if $F_i(t) < 1$ and $g_i(t) = 0$ if $F_i(t) = 1$, $i = 1, \ldots, r$, where $f_i(t)$ is the probability density function corresponding to $F_i(t)$, $i = 1, \ldots, r$.

Assume that the functions $f_i(t)$, $i = 1, \ldots, r$ are logconcave. Then, by Theorem 2.4 the functions $1 - F_i(t)$, $i = 1, \ldots, r$ are also logconcave. If, in addition, we assume that for any $x, y_1, \ldots, y_r$, satisfying the last three constraints in (3.17), we have $F_i(y_i) < 1$, $i = 1, \ldots, r$, then we can take logarithm on both sides in the first constraint.

The logconcavity of $1 - F_i(t)$ implies that $g_i(t)$ is a decreasing function. Integrating (3.18) we obtain

$$1 - F_i(y_i) = e^{-\int_{-\infty}^{y_i} g_i(t)\, \mathrm{d}t}.$$

We estimate the functions $g_i(t)$ from samples.

Let $g_i^{(n)}$ denote an original estimator of g_i for a given n. We take a sample $\{\xi_{ni}\}$ from the population with distribution function F_i, and create a grid $t_{n,1} < t_{n,2} < \ldots < t_{n,N}$. The grid may depend on i but we suppress it, for the sake of simplicity. The original estimator $g_i^{(n)}$ can then be defined as

$$g_i^{(n)}(t) = \frac{F_i^{(n)}(t_{n,j+1}) - F_i^{(n)}(t_{n,j})}{(t_{n,j+1} - t_{n,j})(1 - F_i^{(n)}(t_{n,j}))}, \quad t_{n,j} < x \le t_{n,j+1},$$

where $F_i^{(n)}$ is the empirical distribution function corresponding to F_i, $i = 1, \ldots, r$.

The next step is to choose a point $x_{n,j}$ from the window $(t_{n,j}, t_{n,j+1}]$ and assign a weight to it: $w(x_{n,j})$. Then solve the problem

$$\inf_{U_j \text{ increasing}} \sum_{j=1}^{N} \left(U_j - g_i^{(n)}(x_{n,j})\right)^2 w(x_{n,j}).$$

Let $\hat{g}_i^{(n)}(x_{n,j})$ be the optimal solution and assign this value to each element of the window $(t_{n,j}, t_{n+1,j}]$. Then $\hat{g}_i^{(n)}$ is a nondecreasing step function approximating $g_i^{(n)}$ in the least square sense. Now our approximation to $1 - F_i(y_i)$ is

$$1 - \hat{F}_i^{(n)}(y_i) = e^{-\int_{-\infty}^{y_i} \hat{g}_i^{(n)}(t)\, dt}.$$

This estimate has several good properties.

It remains to show how this estimate can be used to solve problem (3.17). Since

$$\log\left(1 - \hat{F}_i^{(n)}(y_i)\right) = -\int_{-\infty}^{y_i} \hat{g}_i^{(n)}(t)\, dt,$$

this function is piecewise linear and concave. Assume that the function consists of a finite number of linear pieces, the equations of which are

$$a_{ij}^T y + b_{ij}, \quad j = 1, \ldots, J_i, \quad i = 1, \ldots, r.$$

The problem (3.17) is equivalent to the following LP:

$$\min c^T x$$

subject to

$$y_i \geq a_{ij}^T y + b_{ij}, \quad j = 1, \ldots, J_i$$

$$T_i x = y_i, \quad i = 1, \ldots, r$$

$$Ax = b$$

$$x \geq 0.$$

Solution by a regression method

We solve problem (1.2) under the following conditions: the random vector ξ has continuous distribution and logconcave density; $g_1, \ldots, g_r$ are concave or quasi-concave functions; the other constraints are linear that we write in the form $Ax = b$, $x \geq 0$; the objective function is linear: $h(x) = c^T x$. The method works under more general assumptions as well but we restrict ourselves to a relatively simple case.

Under the above assumptions the constraining function

$$G(x) = P(g_1(x, \xi) \geq 0, \ldots, g_r(x, \xi) \geq 0) \tag{3.19}$$

is logconcave in x. We approximate $\log G(x)$ by a quadratic function $x^T Dx + b^T x + c$, where $x^T Dx$ is negative definite, solve the approximate problem, take a new feasible point to improve on the approximation, again solve the problem etc. The function values $G(x)$ computed for the approximation may be noisy, a least square approximation procedure will eliminate much of the noise. The solution algorithm can be described as follows.

Step 1. Find $x_0, \ldots, x_{k-1}$ which are feasible solutions to the problem. Compute the corresponding values $p_i = \log G_i(x)$, $i = 1, \ldots, k-1$.

Step 2. Let

$$q_k(x) = x^T D_k x + b_k^T x + c_k$$

be a quadratic function, where D_k, b_k, c_k are obtained from the least square approximation

$$\min \sum_{i=0}^{k-1} (p_i - q_k(x_i))^2.$$

Step 3. Solve the approximate problem

$$\min c^T x$$

subject to

$$x^T D_k x + b_k^T x + c_k \geq \log p$$

$$Ax = b$$

$$x \geq 0.$$

Let x_k be an optimal solution. Suppose that we have a criterion to decide if x_k is "good enough" as a solution to the original problem. In principle the KKT conditions are the best from this point of view but we may have some other stopping rule as well, e.g., the optimum values of the approximate problem do not change significantly in several subsequent iterations. Now, if x_k is "good enough", then stop and accept it as optimal solution to the original problem. Otherwise let $k \leftarrow k + 1$ and go to Step 2.

3.2 The case of discrete distribution

We look at problem (1.3), where we assume that the random vector ξ has discrete distribution. First we formulate the definition of a p-level efficient point (or PLEP). Let F designate the distribution function of ξ.

Definition 3.1. A point $z \in \mathbb{R}^r$ is said to be a p-level efficient point of the probability distribution F, if $F(z) \geq p$ and there is no y such that $y \leq z$, $y \neq z$, $F(y) \geq p$.

If $r = 1$, then for any $p \in (0, 1)$ there exists exactly one p-level efficient point. Sometimes we need p-level efficient points defined in connection with functions obtained from a distribution function by holding some of the variables fixed. This is the case, e.g., in the forthcoming enumeration algorithm. Definition 3.1 extends in a trivial way to this case. We have the following

Theorem 3.1. *If the components of the random vector ξ are integer-valued, then for any $p \in$ (0,1) the set of p-level efficient points is nonempty and finite. The set of p-level efficient points serves as the p-quantile of the probability distribution determined by F.*

From the practical point of view we may restrict ourselves to the case where the distribution has a finite support. We do not assume, however, that the support set is part of the integer lattice in $\mathbb{R}^r$.

Let $Z_i = \{z_{i0}, \ldots, z_{ik_i+1}\}$ be the set of possible values of ξ_i, $i = 1, \ldots, r$ and define the direct product

$$Z = Z_1 \times \ \ldots \ \times Z_r. \tag{3.20}$$

The set Z contains all possible values of ξ but may be somewhat larger than the set of possible values. Below we present an algorithm to find all p-level-efficient points of the distribution. We remark that any p-level-efficient point is necessarily an element of Z.

Algorithm to enumerate the p-efficient points
Step 0. Initialize $k \leftarrow 0$. Go to Step 1.
Step 1. Let

$$\begin{aligned} z_{1,j_1} &= \arg\min\{y \mid F(y, z_{2,k_2+1}, \ldots, z_{r,k_r+1}) \geq p\} \\ z_{2,j_2} &= \arg\min\{y \mid F(z_{1,j_1}, y, \ldots, z_{r,k_r+1}) \geq p\} \\ &\vdots \\ z_{r,j_r} &= \arg\min\{y \mid F(z_{1,j_1}, \ldots, z_{r-1,j_{r-1}}, y) \geq p\}. \end{aligned}$$

Go to step 2
Step 2. Let $E \leftarrow \{z_{1,j_1}, \ldots, z_{r,j_r}\}$. Go to Step 3.
Step 3. Let $k \leftarrow k + 1$. If $j_1 + k > k_1 + 1$, then go to Step 5. If $j_1 + k \leq k_1 + 1$, then go to Step 4.
Step 4. Enumerate all p-level-efficient points of the function $F(z_{1,j_1+k}, y)$, where $y \in \mathbb{R}^{r-1}$ and eliminate those which dominate at least one element in E. (y dominates z, if $y \geq z$ and $y \neq z$). If H is the set of the remaining p-level-efficient points, which may be empty, then let $E \leftarrow E \cup H$. Go to Step 3.
Step 5. Stop. E is the set of p-level-efficient points of the distribution function F.

Remark. The above algorithm allows for the enumeration of all p-level-efficient points of functions F which assign probability ≤ 1 to the entire space.

Example. Let $r = 2$, $Z_1 = Z_2 = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$, $p_{ik} = 0.019$ if $0 \leq i \leq 4, 0 \leq k \leq 5$ or $k = 8, 9$; $p_{ik} = 0.038$ if $0 \leq i \leq 4, k = 6$; $p_{ik} = 0$ if $0 \leq i \leq 4, k = 7$; $p_{ik} = 0.001$ if $5 \leq i \leq 9, 0 \leq k \leq 9$ and $p = 0.6$. In Step 1 we obtain

$$\begin{aligned} 3 &= \arg\min\{y \mid F(y, 9) \geq 0.6\} \\ 6 &= \arg\min\{y \mid F(3, y) \geq 0.6\}. \end{aligned}$$

Thus, $(z_{1,j_1}, z_{2,j_2}) = (3, 6)$ and at Step 2 we have $k = 0$, $E = \{(3, 6)\}$. In Step 3 we take $k = 4$ and go to Step 4, where we obtain $H = \{(4, 6)\}$. We eliminate $(4, 6)$ and $E = \{(3, 6)\}$. Now we go to Step 3 and find H empty for $k = 2, 3, 4, 5$. In case of $k = 6$ we obtain $H = \{(9, 5)\}$ and the algorithm terminates. The result is $E = \{(3, 6), (9, 5)\}$.

Cutting plane method for the solution of problem (1.3) with discrete random vector ξ

We assume that the support of ξ is Z, given by (3.20), or a part of it. Let $z^{(1)}, \ldots, z^{(N)}$ designate the p-level-efficient points of the distribution F and suppose that we have already enumerated them. Problem (1.3) is equivalent to the disjunctive programming problem:

$$\begin{aligned} &\min c^T x \\ &\text{subject to} \\ &Tx \geq z^{(i)} \quad \text{for at least one } i = 1, \ldots, N \\ &Ax = b,\ x \geq 0. \end{aligned} \tag{3.21}$$

Problem (3.21) is relaxed as

$$\begin{aligned} &\min c^T x \\ &\text{subject to} \\ &Tx \geq \sum_{i=1}^{N} \lambda_i z^{(i)} \\ &Ax = b,\ x \geq 0 \\ &\sum_{i=1}^{N} \lambda_i = 1,\ \lambda \geq 0 \end{aligned} \tag{3.22}$$

and we propose an algorithm to solve (3.22). Introducing slack variables the problem becomes

$$\begin{aligned} &\min c^T x \\ &\text{subject to} \\ &Tx - u = \sum_{i=1}^{N} \lambda_i z^{(i)} \\ &Ax = b,\ x \geq 0 \\ &\sum_{i=1}^{N} \lambda_i = 1, \quad \lambda \geq 0,\ u \geq 0. \end{aligned} \tag{3.23}$$

Since the set of p-level efficient points may be concentrated on a manifold with dimension smaller than r, first we determine this manifold. Let

$$\overline{z} = \frac{1}{N} \sum_{i=1}^{N} z^{(i)}$$

and consider the system of linear equations with unknown vector w:

$$w^T(z^{(i)} - \overline{z}) = 0, \quad i = 1, \ldots, N. \tag{3.24}$$

If $w_1, \ldots, w_h$ is a maximum number of linearly independent vectors satisfying (3.24), then we append the constraints

$$w_l^T(Tx - u) = 0, \quad l = 1, \ldots, h$$

to the constraints $Ax = b$ and keep them together throughout the procedure. The steps of the algorithm are the following.

Step 1. Enumerate all the p-level efficient points $z^{(1)}, \ldots, z^{(N)}$. Initialize $k \leftarrow 0$ and go to Step 2.

Step 2. Solve the LP:

$$\begin{aligned} &\min c^T x \\ &\text{subject to} \\ &Ax = b, \quad w_l^T(Tx - u) = 0, \quad l = 1, \ldots, h \\ &(w^i)^T(Tx - u - \overline{z}) \geq 0, \quad i = 1, \ldots, k \\ &x \geq 0, u \geq 0. \end{aligned} \tag{3.25}$$

If $k = 0$ then ignore the cuts, i.e., the constraints in the second to the last line in (3.25). Let $(x^{(k)}, u^{(k)})$ be an optimal solution. Go to Step 3.

Step 3. Solve the auxiliary problem

$$\begin{aligned} &\min e^T \mu = \alpha \\ &\text{subject to} \\ &\sum_{i=1}^{N} (z^{(i)} - \overline{z})\mu_i = Tx^{(k)} - u^{(k)} - \overline{z} \\ &\mu \geq 0, \end{aligned} \tag{3.26}$$

where $e = (1, \ldots, 1)^T$ and the decision vector is μ. The solution of (3.26) needs some care because the matrix of the problem may not have full

row rank. In this case we may use a more general variant of the simplex method (see Prékopa (1996)) or use a well-known technique involving artificial variables.

If $\alpha \leq 1$, then Stop, the current $(x^{(k)}, u^{(k)})$ is an optimal solution of problem (3.25). If $\alpha > 1$ then go to Step 4.

Step 4. Let $z^{(i_1)} - \overline{z}, \ldots, z^{(i_r-h)} - \overline{z}$ be an optimal basis to the problem (3.26). Then find a w satisfying

$$w^T w_i = 0, \quad i = 1, \ldots, h$$
$$w^T (z^{(i_j)} - z^{(i_1)}) = 0, \quad j = 2, \ldots, r - h.$$

These $r - 1$ equations determine w up to a constant factor. Assume that we have determined w in such a way that

$$w^T \left(T x^{(k)} - u^{(k)} - \overline{z} \right) < 0.$$

Then define $w^{k+1} = w$ and introduce the cut

$$\left(w^{k+1} \right)^T (Tx - u - \overline{z}) \geq 0.$$

Set $k \leftarrow k + 1$ and go to Step 2.

The finiteness of the above algorithm is guaranteed if $\{z^{(i)}, i = 1, \ldots, N\}$ is a discrete convex set, i.e., $z^{(j)} \notin \text{riconv}\{z^{(i)}, i = 1, \ldots, N\}, j = 1, \ldots, N$.

The cone generation method

The method solves problem (3.22). Before presenting it we introduce the notion of a discrete α-concave function. In Section 2 we already introduced this notion in connection with functions defined on the entire r-space and stated in Theorem 2.9 that if the probability measure is generated by an α-concave density function, then the measure is γ-concave in the sense of Definition 2.4, where $\gamma = \alpha/(1 + r\alpha)$. As we have remarked in Section 2, there is no discrete counterpart of this theorem. Still, the notion of α-concavity can be defined in connection with a function defined on the integer lattice of the r-space.

Definition 3.2. The function $f(z) \geq 0$, defined on the integer lattice of the r-space is said to be α-concave if for any integer component vectors x, y, z and $0 < \lambda < 1$ such that $f(x) > 0, f(y) > 0, z \geq \lambda x + (1 - \lambda)y$, we have the inequality

$$f(\lambda x + (1 - \lambda)y) \geq [\lambda f^{\alpha}(x) + (1 - \lambda) f^{\alpha}(y)]^{1/\alpha}, \tag{3.27}$$

where $-\infty \leq \alpha < \infty$; for $\alpha = -\infty$, $\alpha = 0$ the expression in (3.27) is defined by continuity in the same way as we have defined it in Definition 2.3.

If a function is α-concave in the entire r-space, then it is α-concave on the integer lattice of the same space.

Assume that the probability distribution function $F(z) = P(\xi \leq z)$, $z \in \mathbb{R}^r$ is α-concave on the integer lattice of the space. Let

$$Z_p = \{z \in \mathbb{R}^r \mid F(z) \geq p\}.$$

Then we have the relation (for the proof see Dentcheva et al. (2000))

$$Z_p \cap Z_+^r = \text{conv}\,(Z_p) \cap Z_+^r, \tag{3.28}$$

where Z_+^r is the nonnegative orthant of the r-space, $Z_+^r = \{z \in \mathbb{Z}^r \mid z \geq 0\}$.

Relation (3.28) gives some information regarding the relationship of problems (3.21) and (3.22). Equivalence between the two problems, however, can be stated only if further information is available. If, for example x is also restricted to be integer and the matrix T has integer entries, then, by (3.28), the two problems are equivalent.

The cone generation method consists of the following steps.

Step 1. Find a p-level efficient point $z^{(1)}$ and set $I_1 = \{1\}$, $k = 1$.

Step 2. Solve the master problem

$$\begin{aligned} &\min c^T x \\ &\text{subject to} \\ &Tx \geq \sum_{i \in I_k} \lambda_i z^{(i)} \\ &Ax = b,\ x \geq 0 \\ &\sum_{i \in I_k} \lambda_i = 1,\ \lambda \geq 0. \end{aligned} \tag{3.29}$$

Let u^k be the part of the optimal dual vector which is associated with the first part of the constraints.

Step 3. Calculate an upper bound on the value

$$\min_{i \in I_k} (u^k)^T z^{(i)}. \tag{3.30}$$

If for $j_k \in I_k$ we have $\lambda_{j_k} > 0$ then $(u^k)^T z^{j_k}$ is a suitable upper bound on this value.

Step 4. Find a p-level efficient solution to the problem

$$\min_{F(z) \geq p} \left(u^k\right)^T z, \tag{3.31}$$

where $F(z) = P(\xi \leq z)$, $z \in \mathbb{R}^r$. Let $z^{(k+1)}$ designate an optimal solution to this problem. If the optimal values in (3.31) and (3.30) coincide, then Stop, the optimal solution of problem (3.29) is an optimal solution to problem (3.22). Otherwise set $I_{k+1} \leftarrow I_k \cup \{k+1\}$, $k \leftarrow k+1$ and go to Step 2.

The solution of problem (3.31) may be difficult in the general case. However, if the random variables $\xi_1, \ldots, \xi_r$ are independent, there is a simple way to do the job. We assume that Z is the integer grid of $\mathbb{R}^r$. Let $F_i(z)$ designate the probability distribution function of the random variable ξ_i, $i = 1, \ldots, r$. Then the probabilistic constraint $F(z) \geq p$ can be written in the form

$$\log F(z) = \sum_{i=1}^{r} \log F_i(z_i) \geq \log p.$$

If $F(z) \geq p$, then we also have $F_i(z_i) \geq p$, $i = 1, \ldots, r$. This implies that if l_i is the p-level efficient point of the distribution function F_i, then $z_i \geq l_i$, $i = 1, \ldots, r$. It follows that problem (3.31) is equivalent to the nonlinear knapsack problem:

$$\begin{aligned}
&\min \sum_{i=1}^{r} u_i z_i \\
&\text{subject to} \\
&\sum_{i=1}^{r} \log F_i(z_i) \geq \log p \\
&z_i \geq l_i, \quad z_i \quad \text{integer}, \quad i = 1, \ldots, r.
\end{aligned} \tag{3.32}$$

If b_i is a known upper bound on z_i, $i = 1, \ldots, r$, then problem (3.32) can be transformed into the following equivalent 0–1 variable LP:

$$\begin{aligned}
&\min \sum_{i=1}^{r} \sum_{j=l_i}^{b_i} j u_i y_{i,j} \\
&\text{subject to} \\
&\sum_{i=1}^{r} \sum_{j=l_i}^{b_i} (\log F_i(j)) y_{i,j} \geq \log p \\
&\sum_{j=l_i}^{b_i} y_{i,j} = 1, \quad i = 1, \ldots, r \\
&y_{i,j} \in \{0, 1\}, \quad j = l_i, \ldots, b_i, \quad i = 1, \ldots, r.
\end{aligned} \tag{3.33}$$

The variable z_i in problem (3.32) is replaced by $\sum_{j=l_i}^{b_i} j\, y_{i,j}$ in problem (3.33). If $F_i(z)$ is the distribution function of a logconcave distribution on the integers

$i = 1, \ldots, r$, then we introduce $z_i = l_i + \sum_{j=l_i+1}^{b_i} \delta_{ij}$, $i = 1, \ldots, r$, where δ_{ij} are 0−1 variables. Then we reformulate problem (3.32) as follows:

$$\begin{aligned}
&\min \sum_{i=1}^{r} \sum_{j=l_i+1}^{b_i} u_i \delta_{ij} \\
&\text{subject to} \\
&\sum_{i=1}^{r} \sum_{j=l_i+1}^{b_i} a_{ij} \delta_{ij} \geq q, \\
&\delta_{ij} \in \{0, 1\}, \quad j = l_i + 1, \ldots, b_i, \; i = 1, \ldots, r,
\end{aligned} \tag{3.34}$$

where $a_{ij} = \log F_i(j) - \log F_i(j-1) = P(\xi = j)$ and $q = \log p - \log l$, $l = (l_1, \ldots l_r)^T$. Problem (3.34) is a knapsack problem, for which many efficient solution techniques exist.

The advantage of the cone generation method is that problem (3.31) is separated from the other parts of the problem and, as we have seen, this problem sometimes has simple solution technique.

A branch and bound method

The method solves problem (1.3) with the additional restriction that x and ξ are integers. The matrix T is assumed to have integer entries. As we have remarked in the discussion of the previous method, under these conditions the problem

$$\begin{aligned}
&\min c^T x \\
&\text{subject to} \\
&Tx \geq z^{(i)}, \quad \text{for at least one} \quad i = 1, \ldots, n \\
&Ax = b, \; x \geq 0 \text{ integer},
\end{aligned} \tag{3.35}$$

which is the same as problem (3.1) with the integrality restriction on x, is the same as the problem

$$\begin{aligned}
&\min c^T x \\
&\text{subject to} \\
&Tx \geq \sum_{i=1}^{N} \lambda_i z^{(i)} \\
&Ax = b, \quad x \geq 0 \text{ integer} \\
&\sum_{i=1}^{N} \lambda_i = 1, \; \lambda \geq 0.
\end{aligned} \tag{3.36}$$

A new PLEP generation technique is also proposed by the authors. In order to describe it we need some preparation.

In connection with a vector w satisfying $F(w) \geq p$ for some $p \in (0, 1)$ we define $l_i = l_i(w)$ to be the $p/F(w)$-level efficient point of the conditional marginal distribution $F_i(z_i \mid \xi \leq w)$, i.e.,

$$l_i(w) = \arg\min\left\{j \mid F_i(j \mid \xi \leq w) \geq \frac{p}{F(w)}\right\}, \quad i = 1, \ldots, r.$$

Let $l = l(w)$ designate the vector of components $l_i(w)$, $i = 1, \ldots, r$. The following assertions hold true:

(i) for every p-level efficient point $v \leq w$ we have $v \geq l(w)$;
(ii) if $z \leq w$, then $l(z) \geq l(w)$;
(iii) w is p-level efficient if and only if $l(w) = w$.

To every nonnegative integer component vector z we assign a level $|z| = z_1 + \ldots + z_r$. We also create a graph out of these points as nodes and draw a directed arc from v to w if $|w| = |v| + 1$ and the two vectors differ in one component only. There are two variants of the enumeration technique: the forward and the backward schemes. In the latter one we start at the highest level candidate, in the former one at the lowest level candidate. Since $p \approx 1$, the backward algorithm produces the result faster, in general, therefore we present only that one.

Backward enumeration scheme

Step 0. Let $v = (k_1, \ldots, k_r)$ and set $k = |v|_1$, level counter; $S^k = \emptyset$, the set of PLEP's at level k; $C^k = \{v\}$, the set of candidate points at level k; $J = \emptyset$, the set of all PLEP's.

Step 1. For each $v \in C^k$ generate its predecessors and supplement them to C^{k-1}.

Step 2. For each $v \in C^{k-1}$ compute $l(v)$. If $l(v) = v$, move the point from C^{k-1} to S^{k-1}.

Step 3. Set $J = J \cup S^{K-1}$. If $C^{k-1} = \emptyset$, Stop. Otherwise decrease k by 1 and go to Step 1.

In the implementation an important issue is to avoid generating the same point more than once. Granted, the procedure terminates in a finite number of iterations.

In the solution algorithm of the problem we use the value $z(l^*(v))$ defined by

$$\begin{aligned} &z(l^*(v)) = \min c^T x \\ &\text{subject to} \\ &Tx \geq l^*(v) \\ &Ax = b,\ x \geq 0 \text{ integer}, \end{aligned} \tag{3.37}$$

where $l^*(v)$ designates the smallest vector in Z such that $l^*(v) \geq l(v)$. The value $z(l^*(v))$ is a lower bound on the optimal values of all predecessors of v.

Let z_I designate the best known feasible solution to the problem, at some point of the algorithm. Then, if $z(l^*(v)) \geq z_I$, node v can be discarded. If $z(l^*(v)) < z_I$, then v is either stored as candidate for further use (if the solution is fractional), or replaces z_I (if the solution is integer).

The branch and bound algorithm

Step 0. Let $v = (k_1, \ldots, k_r)$. Compute an initial upper bound z_I and set $k = |v|_1$, level counter; $S^k = \emptyset$, set of PLEP's at level k; $C^K = \{v\}$, set of candidate points at level k; $J = \emptyset$, set of all PLEP's; $M = \emptyset$, set of PLEP's corresponding to integer problems.

Step 1. For each $v \in C^k$ generate the predecessors and supplement them to C^{k-1}.

Step 2. For each $v \in C^{k-1}$ compute the conditional lower bound $l(v)$ on PLEP's. If $l(v) = v$, move the point from C^{k-1} to S^{k-1}.

Step 3. Set $J = J \cup S^{k-1}$. If $C^{k-1} = \emptyset$ and $S^{k-1} = \emptyset$, Stop. Otherwise decrease k by 1.

Step 4. For each point $v \in S^k \cup C^k$ solve a relaxation of problem (3.37). Let $x(l^*(v))$ be an optimal solution.

(1) If $z(l^*(v)) \geq z_I$, discard the point v.

(2) If $z(l^*(v)) < z_I$ and $v \in S^k$, then (I) if $x(l^*(v))$ is fractional, then supplement the point v to M; (II) otherwise update z_I and remove from M all points w having worse lower bound values $z(l^*(v))$.

If $C^k = \emptyset$, Stop. Otherwise go to Step 1.

The algorithm terminates in a finite number of steps. At the end M contains all PLEP's corresponding to the integer problem.

As a special case of the above described probabilistic integer programming problem the probabilistic set covering problem has been considered and solved, where the components of the decision vector x are 0–1 variables and the random vector ξ has 0–1 components as well. The special structure of the problem is exploited in the solution algorithm.

Finally, we mention that an algebraic geometry approach has been developed for the solution of the probabilistic constrained integer programming problem, where the random variables are also integer valued and are located in the technology matrix. The method uses Gröbner bases. The description of it is, however, rather lengthy therefore we disregard its presentation here.

3.3 Bibliographical notes

The method of feasible direction is due to Zoutendijk (1960). Its application (algorithm P2) to solve probabilistic constrained stochastic programming problems was proposed by Prékopa (1970). Deák (1971) implemented it together with the calculation of multivariate normal distribution function values and

their gradients. Prékopa et al. (1980) applied the developed method for the solution of an economic problem concerning electrical energy. Problem (1.5) with normal distribution was also efficiently solved by Szántai (1985). The use of SUMT (see Fiacco and McCormick (1968) for its general description) for the same problem was suggested by Prékopa (1972a) and implemented by Rapcsák (1974). It was first applied (in the probabilistic constrained framework) by Prékopa et al. (1978). Further application is mentioned in Section 7. The application of a variant of the supporting hyperplane method of Veinott (1967), to solve problem (1.3) was proposed by Prékopa and Szántai (1978b). The solution code, due to Szántai, is described in his (Szántai, 1988) paper. The reduced gradient method, developed by Abadie and Carpentier (1969), was applied by Mayer (1979, 1980, 1988, 2000) to solve problem (1.3). See also his summarizing paper (1992) (Mayer (1992)) and book (1998) (Mayer (1998)) . The paper by Kall and Mayer (1996) presents, among others, the solution of (1.3), by the use of the reduced gradient method, in a general model management framework. The next, primal-dual, method is due to Komáromi (1986b, 1987).

The primal-dual interior point algorithm is due to Vanderbei and Shanno (2000). The ill-conditioning in the Hessian can be balanced by a method due to Nash and Sofer (1993). Other refinement is due to the same authors (1998) (Nash and Sofer, 1998). It addresses the problem of the infeasibility of a returned μ.

The method, using nonparametric estimates of the distribution functions, when in problem (1.3) the random vector ξ has independent components, is due to Gröwe (1995, 1997). The regression method is due to Deák (2000). Another approximation method is presented in Salinetti (1983).

The concept of a p-level efficient point PLEP was introduced by Prékopa (1990a). It is also termed p-efficient point and sometimes abbreviated as PEP.

The algorithm to enumerate the p-level efficient points and the subsequent cutting plane method is due to Prékopa, Vízvári and Badics (1998). The cone generation method, together with the embedded enumeration of the p-level efficient points is due to Dentcheva, Prékopa and Ruszczyński (2000). Vízvári (2002) has revised it from the point of view of integer programming. The next, branch and bound method and the embedded enumeration technique of the p-level efficient points is due to Beraldi and Ruszczyński (2001). They have another, paper (2002), where they solve a stochastic set covering problem by a similar method.

The algebraic-geometry method, using Gröbner bases (see Becker and Weispfenning), is due to Tayur et al. (1995). Inequalities in connection with probabilistic constrained problems with discrete random variables are presented in Sen (1992).

4 Dynamic type stochastic programming problems with probabilistic constraints

The simplest dynamic type stochastic programming problem is the two-stage programming under uncertainty, or, stochastic programming

with recourse. This can be formulated as

$$\begin{aligned} &\min\{c^T x + E(q(x, \xi))\} \\ &\text{subject to} \\ &Ax = b,\ x \geq 0, \\ &x \in K, \end{aligned} \tag{4.1}$$

where

$$\begin{aligned} &q(x, \xi) = \min\ q^T y \\ &\text{subject to} \\ &Wy \geq \xi - Tx,\ y \geq 0. \end{aligned} \tag{4.2}$$

Problems (4.1) and (4.2) are called the first and the second stage problems, respectively. The set K in problem (4.1) is the set of all x vectors for which problem (4.2) has feasible solution for any possible values of the random vector ξ. Since the projection of the convex polyhedron

$$\{(x, \xi, y) \mid Wy \geq \xi - Tx,\ y \geq 0\}$$

onto the space of the x, ξ vectors can be described by the (in x and ξ) homogeneous linear inequalities

$$Hx \geq G\xi, \tag{4.3}$$

it follows that

$$K = \{x \mid Hx \geq h\}, \tag{4.4}$$

where

$$h_i = \sup_{\xi \in \Xi} (G\xi)_i$$

and Ξ is the support of the random vector ξ. By (4.4), K is a convex polyhedron.

If we assume that the dual of problem (4.2) has feasible solution and $E(\xi)$ exists, then the optimum value of (4.2) exists for any $x \in K, \xi \in \Xi$ and $E(q(x, \xi))$ exists for any $x \in K$.

The condition that the second stage problem be solvable for any $\xi \in \Xi$ is frequently too restrictive in practice. In the power system engineering, for example, we cannot design power systems in such a way that no blackout should occur, ever. Similarly, in water resources engineering we cannot design a dam in such a way that no flood should occur, ever. We have to allow

disaster to occur but we may limit the frequency of its occurrence. A model where the solvability of the second stage problem is ensured only by a (large) probability has been formulated by Prékopa (1973a). Since (4.3) is a necessary and sufficient condition for the solvability of problem (4.2), we formulate our new problem in such a way that impose a probabilistic constraint on the inequalities (4.3). This raises the question that what optimum value shall we enter into problem (4.1) if problem (4.2) is not always solvable. To overcome this difficulty we introduce new variables (components of) z into problem (4.2) that we add to the left hand side of the constraints and enter the same variables with high costs into the objective function. The high costs should render $z = 0$ automatically whenever the original second stage problem (4.2) is solvable. Our new two-stage problem with probabilistic constraint is the following:

$$\begin{aligned} &\min\{c^T x + E(q(x, \xi))\} \\ &\text{subject to} \\ &P(Hx \geq G\xi) \geq p \\ &Ax = b,\ x \geq 0, \end{aligned} \tag{4.5}$$

where

$$\begin{aligned} &q(x, \xi) = \min\{q^T y + d^T z\} \\ &\text{subject to} \\ &Wy + z \geq \xi - Tx,\ y \geq 0. \end{aligned} \tag{4.6}$$

Prékopa (1980b, 1995) formulated the power system expansion problem as a special case of problem (4.5)–(4.6). A solution technique for the above problem is proposed by Deák (2001). Similar models can be formulated for the multiperiod case. A practical way, however, to incorporate probabilistic constraints into dynamic type models is to include them in rolling horizon models, where we solve static models subsequently in time. Examples will be presented in Section 7.

5 Bounding, approximation and simulation of probabilities

In Section 3 we presented nonlinear programming procedures suitable to solve probabilistic constrained stochastic programming problems. We left open the question how to compute the constraining function values and their gradients in the probabilistic constraint. First we look at the function values which are joint probabilities of finite numbers of random events, where each event is determined by some relation involving multivariate functions. For example the constraining function $P(Tx \geq \xi)$ in the probabilistic constraint is the joint probability of r random events: $T_i x \geq \xi_i$, $i = 1, \ldots, r$,

for every fixed x. If r is large, then we may not expect that the joint probability of these events can be computed, therefore we look for bounding, approximation and simulation procedures.

The combined use of simulation and optimization appeared in the probabilistic constrained stochastic programming publications as early as 1974 (see Prékopa et al. (1980)). More recent is the use of bounding techniques which, among other methods, serve for approximation of probabilities.

First we look at the bounding techniques because some of them are used also in the approximation and simulation procedures.

5.1 *Bounding probabilities of boolean functions of events*

Let $A_1, \ldots, A_r$ be events in an arbitrary probability space. We intend to give lower and upper bounds for some Boolean functions of them. We are primarily interested in the union $\cup_{i=1}^r A_i$ and the intersection $\cap_{i=1}^r A_i$ of the events as their Boolean functions. The intersection of r events appears in the stochastic constraint $Tx \geq \xi$, where $A_i = \{T_i x \geq \xi_i\}$, $i = 1, \ldots, r$. The union of events has also significance here because, by De Morgan's equality,

$$\bigcap_{i=1}^{r} A_i = \overline{\bigcup_{i=}^{r} \overline{A}_i}$$

and, consequently,

$$P\left(\bigcap_{i=1}^{r} A_i\right) = 1 - P\left(\bigcup_{i=1}^{r} \overline{A}_i\right).$$

Sometimes it is easier to present bounding formulas for the union.

Let us introduce the notation

$$S_k = \sum_{1 \leq i_1 < \cdots < i_k \leq r} P\left(A_{i_1} \cap \ldots \cap A_{i_r}\right), \quad k = 0, 1, \ldots, r,$$

where $S_0 = 1$. These values appear in the inclusion-exclusion formula:

$$P(A_1 \cup \ldots \cup A_r) = S_1 - S_2 + \cdots + (-1)^{r-1} S_r. \tag{5.1}$$

Formula (5.1) provides us, in principle, with the possibility to find the probability of the union of events, provided that we can find the probabilities of the intersections of any number of them. However, if r is large, then this is not the case and in practice we are able to find only a few of $S_1, S_2, \ldots$.

Let ν designate the number of events, out of $A_1, \ldots, A_r$, which occur. Then we have the following

Theorem 5.1. *The following equalities hold true*

$$E\left[\binom{\nu}{k}\right] = S_k, \quad k = 1, \ldots, r.$$

For a proof see, e.g., Prékopa (1995).

In view of Theorem 5.1 we call the values $S_1, S_2, \ldots$ binomial moments. Let

$$v_i = P(\nu = i), \quad i = 0, 1, \ldots, r.$$

Then Theorem 5.1 can be stated in the equivalent form

$$\sum_{i=0}^{r} \binom{i}{k} v_i = S_k, \quad k = 0, 1, \ldots, r. \tag{5.2}$$

The $v_0, \ldots, v_r$ and $S_1, \ldots, S_r$ values uniquely determine each other through the relation (5.2).

If only $S_1, \ldots, S_m$ are known in some situation, then we can formulate the question: what are the best lower and upper bounds, for the probability of the union, that can be given, based on this information. The answer to this question is given by the pair of linear programming problems:

$$\begin{aligned}
&\min(\max) \sum_{i=1}^{r} v_i \\
&\text{subject to} \\
&\sum_{i=0}^{r} \binom{i}{k} v_i = S_k, \quad k = 0, 1, \ldots, m \\
&v_i \geq 0, \quad i = 0, 1, \ldots, r.
\end{aligned} \tag{5.3}$$

In (5.3) $v_0, \ldots, v_r$ are decision variables, they are no longer uniquely determined by the available binomial moments $S_1, \ldots, S_m$. A slightly more convenient formulation of these LP's is:

$$\begin{aligned}
&\min(\max) \sum_{i=1}^{r} v_i \\
&\text{subject to} \\
&\sum_{i=1}^{r} \binom{i}{k} v_i = S_k, \quad k = 1, \ldots, m \\
&v_i \geq 0, \quad i = 1, \ldots, r.
\end{aligned} \tag{5.4}$$

Problem (5.4) arises from problem (5.3) in such a way that we remove v_0 as well as the constraint involving S_0. If $V_{\min}$ and $V_{\max}$ designate the optimum values of problems (5.4), then $V_{\min}$ is also the optimum value of the (5.3) minimum problem and $\min(V_{\max}, 1)$ is the optimum value of the (5.3) maximum problem. Problems (5.3) and (5.4) are called binomial moment problems.

The dual of problem (5.4) can be written as

$$\max(\min) \sum_{k=1}^{m} y_k S_k$$

subject to

$$\sum_{k=1}^{m} \binom{i}{k} y_k \underset{(\geq)}{\leq} 1. \tag{5.5}$$

While in problem (5.4) the input data (the binomial moments $S_1, \ldots, S_m$) are in the constraints, in problem (5.5) the constraints are universal, independent on the special events, and the input data appear only in the objective function.

Let $A = (a_1, \ldots, a_r)$, b, c designate the matrix of the equality constraint, the right hand side and the objective function coefficient vectors, respectively.

Definition 5.1. A basis B in the minimization (maximization) problem (5.4) is said to be dual feasible if

$$c_B^T B^{-1} a_k \leq c_k, \quad k = 1, \ldots, r$$
$$\left(c_B^T B^{-1} a_k \geq c_k, \quad k = 1, \ldots, r\right).$$

Note that inequality holds with equality if a_k is a basic vector.

Definition 5.2. A basis B in any of the problems (5.4) is said to be dual nondegenerate if

$$c_B^T B^{-1} a_k \neq c_k \quad \text{for nonbasic } a_k.$$

It is well-known in linear programming theory that a dual feasible basis in the minimization (maximization) problem has objective function value which is a lower (upper) bound for the optimum value. In view of this we have the following relations

$$c_{B_1}^T B_1^{-1} b \leq V_{\min} \leq P(A_1 \cup \ldots \cup A_r) \leq V_{\max} \leq c_{B_2}^T B_2^{-1} b, \tag{5.6}$$

where $B_1(B_2)$ is a dual feasible basis for the minimum (maximum) problem (5.4).

We have a complete description of dual feasible bases of problem (5.4). This is expressed in

Theorem 5.2. *Every dual feasible basis in problem (5.4) is dual non-degenerate and has one of the following structures described by the subscripts of the vectors:*

	m even	m odd
min problem	$i, i+1, \ldots, j, j+1,$	$i, i+1, \ldots, j, j+1, r$
max problem	$1, i, i+1, \ldots, j, j+1, r$	$1, i, i+1, \ldots, j, j+1.$

In other words, if m is even, then a basis is dual feasible in the min (max) problem iff. the subscript set of the basic vectors consists of consecutive pairs (1, r and consecutive pairs). If m is odd, then in case of the min (max) problem the subscript set is formed by consecutive pairs and r (by 1 and consecutive pairs). The optimal dual vector y, i.e., the optimal solution of problem (5.5) has an interesting property expressed in

Theorem 5.3. *The components $y_1, \ldots, y_m$ of the optimal solution y of problem (5.5) have alternating signs, starting with +, and have the following property*

$$|y_1| \geq |y_2| \geq \ldots \geq |y_m|.$$

Since all bases of problem (5.4) are dual nondegenerate, the optimal basis is unique. If m is small ($m \leq 4$), then we can find the optimal basis in such a way that we look for that basis, among the dual feasible ones, which is also primal feasible. This method provides us with bounds that can be obtained in a relatively simple way for the cases of $m = 2, 3$. For $m = 4$ we present the upper bound. The lower bound is complicated and we disregard its presentation here. The number m, indicating the largest number of intersections of which the probabilities appear in a formula, is called the order of the bound. The bounds presented below are sharp in the sense that under the given information (input data) no better bounds can be obtained.

Second order bounds using S_1, S_2

The lower bound is expressed by

$$P(A_1 \cup \ldots \cup A_r) \geq \frac{2}{i+1} S_1 - \frac{2}{i(i+1)} S_2,$$

where

$$i = 1 + \left\lfloor \frac{2S_2}{S_1} \right\rfloor.$$

The upper bound is expressed by

$$P(A_1 \cup \ldots \cup A_r) \leq \min\left(S_1 - \frac{2}{r}\, S_2,\, 1\right).$$

Third order bounds using S_1, S_2, S_3

The lower bound is expressed by

$$P(A_1 \cup \ldots \cup A_r) \geq \frac{i + 2r - 1}{(i+1)r}\, S_1 - \frac{2(2i + r - 2)}{i(i+1)r}\, S_2 + \frac{6}{i(i+1)r}\, S_3,$$

where

$$i = \left\lfloor \frac{-6S_3 + 2(r-2)S_2}{-2S_2 + (r-1)S_1} \right\rfloor.$$

The upper bound is given by

$$P(A_1 \cup \ldots \cup A_r) \leq S_1 - \frac{2(2i-1)}{i(i+1)}\, S_2 + \frac{6}{i(i+1)}\, S_3,$$

where

$$i = 2 + \left\lfloor \frac{3S_3}{S_2} \right\rfloor.$$

Fourth order upper bound using S_1, S_2, S_3, S_4

The bound is expressed by

$$P(A_1 \cup \ldots \cup A_r) \leq$$
$$\min\left(S_1 - \frac{2((i-1)(i-2)+(2i-1)r)}{i(i+1)r}\, S_2 + \frac{6(2i+r-4)}{i(i+1)r}\, S_3 - \frac{24}{i(i+1)r}\, S_4, 1\right).$$

Any bound, obtained by any method in the literature, is either a special case of our bounds, in the sense that it is the objective function value corresponding to some dual feasible basis in problem (5.4), or it is not the best possible bound and can be majorized by one of our bounds. As an example we mention the Bonferroni bounds. By Theorem 5.2 the basis

$B = (a_1, \ldots, a_m)$ is dual feasible in the minimization (maximization) problem if m is even (odd). Since

$$B^{-1} = \begin{pmatrix} 1 & 2 & 2 & \cdots & m \\ & 1 & \binom{3}{2} & \cdots & \binom{m}{2} \\ & & \ddots & & \vdots \\ & & & & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & -2 & 3 & \cdots & (-1)^{m-1}m \\ & 1 & -\binom{3}{2} & \cdots & (-1)^{m-2}\binom{m}{2} \\ & & & \ddots & \vdots \\ & & & & 1 \end{pmatrix},$$

it follows that

$$c_B^T B^{-1} b = S_1 - S_2 + \cdots + (-1)^{m-1} S_m$$

and, by (5.6), we have established the Bonferroni bounds:

$$P(A_1 \cup \ldots \cup A_r) \geq S_1 - S_2 + \cdots + S_{m-1} - S_m,$$

if m is even and

$$P(A_1 \cup \ldots \cup A_r) \leq S_1 - S_2 + \cdots + S_{m-2} - S_{m-1} + S_m,$$

if m is odd.

For an m for which the bounds are not available in formulas, we can execute a simple dual algorithm, a variant of Lemke's dual method, to find the optimum of problem (5.4). The algorithm presented below is valid for both the minimization and maximization problems (5.4).

Dual algorithm to solve problem 5.4

Step 0. Find an initial dual feasible basis B, by the use of the structural Theorem 5.2.

Step 1. If $B^{-1}b \geq 0$, Stop, the basis B is optimal and $c_B^T B^{-1} b$ gives us the optimal value, i.e., the required bound for the probability of the union. Otherwise go to Step 2.

Step 2. Choose any j such that $(B^{-1}b)_j < 0$ and remove the jth vector from the basis B. Go to Step 3.

Step 3. Include that vector into the basis which restores the dual feasible basis structure described in Theorem 5.2. There is exactly one such vector. Go to Step 1.

Note that the incoming vector can be found by a simple search procedure. The other parts of the algorithm can also be executed in a simple way as described in Prékopa (2001b).

Another probability bounding scheme, which provides us with better bounds but the corresponding LP is more difficult to solve, is the Boolean scheme or Boolean problem. In this case we use the joint probabilities of events individually, rather than just their sums in the binomial moments S_k.

Let again $A_1, \ldots, A_r$ be events in an arbitrary probability space and introduce the notations

$$v_J = P\left(\left(\bigcap_{j\in J} A_j\right)\left(\bigcap_{j\in \overline{J}} \overline{A}_j\right)\right)$$

$$p_I = P\left(\bigcap_{i\in I} A_i\right),$$

where $I, J \subset \{1, \ldots, r\}$. Here v_J is the probability that the events $A_j, j \in J$ occur but the events $A_j, j \in \overline{J}$ do not occur; p_I is the probability that the events $A_i, i \in I$ occur. If we introduce the incidence matrix

$$a_{IJ} = \begin{cases} 1 & \text{if } I \subset J \\ 0 & \text{if } I \not\subset J, \end{cases}$$

then we have the equation

$$\sum_{J\subset\{1,\ldots,r\}} a_{IJ} v_j = p_I, \quad I \subset \{1, \ldots, r\}. \tag{5.7}$$

If for input data those probabilities p_I are available for which $|I| \leq m$ (where $|I|$ designates the number of elements of the set I), then (5.7), restricted to $|I| \leq m$, does not determine uniquely the probabilities v_J. We can, however, write up minimization and maximization LP's which provide us with the best lower and upper bounds for the probability of the union, under the given information. As we have done in problem (5.4), here too, we disregard one variable and one constraint, those which correspond to $J = \emptyset$ and $I = \emptyset$, respectively. So we are led to the following problems

$$\min(\max) \sum_{\emptyset\neq J\subset\{1,\ldots,r\}} v_J$$

subject to

$$\sum_{\emptyset\neq J\subset\{1,\ldots,r\}} a_{IJ} v_J = p_I, \emptyset \neq I \subset \{1, \ldots, r\}, \quad |I| \leq m$$

$$v_J \geq 0, \emptyset \neq J \subset \{1, \ldots, r\}. \tag{5.8}$$

Problem (5.8) is called Boolean probability bounding scheme.

Problems (5.4) can be regarded as aggregated problems of those in (5.8) and problems (5.8) are disaggregated as compared to those in (5.4). Problems (5.8) provide us with better bounds than problems (5.4). However, no general dual feasible basis structure theorem is available for problems (5.8) and its numerical solution is computationally intensive as there are 2^r-1 variables in it. Still, a number of important bounds can be derived from (5.8) and some bounds that have been known in the literature can be recovered as objective function values corresponding to some dual feasible bases in (5.8). In the first category we mention the cherry tree bound and its generalizations. In the second category noteworthy is the classical Hunter's bound.

Hunter's upper bound for the probability of the union of events

Let $A_1, \ldots, A_r$ be events in an arbitrary probability space and $p_i = P(A_i)$ $p_{ij} = P(A_i \cap A_j)$ for $i \neq j$. Create the complete graph with nodes $1, \ldots, r$ and assign to arc (i, j) the weight p_{ij}, $i \neq j$. Let T be any spanning tree in this graph. Then we have the relation

$$P(A_1 \cup \ldots \cup A_r) \leq S_1 - \sum_{(i,j)\in T} p_{ij}, \tag{5.9}$$

where $S_1 = p_1 + \cdots + p_r$. The best upper bound of the type (5.9) is obtained from the heaviest spanning tree T^*:

$$\sum_{(i,j)\in T^*} p_{ij} = \max_{T \text{ spanning tree}} \sum_{(i,j)\in T} p_{ij}.$$

Inequality (5.9) provides us with Hunter's upper bound if we choose $T = T^*$.

To find the heaviest spanning tree Kruskal's algorithm is available. It consists of the following steps:

Step 1. Initialize $k = 1$ and find the heaviest arc. Go to Step 2.
Step 2. Increase k by 1. If $k = n$ then Stop, the heaviest spanning tree T^* has been found. Otherwise go to Step 3.
Step 3. Find the heaviest arc that does not create cycle. Go to Step 2.

Any upper bound in (5.9) can be represented as the objective function value corresponding to a suitable dual feasible basis in the maximization problem (5.8) when $m = 2$. The basis can be constructed as follows: take all paths that can be created by the use of the nodes and arcs in T; in case of any of these paths take the set of nodes used by the path and consider it as a label set J in problem (5.8); the columns in problem (5.8) corresponding to these labels provide us with the required dual feasible basis.

Hunter's upper bound is always at least as good as the second order binomial moment bound.

Finally, we mention that the probability bounding schemes can be incorporated into the probabilistic constrained stochastic programming problems. We present two examples in this respect.

In the first example we use the simple probability bound

$$P(A_1 \cap \ldots \cap A_r) \geq \sum_{i=1}^{r} P(A_i) - (r-1),$$

replace $T_i x \geq \xi_i$ for A_i and impose a probabilistic constraint on the right hand side to obtain

$$\sum_{i=1}^{r} P(T_i x \geq \xi_i) - (r-1) \geq p.$$

This can replace the constraint $P(Tx \geq \xi) \geq p$ in problem (1.3). However, we are somewhat better off, if we formulate for problem (1.3) the following approximate problem:

$$\begin{aligned}
&\min c^T x \\
&\text{subject to} \\
&P(T_i x \geq \xi_i) \geq p_i, \quad i = 1, \ldots, r \\
&\sum_{i=1}^{r} (1 - p_i) \leq 1 - p \\
&Ax = b,\ x \geq 0,
\end{aligned} \tag{5.10}$$

where p is a fixed probability but $p_1, \ldots, p_r$ are variables. It can be shown that the first two sets of constraints imply that $P(Tx \geq \xi) \geq p$. If $F_i(z)$ is the probability distribution function of the random variable ξ_i, then the constraint $P(T_i x \geq \xi_i) \geq p_i$ is equivalent to $T_i x \geq F^{-1}(p_i)$, $i = 1, \ldots, r$. If we replace problem (5.10) for problem (1.3), then the set of feasible solutions shrinks, in general, and the optimum value of problem (5.10) will be larger than that of problem (1.3). The same is true for the second example that we present below.

Let us introduce the functions:

$$\begin{aligned}
&S_k(x) = \sum_{1 \leq i_1 < \ldots < i_k \leq r} P\big(T_{i_1} x \geq \xi_{i_1}, \ldots, T_{i_r} x \geq \xi_{i_r}\big), \\
&k = 0, 1, \ldots, m,
\end{aligned}$$

where $S_0(x) \equiv 1$. Out of problem (1.3) we create the new problem:

$$\begin{aligned}
&\min c^T x \\
&\text{subject to} \\
&\sum_{i=0}^{r} \binom{i}{k} v_i = S_k(x), \quad k = 0, 1, \ldots, m \\
&v_r \geq p \\
&Ax = b,\ x \geq 0\ .
\end{aligned} \tag{5.11}$$

Similar problem can be formulated by the use of the Boolean minimization problem but in this case we have to reestablish the constraint as well as the variable corresponding to the empty set. In fact, in problem (5.11) we have used the binomial moment problem (5.3), and not (5.4), because problems (5.3) and (5.4) are equivalent only in the case of bounding the union whereas in problem (5.11) we create a bound for the intersection.

There is a considerable recent literature on the binomial and Boolean probability bounding schemes.

5.2 *Approximation and simulation of probabilities*

Programming under probabilistic constraints is the research area where optimization combined with simulation first appeared as a problem solving methodology.

The joint probabilistic constraint with stochastically dependent random variables presumes that we are able to compute or at least estimate the values and gradients of multivariate distribution functions. The normal distribution seems to be the most frequent among the multivariate distributions hence we pay special attention to it.

The r-variate nondegenerate standard normal probability density function has the form

$$\varphi(z;\ R) = \frac{1}{(2\pi)^{r/2}\sqrt{|R|}}\ e^{-z^T R^{-1} z}, \quad z \in \mathbb{R}^r.$$

A random variable ξ that has this probability density function can be represented in the form

$$\xi = \chi_r T \zeta,$$

where χ_r is χ-distributed random variable with r degrees of freedom, ζ is an r-variate random vector, uniformly distributed on the surface of the unit sphere:

$$S^r = \left\{ z \mid \sum_{i=1}^{r} z_i^2 = 1 \right\}$$

and T is lower triangular matrix such that

$$TT^T = R.$$

The most important problem is to approximate, or estimate the integral

$$p = \int_Q \varphi(z;\ R)\, \mathrm{d}z, \tag{5.12}$$

where Q is a finite or infinite rectangle. An efficient method works in the following way.

Let the rectangle Q be bounded, for the sake of simplicity, and given by

$$Q = \{z \mid a \leq z \leq b\}.$$

Let further $k_r(t)$ designate the probability density function of the χ-distribution with r degrees of freedom and

$$z_1(v) = \min \rho, \quad \text{subject to} \quad a \leq \rho T v \leq b$$
$$z_2(v) = \max \rho, \quad \text{subject to} \quad a \leq \rho T v \leq b.$$

The probability in (5.12) can be expressed in the form

$$p = \int_Q \varphi(z;\ R)\, \mathrm{d}z = \int_{S^r} \left(\int_{z_1(v)}^{z_2(v)} k_r(t)\, \mathrm{d}t \right) \mathrm{d}U(v), \tag{5.13}$$

where U is the probability distribution function of ζ.

To estimate the probability p in (5.13) we can generate a sample $v_1, \ldots, v_N$ for the random vector ζ and take

$$\frac{1}{N} \sum_{i=1}^{N} e(v_i),$$

where $e(v)$ designates the interior integral on the right hand side of (5.13). This way, however, the procedure is slow. In order to speed it up, the generation of the sampling elements of S^r by the use of random orthonormalized systems has been proposed. This means that we randomly pick an orthonormalized system of vectors $g_1, \ldots, g_r$ ($g_i^T g_j = \delta_{ij}$, where δ_{ij} is Kronecker's delta), then choose k out of the r in all possible ways and for each choice of k vectors we multiply them by $+1$ and -1 in all possible ways. The number of orthonormal systems generated this way out of the single system $g_1, \ldots, g_r$, is $2^k\binom{r}{k}$. One choice of k vectors can be represented by an index set $I = \{i_1, \ldots, i_k\} \subset \{1, \ldots, r\}$ and a set of $+1$, -1 multipliers is designated by $s_1, \ldots, s_k$. Then we form the sum

$$g(S, I) = \frac{1}{\sqrt{k}} \sum_{j=1}^{k} s_j g_{i_j}.$$

If initially there are N orthonormal systems $g_1^{(l)}, \ldots, g_r^{(l)}$, $l = 1, \ldots, N$ chosen, then our estimation for the probability p is

$$\vartheta_k = \frac{1}{N} \sum_{l=1}^{N} \frac{1}{2^k \binom{r}{k}} \sum_{S,I} e\big(Tg^{(l)}(S, I)\big), \tag{5.14}$$

where the second summation extends over all k element subsets of $\{1, \ldots, r\}$ and all 2^k k-component vectors with components -1, $+1$. The value ϑ_k is an unbiased estimator of p.

The main advantage of the above procedure is that whenever we generate one orthonormalized system, we immediately produce $2^k\binom{r}{k}$ out of it. There is, however, another important advantage of it offered by the formula (5.14). When we compute $Tg^{(l)}(S, I)$, then first we form the products $Tg_1^{(l)}, \ldots, Tg_r^{(l)}$ and only after that pick the k vectors and multiply them by $s_1, \ldots, s_k$. This arrangement of computation also saves considerable time.

The method has been extended to find probabilities of convex polyhedra and other convex sets in $\mathbb{R}^n$. The only difficulty in the more general cases is to find the intersections $z_1(v)$, $z_2(v)$ of the straight line $\rho^T v$ $(-\infty < \rho < \infty)$ with the boundary of the convex set. It is reported that three digit accuracy can be obtained in less than 1 s. for all $r \leq 20$.

The next simulation technique that we describe is more general, it can be applied in principle for arbitrary multivariate probability distributions. However, numerical results are available only for three multivariate distributions: normal, gamma and Dirichlet. We describe the general method for the case where $A_i = \{\xi_i \leq x_i\}$, $i = 1, \ldots, r$ and we want to estimate the joint probability distribution function of $\xi_1, \ldots, \xi_r$:

$$F(x_1, \ldots, x_r) = P(A_1 \cap \ldots \cap A_r).$$

We pass to the complementary events $\overline{A}_i = \{\xi_i > x_i\}$, $i = 1, \ldots, r$, and define

$$\overline{S}_k = \sum_{1 \le i_1 < \ldots < i_k \le r} P\left(\overline{A}_{i_1} \cap \ldots \cap \overline{A}_{i_k}\right), \quad k = 1, \ldots, r.$$

If we use $\overline{S}_1$, $\overline{S}_2$, $\overline{S}_3$, then, using the bounds described in Section 5.1, we can create lower and upper bounds for the probability of the union $\overline{A}_1 \cup \ldots \cup \overline{A}_r$, and, in turn, for the probability of the intersection which is $F(x_1, \ldots, x_r)$. Three lower bounds: L_1, L_2, L_3 and two upper bounds: U_1, U_2 are used. These are the following

$$L_1 = 1 - \overline{S}_1 \quad \text{first order binomial moment bound}$$

$$L_2 = 1 - \overline{S}_1 + \frac{2}{r}\overline{S}_2 \quad \text{second order binomial moment bound}$$

$$L_3 = 1 - \overline{S}_1 + \sum_{(i,j) \in T^*} P\left(\overline{A}_i \cap \overline{A}_j\right) \quad \text{Hunter's bound}$$

$$U_1 = 1 - \overline{S}_1 + \overline{S}_2 \quad \text{second order Bonferroni bound}$$

$$U_2 = 1 - \frac{2}{i+1}\,\overline{S}_1 + \frac{2}{i(i+1)}\overline{S}_2 \text{ second order binomial moment bound,}$$

where $i = 1 + \left\lfloor \dfrac{2\overline{S}_2}{\overline{S}_1} \right\rfloor$.

We have the relations:

$$F(x_1, \ldots, x_r) - L_1 = \overline{S}_2 + \overline{S}_3 - \cdots + (-1)^r \overline{S}_r$$

$$F(x_1, \ldots, x_r) - L_2 = \left(1 - \frac{2}{r}\right)\overline{S}_2 - \overline{S}_3 + \cdots + (-1)^r \overline{S}_r$$

$$F(x_1, \ldots, x_r) - L_3 = -\sum_{(ij) \in T^*} P\left(\overline{A}_i \cap \overline{A}_j\right) + \overline{S}_2 - \overline{S}_3 + \cdots + (-1)^r \overline{S}_r$$

$$F(x_1, \ldots, x_r) - U_1 = \overline{S}_3 + \cdots + (-1)^r \overline{S}_r$$

$$F(x_1, \ldots, x_r) - U_2 = \left(\frac{k}{i+1} - 1\right)\overline{S}_1 + \left(1 - \frac{2}{i(i+1)}\right)\overline{S}_2 - \overline{S}_3 + \cdots + (-1)^r \overline{S}_r. \tag{5.15}$$

The simulation technique uses the exact values of the univariate and bivariate marginal probability distribution function values. We compute them

by deterministic numerical integration method. These values enter into L_1, L_2, L_3, U_1, U_2. Then we simulate the values on the right hand sides of ((5.15)), to obtain estimations of $F(x_1, \ldots, x_r)$.

Let $(\xi_1^{(s)}, \ldots, \xi_r^{(s)})$, $s = 1, \ldots, N$ be a sample of size N for the random vector $(\xi_1, \ldots, \xi_r)$. For fixed s let $\kappa^{(s)}$ designate the number of those inequalities $\xi_1^{(s)} \le x_1, \ldots, \xi_r^{(s)} \le x_r$ which are not fulfilled. Let further $\lambda^{(s)}$ designate the number of those pairs $(i,j) \in T^*$ for which we have $\xi_i^{(s)} > x_i$, $\xi_j^{(s)} > x_j$. Since we have the relations

$$E\left[\binom{\kappa^{(s)}}{k}\right] = \overline{S}_k, \quad k = 0, 1, \ldots, r; \quad s = 1, \ldots, N$$

$$E\left(\lambda^{(s)}\right) = \sum_{(i,j)\in T^*} P\left(\overline{A}_i \cap \overline{A}_j\right),$$

it follows that $\binom{\kappa^{(s)}}{k}$ and $\lambda^{(s)}$ are unbiased estimators of the right hand side values in (5.16), respectively. If we use this and the relation

$$(1-1)^{\kappa^{(s)}} = \sum_{j=0}^{\kappa^{(s)}} (-1)^j \binom{\kappa^{(s)}}{j} = 0,$$

then we can easily show that the follwooing random variables are unbiased estimators of the right hand sides of (5.15):

$$\nu_{L_1}^{(s)} = \begin{cases} \kappa^{(s)} - 1, & \text{if } \kappa^{(s)} \ge 2 \\ 0 & \text{otherwise} \end{cases}$$

$$\nu_{L_2}^{(s)} = \begin{cases} \frac{1}{r}\left(\kappa^{(s)} - 1\right)\left(r - \kappa^{(s)}\right), & \text{if } \kappa^{(s)} \ge 2 \\ 0 & \text{otherwise} \end{cases}$$

$$\nu_{L_3}^{(s)} = \begin{cases} \kappa^{(s)} - 1 - \lambda^{(s)}, & \text{if } \kappa^{(s)} \ge 2 \\ 0 & \text{otherwise} \end{cases}$$

$$\nu_{U_1}^{(s)} = \begin{cases} \frac{1}{2}\left(\kappa^{(s)} - 1\right)\left(2 - \kappa^{(s)}\right), & \text{if } \kappa^{(s)} \ge 3 \\ 0 & \text{otherwise} \end{cases}$$

$$\nu_{U_2}^{(s)} = \begin{cases} \frac{\left(i-\kappa^{(s)}\right)\left(\kappa^{(s)}-i-1\right)}{i(i+1)}, & \text{if } \kappa^{(s)} \ge 1 \\ 0 & \text{otherwise} \end{cases}$$

Taking averages with respect to s we obtain five unbiased estimators for $F(x_1, \ldots, x_r)$:

$$\nu_{L_j} = L_j + \frac{1}{N} \sum_{s=1}^{N} \nu_{L_j}^{(s)}, \quad j = 1, 2, 3$$

$$\nu_{U_j} = U_j + \frac{1}{N} \sum_{s=1}^{N} \nu_{U_j}^{(s)}, \quad j = 1, 2. \tag{5.17}$$

A sixth unbiased estimator is given by

$$\nu_0 = \frac{1}{N} \sum_{s=1}^{N} \nu_0^{(s)}, \tag{5.18}$$

where $\nu_0^{(s)} = 1$, if all relations $\xi_1^{(s)} \le x_1, \ldots, \xi_r^{(s)} \le x_r$ are satisfied and $\nu_0^{(s)} = 0$ otherwise.

Out of the six estimators in (5.17) and (5.18) one estimator is formed:

$$\nu = w_0\nu_0 + w_{L_1}\nu_{L_1} + w_{L_2}\nu_{L_2} + w_{L_3}\nu_{L_3} + w_{U_1}\nu_{U_1} + w_{U_2}\nu_{U_2},$$

where the sum of hte weights is equal to one and weights are computed in such a way that the variance of ν should be minimum. The covariances of the six estimators for this minimization problem are estimated from the sample.

Based on the above two simulation methods a hybrid method can be created. Let again $C_j = [a_j, b_j], j = 1, \ldots, r$ and define the sets

$$D_i(v) = \{\rho \mid a_j \le \rho T_j v \le b_j \text{ is violated for exactly } i \text{ indices } j\}.$$

We estimate the probability $P(\overline{A}_1 \cup \cdots \cup \overline{A}_r)$, where $\overline{A}_j = \{\xi_j \notin C_j\}, j = 1, \ldots, r$. Let $\overline{p}_j = P(\overline{A}_j), j = 1, \ldots, r$. Then we have

$$\overline{S}_k = \sum_{i=1}^{r} \binom{i}{k} \overline{p}_i, \quad k = 1, \ldots, r.$$

By Theorem 5.1 we have the equalities

$$\overline{S}_2 - \overline{S}_3 + \cdots + (-1)^r \overline{S}_r = \sum_{i=2}^{r} (i-1)\overline{p}_i \tag{5.19}$$

$$-\overline{S}_3 + \cdots + (-1)^r \overline{S}_r = -\sum_{i=3}^{r} \binom{i-1}{2} \overline{p}_i. \tag{5.20}$$

In addition, from the previous descriptions we know that

$$\overline{p}_i = \int_{S^r} \left(\int_{D_i(v)} k_r(t)\, \mathrm{d}t \right) \mathrm{d}U(v), \quad i = 1, \ldots, r. \tag{5.21}$$

Combining the above–mentioned equalities we obtain

$$\begin{aligned} P(A_1 \cap \cdots \cap A_r) &= P\left(\overline{A}_1 \cup \cdots \cup \overline{A}_r\right) \\ &= 1 - \overline{S}_1 + \left(\overline{S}_2 + \cdots + (-1)^r \overline{S}_r\right) \\ &= 1 - \overline{S}_1 + \int_{S^r} \left(\sum_{i=2}^{r} (i-1) \int_{D_i(v)} k_r(t)\, \mathrm{d}t \right) \mathrm{d}U(v) \end{aligned} \tag{5.22}$$

and

$$\begin{aligned} P(A_1 \cap \cdots \cap A_r) &= P\left(\overline{A}_1 \cup \cdots \cup \overline{A}_r\right) \\ &= 1 - \overline{S}_1 + \overline{S}_2 + \left(-\overline{S}_3 + \cdots + (-1)^r \overline{S}_r\right) \\ &= 1 - \overline{S}_1 + \overline{S}_2 - \int_{S^r} \left(\sum_{i=3}^{r} \binom{i-1}{2} \int_{D_i(v)} k_r(t)\, \mathrm{d}t \right) \mathrm{d}U(v). \end{aligned} \tag{5.23}$$

If we compute $\overline{S}_1$ and $\overline{S}_2$ exactly, by some numerical quadrature and choose randomly the vector v, that appear on the right hand sides of (5.21)–(5.23), and designate by ν_1 and ν_2, respectively the obtained last terms in these equations, then we have two estimators for $P(A_1 \cap \cdots \cap A_r)$. These are

$$\hat{P}_1 = 1 - \overline{S}_1 + \nu_1, \hat{P}_2 = 1 - \overline{S}_1 + \overline{S}_2 + \nu_2.$$

We can use $\hat{P}_0 = \nu_0$ in (5.18) as a third estimator. Our final estimator is a minimum variance linear combination of the three:

$$\hat{P} = w_0 \hat{P}_0 + w_1 \hat{P}_1 + w_2 \hat{P}_2$$

where $w_0 + w_1 + w_2 = 1$. The covariance of $\hat{P}_0, \hat{P}_1, \hat{P}_2$ can be estimated from the sample and w_0, w_1, w_2 can be obtained as optimal solution of the problem:

$$\begin{aligned} &\min w^T C w \\ &\text{subject to} \\ &w_0 + w_1 + w_2 = 1. \end{aligned}$$

A simple procedure is available to find the sets $D_i(v)$, $i = 1, \ldots, r$, for a given v.

The above described method has been compared for the case of a multivariate normal distribution to methods of deterministic numerical integration. It is suggested that before choosing the method to approximate the probability $p = P(A_1 \cap \cdots \cap A_r)$, a few trial point should be generated to obtain a preliminary indication about the magnitude of the probability p. It is stated that, choosing the most suitable method, a twenty dimensional probability can be computed with four digit accuracy in less than six minutes on an ordinary desktop computer.

To close this section, we briefly describe a recently developed numerical integration method to find the values of the multivariate normal integrals. The method works both in the nondegenerate and degenerate cases and its code is publicly available on the internet.

Assume that the distribution is a nondegenerate standard normal distribution with correlation matrix Σ of which the Cholesky factorization $\Sigma = CC^T$ is known. Let a and b be the lower and upper boundary points of the r-dimensional rectangular set, respectively. We want to compute the integral

$$p = \int_{a_1}^{b_1} \cdots \int_{a_r}^{b_r} \frac{1}{|\Sigma|^{1/2}(2\pi)^{r/2}} \, e^{-\frac{1}{2} x^T \Sigma^{-1} x} \, \mathrm{d}x_r \, \cdots \, \mathrm{d}x_1. \tag{5.24}$$

If we use the transformation $x = Cy$, then we have to integrate with respect to the components of y satisfying

$$\begin{aligned} a_1 \le y_1 \le b_1 \\ a_i'(y_1, \ldots, y_{i-1}) = \left(a_i - \sum_{j<i} c_{ij} y_j\right) \frac{1}{c_{ii}} \\ \le y_i \\ \le \left(b_i - \sum_{j<i} c_{ij} y_j\right) \frac{1}{c_{ii}} = b_i'(y_1, \ldots, y_{i-1}), \quad i = 2, \ldots, r. \end{aligned}$$

With this transformation the integral (5.24) becomes (φ is the univariate standard normal probability density function):

$$\int_{a_1}^{b_1} \varphi(y_1) \int_{a_2'(y_1)}^{b_2'(y_1)} \varphi(y_2) \cdots \int_{a_r'(y_1,\ldots,y_{r-1})}^{b_r'(y_1,\ldots,y_{r-1})} \varphi(y_r)\, \mathrm{d}y_r \cdots \mathrm{d}y_1.$$

If we introduce the further transformation $y_i = \Phi^{-1}(z_i)$, $i = 1, \ldots, r$, where Φ is the univariate standard normal distribution function, and introduce the notations

$$g_1 = \Phi(a_1), \quad h_1 = \Phi(b_1)$$
$$g_i(z_1, \ldots, z_{i-1}) = \Phi\left(\left(a_i - \sum_{j=1}^{i-1} c_{ij}\Phi^{-1}(z_j)\right) \Big/ c_{ii}\right)$$
$$h_i(z_1, \ldots, z_{i-1}) = \Phi\left(\left(b_i - \sum_{j=1}^{i-1} c_{ij}\Phi^{-1}(z_j)\right) \Big/ c_{ii}\right), \quad i = 2, \ldots, r,$$

then the integral (5.24) transforms into

$$p = \int_{g_1}^{h_1} \int_{g_2(z_1)}^{h_2(z_1)} \cdots \int_{g_r(z_1,\ldots,z_{r-1})}^{h_r(z_1,\ldots,z_{r-1})} \mathrm{d}z_r \cdots \mathrm{d}z_1.$$

The final transformation $z_i = g_i + w_i(h_i - g_i)$, $i = 1, \ldots, r$ transforms the integral into

$$p = (h_1 - g_1) \int_0^1 (h_2(w_1) - g_2(w_1)) \int_0^1 \cdots$$
$$\int_0^1 (h_r(w_1, \ldots, w_{r-1}) - g_r(w_1, \ldots, w_{r-1}))\, \mathrm{d}w_{r-1} \cdots \mathrm{d}w_1$$

where the integration region is standardized.

The above procedure is applicable in any ordering of the variables. The method works best if the innermost integral carries the most weight, then comes the second etc. These weights can be ranked by ranking the probabilities $\Phi(b_i) - \Phi(a_i)$, $i = 1, \ldots, r$.

5.3 Calculation of the gradient values

Let $\xi = (\xi_1, \ldots, \xi_r)$ be a continuously distributed random vector, $F(z_1, \ldots, z_r)$ its probability distribution function and $f_i(z)$ the probability density function of ξ_i, $i = 1, \ldots, r$. Let further

$$F(z_1, \ldots, z_{i-1}, z_{i+1}, \ldots, z_r \mid z_i)$$
$$= P(\xi_1 \leq z_1, \ldots, \xi_{i-1} \leq z_{i-1}, \xi_{i+1} \leq z_{i+1}, \ldots, \xi_r \leq z_r \mid \xi_i = z_i)$$
$$i = 1, \ldots, r.$$

It is easy to see that

$$\frac{\partial F(z_1, \ldots, z_r)}{\partial z_i} = F(z_1, \ldots, z_{i-1}, z_{i+1}, \ldots, z_r \mid z_i) f_i(z_i)$$
$$i = 1, \ldots, r \tag{5.25}$$

and this formula provides us with a general method to compute the gradients of F. In case of many known probability distributions the conditional distribution function $F(z_1, \ldots, z_{i-1}, z_{i+1}, \ldots, z_r \mid z_i)$ is of the same type as the original distribution function, hence using the same code and a code to calculate $f_i(z_i)$, we can obtain the ith component of ∇F.

As an example we present the gradient of the multivariate standard normal probability distribution function $\Phi(z_1, \ldots, z_r; R)$, where $R = (\rho_{ij})$ is the correlation matrix. We can use formula (5.25), where we replace

$$F(z_1, \ldots, z_{i-1}, z_{i+1}, \ldots, z_r \mid z_i)$$
$$= \Phi\left(\frac{z_1 - \rho_{1,i} z_i}{\sqrt{1 - \rho_{1,i}^2}}, \ldots, \frac{z_{i-1} - \rho_{i-1,i} z_i}{\sqrt{1 - \rho_{i-1,i}^2}}, \frac{z_{i+1} - \rho_{i+1,i} z_i}{\sqrt{1 - \rho_{i+1,i}^2}}, \ldots, \frac{z_r - \rho_{r,i} z_i}{\sqrt{1 - \rho_{r,i}^2}}; R\right),$$

and R is the $(r-1) \times (r-1)$ correlation matrix with entries

$$s_{j,k} = \frac{\rho_{j,k} - \rho_{j,i}\, \rho_{k,i}}{\sqrt{1 - \rho_{j,i}^2}\sqrt{1 - \rho_{k,i}^2}}, \quad j, k = 1, \ldots, r, j \neq i, k \neq i.$$

In addition we replace $\varphi(z_i)$ for $f_i(z_i)$, where φ is the standard normal probability density function. We see that $\nabla\Phi(z; R)$ can be computed by computing the values of $r-1$-variate standard normal distribution functions and the values of the univariate standard normal probability density function.

Note that in the probabilistic constraint $P(Tx \geq \xi) \geq p$ the random vector can be standardized and we can write the constraint in the form:

$$\Phi\left(\frac{T_i x - \mu_i}{\sigma_i},\ i = 1, \ldots, r;\ R\right) \geq p.$$

To compute the gradient of the constraining function the just obtained formula can be used.

Other examples, where the gradients can be computed by the use of the same type but lower dimensional distribution functions, include the gamma and the Dirichlet distributions.

General formulas for gradients of "probability functions" are also available. We look at the function

$$F(x) = \int\limits_{f(x,\,y)\leq 0} p(x, y)\,\mathrm{d}y,$$

where $p(x, y)$ is a probability density function and $f(x, y) \leq 0$ is the compact form of the inequalities $f_i(x, y) \leq 0$, $i = 1, \ldots, k$. There are altogether three formulas, the first one is called the integral over the surface formula, the second one the integral over the volume formula and the third one is a general formula.

The integral over the surface formula

Let $\mu(x) = \{y \,|\, f(x, y) \leq 0\}$ and $\partial\mu(x)$ the boundary of the set $\mu(x)$. Let further $\partial\mu_i(x) = \mu(x) \cap \{y \,|\, f_i(x, y) = 0\}$, i.e., that part of the boundary set $\partial\mu(x)$ which is determined by f_i. Then we have the equation

$$\nabla_x F(x) = \int\limits_{\mu(x)} \nabla_x p(x, y)\,\mathrm{d}y - \sum_{i=1}^{k} \int\limits_{\partial\mu_i(x)} \frac{p(x, y)}{\|\nabla_y f_i(x, y)\|}\ \nabla_x f_i(x, y)\,\mathrm{d}S.$$

The integral over the volume formula

We introduce the notations ($x \in \mathbb{R}^n, y \in \mathbb{R}^m$):

$$f_{1,\,l}(x, y) = \begin{pmatrix} f_1(x, y) \\ \vdots \\ f_l(x, y) \end{pmatrix}, \quad f(x, y) = f_{1,\,k}(x, y)$$

$$\nabla_y f(x, y) = \begin{pmatrix} \frac{\partial f_1(x, y)}{\partial y_1} & \cdots & \frac{\partial f_k(x, y)}{\partial y_1} \\ \vdots & \vdots & \vdots \\ \frac{\partial f_1(x, y)}{\partial y_m} & \cdots & \frac{\partial f_k(x, y)}{\partial y_m} \end{pmatrix}$$

$$H = \begin{pmatrix} h_{1,1} & \cdots & h_{1,m} \\ \vdots & \vdots & \vdots \\ h_{n,1} & \cdots & h_{n,m} \end{pmatrix}, \quad h_{i,j} = h_{i,j}(x, y)$$

$$div_y H = \begin{pmatrix} \sum_{i=1}^{m} \frac{\partial h_{1,j}}{\partial y_j} \\ \vdots \\ \sum_{i=1}^{m} \frac{\partial h_{n,j}}{\partial y_j} \end{pmatrix}, \quad j = 1, \ldots, m.$$

We have the gradient formula

$$\nabla_x F(x) = \int_{\mu(x)} \nabla_x p(x, y)\, \mathrm{d}y + \int_{\mu(x)} div_y(p(x, y)H(x, y))\, \mathrm{d}y,$$

where $H(x, y)$ satisfies the equation

$$H(x, y)\nabla_y f(x, y) + \nabla_x f(x, y) = 0.$$

This last equation may not have a solution and in that case the general formula may be useful.

The general formula for the gradient

We split the set of constraints into two groups and designate the corresponding subscript sets by $K_1 = \{1, \ldots, l\}$ and $K_2 = \{l + 1, \ldots, k\}$. We have the formula

$$\nabla_x F(x) = \int_{\mu(x)} \nabla_x p(x, y)\, \mathrm{d}y + \int_{\mu(x)} div_y(p(x, y)H_l(x, y))\, \mathrm{d}y$$
$$- \sum_{i=l+1}^{k} \int_{\partial_i \mu(x)} \frac{p(x, y)}{\|\nabla_y f_i(x, y)\|} \left[\nabla_x f_i(x, y) + H_l(x, y)\nabla_y f_i(x, y)\right] \mathrm{d}S,$$

where the $n \times m$ matrix $H_l(x, y)$ satisfies the equation

$$H_l(x, y)\nabla_y f_{1,l}(x, y) + \nabla_x f_{1,l}(x, y) = 0.$$

5.4 Bibliographical notes

Combined use of simulation and optimization appeared first, in the stochastic programming context, in a paper by Prékopa et al. (1980). The paper was presented at the First International Conference on Stochastic Programming, in Oxford, England, 1974.

For the history of inclusion–exclusion formula see Takács (1967). For the proof of Theorem 5.1 see Takács (1967) and Prékopa (1995). Problems (5.3)–(5.5) and Theorem 5.2 are taken from Prékopa (1988). Since S_k is the kth binomial moment of v, problems (5.3) and (5.4) are termed binomial moment problems. Theorem 5.3 is from Boros and Prékopa (1989).

The second order lower bound, based on S_1, S_2, was obtained by Dawson and Sankoff (1967). The upper bound is due to Kwerel (1975a,b) and Sathe et al. (1980). The third order lower and upper bounds have been obtained by Kwerel (1975a,b) and Prékopa and Boros (1989). Kwerel used linear programming theory specialized for the cases $m = 2, 3$, without writing up the more general problems (5.3), (5.4). Boros and Prékopa (1989) presented a variety of bounds, based on the dual feasible basis structure theorem (Theorem 5.2). The fourth order bound, based on S_1, S_2, S_3, S_4 is also from Boros and Prékopa (1989). The Bonferroni bounds are due to Bonferroni (1937). The dual algorithm to solve problem (5.4) is due to Prékopa (1988). A more refined version of it, formulated for the (equivalent) power moment problem, was presented in Prékopa (2001b).

Problem (5.8), more exactly its dual was initiated by Boole (1854). Its exact formulation is due to Hailperin (1965). Hunter's upper bound is due to Hunter (1976) and Worsley (1983). The algorithm to find the heaviest spanning tree is due to Kruskal (1956). One generalization of Hunter's bound is presented in Prékopa et al. (2001) and another one, the cherry tree bound, in Bukszár and Prékopa (2001). The use of bounds in probabilistic constrained stochastic programming problems, presented at the end of Section 5.1, is from Prékopa (1999).

Other useful probability bounds are presented in Galambos and Simonelli (1996), Bukszár and Szántai (2002), Bukszár (2001) etc.

The simulation method based on (5.13), to estimate the normal probability distribution function value, is due to Deák (1980, 1986, 1988, 1990, 2000b, 2002). The next, more general method that can be used to estimate the values of arbitrary probability distribution functions, is due to Szántai (1986, 2000). The hybrid method was proposed by Gassmann (1988) and it is called Deák, Szántai, Gassmann (DSG) method. A more recent paper by Gassmann et al. (2002)

improves on this and revises other methods to compute multivariate normal integrals. The last method in Section 5.2 is due to Genz (1992).

The formula to compute gradients of multivariate normal integrals was first proposed in Prékopa (1970) and used in computerized problem solution in Prékopa et al. (1980). Gradient formula for the multivariate gamma distribution, presented in Section 2.2 (distribution (8)) is described in Prékopa and Szántai (1978a). Szántai (1985) derived the gradient of the Dirichlet distribution. For its description see also Prékopa (1995).

The general surface and volume integral formulas are due to Uryasev (1989, 2001). Some special formulas have been derived by Raik (1972), Kibzun and Kurbakooskij (1991), Kibzun and Kan (1996) and Marti (1988).

6 Duality and stability

Probabilistic constrained stochastic programming problems are nonlinear optimization problems, hence nonlinear duality theory has a straightforward application there. There is, however, one primal-dual relationship, involving two special probabilistic constrained problems, and a corresponding duality theorem which are special and deserve presentation here. The theory that we describe is due to Komáromi (1986).

Let A be an $m \times n$ matrix and $\beta \in \mathbb{R}^m$, $\tau \in \mathbb{R}^n$ two continuously distributed random variables. Designate by F and G the probability distribution functions of β and τ, respectively, and suppose that they are quasi-concave functions. Let *supp* F and *supp* G designate the supports of F and G, respectively, i.e., the smallest closed sets where the probability measures, generated by these distribution functions, are equal to one. Introduce the notations

$$
\begin{aligned}
B &= \{b \mid F(b) \geq p,\ b \in \operatorname{supp} F\} \\
X(b) &= \{x \mid Ax \geq b,\ x \geq 0\} \\
C &= \{c \mid G(-c) \geq q,\ -c \in \operatorname{supp} G\} \\
Y(c) &= \{y \mid A^T y \leq c,\ y \geq 0\}.
\end{aligned}
$$

Consider the pair of primal-dual problems:

$$
\begin{aligned}
&\min \sup_{c \in C} c^T x \\
&\text{subject to} \\
&F(b) \geq p,\ b \in \operatorname{supp} F \\
&Ax \geq b,\ x \geq 0
\end{aligned}
\tag{6.1}
$$

and

$$\begin{aligned} &\max \inf_{b \in B} b^T y \\ &\text{subject to} \\ &G(-c) \geq q, \ -c \in \operatorname{supp} G \\ &A^T y \leq c, \ y \geq 0. \end{aligned} \tag{6.2}$$

Note that the constraints in problem (6.1) are equivalent to $P(Ax \geq \beta) \geq p$, $x \geq 0$ and the constraints in problem (6.2) are equivalent to $P(A^T y \leq \tau) \geq q$, $y \geq 0$. We have the following duality theorem.

Theorem 6.1. *The following three assertions hold true.*

(a) *Suppose that the function F is strictly increasing in each variable,* $\operatorname{supp} G$ *is bounded and Slater's condition: $int\{b \mid b \in B, X(b) \neq \emptyset\} \neq \emptyset$ holds. If the objective function of problem (6.1) is unbounded, then the problem (6.2) has no feasible solution. Otherwise, problem (6.2) has feasible solution, the optimum values of the two problems are equal and that value is attained in problem (6.2).*

(b) *Suppose that the function G is strictly increasing in each variable,* $\operatorname{supp} F$ *is bounded and Slater's condition: $int\{c \mid c \in C, Y(c) \neq \emptyset\} \neq \emptyset$ holds. If the objective function of problem (6.2) is unbounded, then problem (6.1) has no feasible solution. Otherwise, problem (6.1) has feasible solution, the optimum values of the two problems are equal and that value is attained in problem (6.1).*

(c) *Suppose that both F and G are strictly increasing functions in each variable,* $\operatorname{supp} F$ *and* $\operatorname{supp} G$ *are bounded and Slater's conditions: $int\{b \mid b \in B, X(b) \neq \emptyset\} \neq \emptyset$, $int\{c \mid c \in C, Y(c) \neq \emptyset\} \neq \emptyset$ hold. Then both problems (6.1) and (6.2) have optimal solutions. If (x^0, y^0) is a pair of optimal solutions of the problems, then it is a saddle point of $y^T Ax$ with respect to minimizing over X and maximizing over Y.*

A logconcave probability distribution function is also quasi-concave, hence the above theory applies to this case.

Another interesting duality theorem, involving logconcave probability distributions, was obtained by Luc (1983).

Stability and sensitivity are important issues in probabilistic constrained stochastic programming problems. Since decision making takes place based on tail probabilities, it is particularly important to know how the optimum value of the problem changes if the probability distribution of the random variables in the model changes. If the probability distribution belongs to some known class, the question can be formulated in such a way that: how the change of the distribution parameters influences the optimum value or the problem in general?

The case of the normal distribution offers nice illustration. Consider problem (1.3), where ξ has a multivariate normal distribution with expectations $E(\xi_i) = \mu_i$, variances $Var(\xi_i) = \sigma_i^2$, $i = 1, \ldots, r$ and correlation matrix R. The probabilistic constraint can equivalently be written in the form

$$P(Tx \geq \xi) = P\left(\frac{T_i x - \mu_i}{\sigma_i} \geq \frac{\xi_i - \mu_i}{\sigma_i},\ i = 1, \ldots, r\right)$$
$$= \Phi(L_1(x), \ldots, L_r(x);\ R) \geq p,$$

where $L_i(x) = (T_i(x) - \mu_i)/\sigma_i$, $i = 1, \ldots, r$. If we introduce the new variables $y_i = L_i(x)$, $i = 1, \ldots, r$, then problem (1.3) can be written in the equivalent form:

$$\begin{aligned} &\min\ c^T x \\ &\Phi(y;\ R) \geq p \\ &L(x) = y \\ &Ax = b,\ x \geq 0. \end{aligned} \tag{6.3}$$

The changes in μ_i, σ_i, $i = 1, \ldots, r$ change only the linear constraint $L(x) = y$ and its effect to the optimum value can be computed by the use of some standard methods (see, e.g., Prékopa (1995, Chapter 15)). When we apply a method of this kind first we linearize the probabilistic constraining function around the optimal y but then concentrate on the effect of the change in the next linear constraint.

The problem is more complicated if the change occurs in the correlation matrix R. We present a few facts that can be used to do further analysis.

Let $\varphi(y;\ R)$ designate the r-variate standard normal probability density function with correlation matrix $R = (\rho_{i,j})$. Then, as it is easy to check, we have the equality

$$\frac{\partial \varphi}{\partial \rho_{i,j}} = \frac{\partial^2 \varphi}{\partial y_i \partial y_j}.$$

Integrating on both sides with respect to y, in $-\infty < y_i \leq z_i$, $i = 1, \ldots, r$, and choosing $i = 1, j = 2$, for the sake of simplicity, we obtain the equation

$$\frac{\partial \Phi(z;\ R)}{\partial \rho_{1,2}} = \int_{-\infty}^{z_1} \cdots \int_{-\infty}^{z_r} \frac{\partial^2 \varphi(y;R)}{\partial y_1 \partial y_2}\, \mathrm{d}y_r \ \ldots\ \mathrm{d}y_2\, \mathrm{d}y_1$$
$$= \int_{-\infty}^{z_3} \cdots \int_{-\infty}^{z_r} \varphi(y;R)\, \mathrm{d}y_r \ \ldots\ \mathrm{d}y_3. \tag{6.4}$$

If we choose another correlation matrix $K = (\kappa_{i,j})$ and take

$$\nu_{i,j} = \lambda\rho_{i,j} + (1-\lambda)\kappa_{i,j}$$

as functions of the variable $0 \le \lambda \le 1$, then, in view of (6.4), we get

$$\frac{d\Phi(z;\ G)}{d\lambda} = \sum_{i<j} \frac{\partial\Phi(z;\ G)}{\partial\nu_{i,j}}(\rho_{i,j} - \kappa_{i,j}), \tag{6.5}$$

where $G = (\nu_{i,j})$. If $\rho_{i,j} \ge \kappa_{i,j}$, then the derivative in (6.5) is positive and integrating on the left hand side of (6.5) with respect to λ from 0 to 1, we obtain the inequality

$$\Phi(z; R) \ge \Phi(z; K), \quad \text{for } R \ge K. \tag{6.6}$$

Inequality (6.6) is known as Slepian's inequality (see Slepian (1962)).

A large number of stability results have been obtained over the past fifteen years by Dupačová (1991) Kall (1987), Römisch and Schultz (1991), Henrion and Römisch (1998, 2000), Henrion (2000) and others. As the results and their descriptions are rather involved we only demonstrate the flavor of the more recent results by presenting a theorem from Henrion and Römisch (1998).

Consider problem (1.3) and assume that ξ has an α-concave probability density function (see Theorem 2.9 for the implication with respect to the distribution of ξ), the problem has feasible solution and finite optimum. Let $\Psi(P)$ designate the set of optimal solutions. Together with P we consider another probability distribution Q and designate by F_p and F_Q, respectively, the corresponding distribution functions. The set of optimal solutions under the distribution Q is designated by $\Psi(Q)$. We assume that there exists an $\overline{x}$ such that $A\overline{x} = b$, $\overline{x} \ge 0$ and $F_P(T\overline{x}) > p$. Moreover, we assume that F_p is strictly convex in a convex neighborhood of $A\Psi(P)$. We have the following.

Theorem 6.2. *Under the above conditions there exist constants* $L > 0$, $\delta > 0$ *such that*

$$d_H(\Psi(P), \Psi(Q)) \le L\|F_P - F_Q\|_\infty^{\frac{1}{2}},$$

whenever

$$\|F_P - F_Q\|_\infty < \delta.$$

7 Selected applications

7.1 Energy problems

An optimal investment problem, formulated for the electrical energy sector of the Hungarian economy, was the first real life application of probabilistic constrained stochastic programming with stochastically dependent random variables. The application was done by Prékopa et al. and was presented at the First International Conference on Stochastic Programming held in Oxford, England, 1974 (see Prékopa et al. (1980)). The model was based on a deterministic model that had been formulated earlier. The number of stochastic constraints is four in the stochastic model. Let $G(x)$ designate the probability that all these are satisfied. Let further x_{lin} and x_{stoch} designate the optimal solutions of the deterministic and stochastic problems, respectively. An interesting phenomenon turned up. The reliability level of the joint probability of the stochastic constraints, taken with x_{lin}, was only 0.1, whereas the optimum value $c^T x$ subject to $G(x) \geq p$, and the other deterministic constraints, came out as (almost) the same as $c^T x_{lin}$, i.e., $c^T x_{lin} = c^T x_{stoch}$, in both the $p = 0.9$ and $p = 0.95$ cases. Thus, in case of each p, an optimal solution, providing us with high reliability in power service, could be obtained with no additional cost as compared to the optimal cost in the deterministic problem. The optimal solutions x_{lin} and x_{stoch}, however, were different. Normal distribution was used and the method of feasible directions to solve the problem.

Two-stage optimal capacity design in power networks, with probabilistic constraint for the solvability of the second stage problem, was formulated in Prékopa (1980). Numerical solution for a special case of this problem was presented by Prékopa et al. (1998). Optimal capacity design problems with probabilistic constraints have been formulated by several authors. Among them we mention Bloom et al. (1984), Bloom (1988), Coté and Laughton (1982), Bisthoven et al. (1988) and Dodu et al. (1981).

7.2 Water resources

One of the revealing problems, where the use of joint probabilistic constraint is the only justified possibility, is a flood control reservoir system design problem. In its simplest (but nontrivial) version there are only two possible reservoir sites where capacities x_1, x_2 have to be determined and these serve to protect a downstream area from flood that may happen once in a year, say. If ξ_1, ξ_2 are the water amounts to be retained by the reservoirs (see Fig. 1 for the topology of the system).

Then the flood will be retained if and only if $x_1 + x_2 \geq \xi_1 + \xi_2$, $x_2 \geq \xi_2$ are satisfied. Since ξ_1, ξ_2 are random variables, the fulfilment of these inequalities can be guaranteed only on a probability level p, chosen by ourselves.

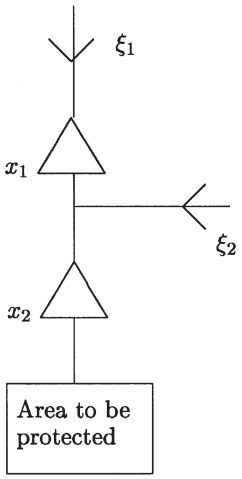

Fig. 1

If $c(x_1, x_2)$ is the reservoir building cost function, then our stochastic programming problem is:

$$\min c(x_1, x_2)$$

subject to

$$P\begin{pmatrix} x_1 + x_2 \geq \xi_1 + \xi_2 \\ x_2 \geq \xi_2 \end{pmatrix} \geq p$$

$$0 \leq x_1 \leq V_1,\ 0 \leq x_2 \leq V_2, \tag{7.1}$$

where V_1, V_2 are upper bounds determined by the local geographic situation. We can see in problem (7.1) that to prescribe probabilistic constraints separately for the two stochastic constraints has no meaning at all because they jointly ensure the reliability of the system. The problem was solved under normal and gamma distributions. The above simple and a more general flood control problem was formulated and solved by Prékopa and Szántai (1978b). The method of supporting hyperplanes was applied to solve the problem. Further results in this respect are in Kelman et al. (1989), Prékopa Rapcsák and Zsuffa (1978) formulated and solved a reservoir system design problem, where the possible sites are located along one river. The probabilistic constraint prescribes the simultaneous probability of retaining the streamflow and serve all demands. Normal distribution was used and the SUMT to solve the problem.

Prékopa and Szántai (1976) formulated and solved a multi-period reservoir system operation problem, using the rolling horizon principle. A multivariate

gamma distribution was fitted to the empirical data and was used, together with the supporting hyperplane method, in the solution of the optimization problem.

A sequential probability maximization problem was formulated for the water level regulation of Lake Balaton in Hungary, the largest lake in Central and Western Europe. A Gaussian process was used to describe the inflow process. The result enabled to improve on the water level regulation reliability (to keep the water level within prescribed limits) from the former 80% to 97.5%.

Dupačová et al. (1991) compared the different reservoir system operation models and solutions under different probability distributions.

7.3 Production and inventory

A reliability type multi-item inventory problem was formulated by Prékopa and Kelle (1978). Further results are in Kelle (1984, 1985). It is assumed that during a given period of time the total delivery is the same as the total demand in each item. However, deliveries take place at random epochs and random quantities. In order to ensure that all demands be met during the given period, initial safety stocks are needed. The problem is to minimize their total holding costs subject to a reliability constraint that serves to ensure the above-mentioned requirement. The SUMT was used to solve the optimization problem.

Murr and Prékopa (2000) have solved a product substitution problem in connection with fiber manufacturing. The manufacturing process produces random yield and the problem is to set the original production goals so that all demands be met, on a given reliability level, with minimum cost. Normal distribution was assumed and the method of feasible direction was applied (Szántai's code).

Beraldi and Ruszczyński (2002) formulated and solved the stochastic set covering problem which has many applications in the production and service industry. The right hand side random variables ξ_i may represent occurrences of requests for service and the jth column of the technology matrix has entries that describe capabilities of the different facilities to respond to these requests. The authors have developed their own method to solve the problem, where both the random variables and the decision variables are discrete (0–1-valued).

Beraldi and Ruszczyński (2001) have formulated a probabilistic lot sizing problem, as an application of their general method to solve probabilistic constrained stochastic programming problems with integer valued right hand side random variables. The problem is to minimize the total setup, production and holding costs subject to the condition that all demands should be met in the course of the planning horizon.

Singh, Abraham and Akella (1990) formulated and solved a chip manufacturing problem. Given a number of possible chip sites and chip types, the problem is to find an optimal allocation of types to sites so that the probability of getting a prescribed non-defective chip composition will be

maximized. In this problem both the random and the decision variables are nonnegative integer valued.

Henrion et al. (2001) and Henrion and Möller (2002) have formulated a model for a continuous distillation process under stochastic inflows in a feed tank. The problem is to control an extracting process so that lower and upper level constraints in the feed tank should be met by a large probability. In the most important case, analysed in this paper regarding the type of randomness, the inflow process is supposed to be Gaussian. For the numerical solution of the problem Szántai's code, to solve the probabilistic constrained problems, was applied. Similar problem is dealt with in another paper by Henrion et al. (2001). The above problem appears to have strong connection to the water level regulation problem of Lake Balaton mentioned in Section 2.

7.4 Telecommunication problems

Some of the most recent works in this area are the following. Dentcheva, Prékopa and Ruszczyński (2000) formulated and solved a traffic assignment problem for Time Division Multiple Access (TDMA) satellite communication systems. The problem has the form:

$$\min \sum_{i=1}^{n} x_i$$

subject to

$$P\left(\sum_{i=1}^{n} Q^{(i)} x_i \geq D\right) \geq p$$

$$x \geq 0, \text{ integer},$$

where $Q^{(i)}$, $i = 1, \ldots, n$ are $m \times m$ permutation matrices and D is an $m \times m$ matrix with nonnegative integer entries (representing demands). The solution technique that solves the problem, the cone generation technique, is described in Section 3.2.

Medova and Scott (2000) formulated a quality of service management problem, where an upper bound is imposed on blocking probabilities. These are transformed into simpler constraints, by the use of large deviation theory. Then the total cost of link capacities minus revenues is minimized subject to the above mentioned constraints.

In her Thesis Heikkinen (2001) formulated an elegant model for the stochastic power control problem in mobile telecommunication systems. The model has interesting connection to game theory and von Neumann's economic model. Further results are in Heikkinen and Prékopa (2002) and Gao and Prékopa (2001).

7.5 Diet problems and food service management

A classical formulation of the deterministic diet problem is an LP: $\min c^T x$, subject to $Ax \geq b$, $x \geq 0$. Here A is an $m \times n$ matrix with entries equal to the nutrient contents of the different foods, m is the number of nutrients, n is the number of foods, b is the nutrient requirement vector and c is the vector of costs of the unit amounts of the different foods.

If one thinks that the food is served for a population, where each individual has his/her own nutrient requirement vector, then the probabilistic constrained problem (where we use β rather than b on the right hand side) can be formulated as:

$$\begin{aligned} &\min c^T x \\ &\text{subject to} \\ &P(Ax \geq \beta) \geq p \\ &x \geq 0, \end{aligned} \tag{7.2}$$

to decide on the quantities of foods to be served. In practical problems A may also be random in which case the solution of the problem becomes hard. Approximation and solution to problem (7.2), for this case, is presented in Armstrong and Balintfy (1975). Further papers on the problem include Balintfy and Prékopa (1966), Balintfy and Armstrong (1980), Lancaster (1992). There are other formulations of the diet problem too, and given the underlying problem, there are stochastic programming formulations for the problem, other than (7.2), as well. Problem (7.2) prescribes that $100p\%$ of the population should receive all nutrients on at least minimum level. If we formulate the problem by the use of conditional expectation constraints (see Section 2.4), then we prescribe upper bounds on the averages of the unserved nutrients.

In this area another problem is the minimum cost animal feed problem. A classical work in this respect is the one published by van de Panne and Popp (1963).

7.6 Finance problems

There are a large number of stochastic programming models applied to financial problems. Most of them, however, that have been formulated so far, belong to the class of recourse problems. Recently, safety type considerations, that have existed since the nineteen fifties but have not gained enough attention during the past decades, came into prominence. The important step

in this direction was the formulation of the concept of Value at Risk (VaR) and its variants. These notions have already existed in a probabilistic/statistical framework with the name of quantile and its generalizations or variants. Before presenting some results in this respect we mention an application of probabilistic constrained stochastic programming to optimal portfolio composition.

The classical portfolio models of Markowitz (1952, 1959, 1987) have safety aspect but it is incorporated into the models in the form of the variance of the return. Given the expected return, the smaller the variance of the return, the better the portfolio. Markowitz looks for efficient portfolios which means that given the expectation of the return, its variance cannot be decreased and given the variance, the expectation cannot be increased. To illustrate the power of probabilistic constrained stochastic programming formulation we present a bond portfolio construction model. Let us introduce the notations:

- n number of bond types which are candidates for inclusion into the portfolio
- m number of periods
- a_{ik} cash flow of a bond of type k in period i, $k = 1, \dots, n \quad i = 1, \dots, m$
- p_k unit price of bond of type k
- ξ_i random liability value in period i, $i = 1, \dots, m$
- x_k decision variable, number of bonds of type k to include into the portfolio
- z_i cash carried forward from period i to period $i+1$, $i = 1, \dots, m$, where z_1 is an initial cash amount that we include into the portfolio and $z_{m+1} = 0$; z_i, $i = 1, \dots, m$ are decision variables
- ρ_i rate of interest in period i, $i = 1, \dots, m$.

If the liabilities were deterministic values then our optimal bond portfolio model would be the following

$$\min\left\{\sum_{k=1}^{n} p_k x_k + z_1\right\}$$

subject to

$$\sum_{k=1}^{n} a_{ik} x_k + (1-\rho_i) z_i - z_{i+1} \geq \xi_i, \quad i = 1, \dots, m$$

$$x_k \geq 0, \quad k = 1, \dots, n$$

$$z_i \geq 0, \quad i = 1, \dots, m, \; z_{m+1} = 0. \tag{7.3}$$

The probabilistic constrained variant of it can be formulated as

$$\min\left\{\sum_{k=1}^{n} p_k x_k + z_1\right\}$$
$$\text{subject to}$$
$$P\left(\sum_{k=1}^{n} a_{i\,k} x_k + (1-\rho_i) z_i - z_{i+1} \geq \xi_i,\ i = 1, \ldots, m\right) \geq p$$
$$x_k \geq 0, \quad k = 1, \ldots, n$$
$$z_i \geq 0, \quad i = 1, \ldots, m, \tag{7.4}$$

where p is a safety (reliability) level chosen by ourselves, e.g., $p = 0.8, 0.9, 0.95$ etc. Some of the liability values (e.g., those corresponding to the early periods) may be deterministic. Then they should be removed from the probabilistic constraint in (7.4) and listed separately. The properties and solution methods of the model can be learned from theorems and methods presented in the previous sections.

Value at Risk, or VaR has been defined in connection with a random variable ζ or its probability distribution function $F(z)$ by the equation $F(z) = p$, where $0 < p < 1$. If $F(z)$ is strictly increasing then there is exactly one solution to this equation. Otherwise we may take the smallest z satisfying $F(z) \geq p$. This definition of VaR is suitable in connection with a random variable that designates loss. If η designates revenue then we take $\xi = -\eta$ which means loss and we define VaR for ξ. The VaR can be defined in connection with a portfolio, where the total random return is $\xi^T x$. In this case we may look for the VaR of the probability distribution of $-\xi^T x$, i.e., for that value of z that minimizes $P(-\xi^T x \leq z) = P(\xi^T x \geq -z)$ subject to the constraint that this probability is at least p.

A closely related notion is the Conditional Value at Risk (CVaR), defined in connection with a random variable ζ (or its probability distribution) by

$$\text{CVaR}(\zeta) = E(\zeta \mid \zeta - z > 0), \quad z = \text{VaR}(\zeta).$$

provided that the distribution is continuous. Other forms of CVaR(ζ) are:

$$\text{CVaR} = \frac{1}{1-p} \int_z^\infty u \, \mathrm{d}F(u), \quad z = \text{VaR}(\zeta) \tag{7.5}$$

$$\text{CVaR} = \inf_a \left\{a + \frac{1}{1-p} E([\zeta - a]_+)\right\}. \tag{7.6}$$

This last form of CVaR is due to Rockafellar and Uryasev (2000). Minimizing CVaR can be used as a decision principle to compose optimal portfolios. An example is the following credit risk optimization problem (see Andersson et al. (2001)). Suppose there are n obligors and let x_i, b_i, ξ_i designate the weight, the debt and the lost part of it of the ith obligor. Then the loss function is $f(x, \xi) = (b - \xi)^T x$. If we replace $f(x, \xi)$ for ζ in (7.6), then we obtain the objective function

$$a + \frac{1}{1-p} E([f(x, \xi) - a]_+) \tag{7.7}$$

that is to be minimized with respect to x and a. Suppose that there are some deterministic constraints, expressed by $x \in X$, and we approximate the expectation in (7.7) by the use of a sample with respect to the distribution of $\xi : y_1, \ldots, y_n$, then we obtain the approximate problem:

$$\min_{a, x, z} \left\{ a + \frac{1}{1-p} \frac{1}{N} \sum_{i=1}^{N} z_i \right\}$$

subject to

$$x \in X$$

$$z_i \geq f(x, y_i), \quad z_i \geq 0, i = 1, \ldots, N.$$

Another example is presented in Borgentoft et al. (2001).

There are many other applications of the probabilistic constrained stochastic programming model and its variant, maximizing a probability subject to constraints. E.g., Pickens et al. (1991) applies it to forrestry, Pintér (1991) to environmental problems, Thoft-Christensen and Murotsu (1986) to engineering structures, Singh, Abraham and Akella (1990) to chip manufacturing, Shapiro (1986) to insurance, Cooper et al. (1998) to data envelopment analysis, Kibzun and Kan (1996) to aviation, etc.

References

Abadie, J., J. Carpentier (1969). Generalization of Wolfe reduced gradient method to the case of nonlinear constraints, in: R. Fletcher (ed.), *Optimization*, Academic Press.

Andersson, F., H. Mausser, D. Rosen, S. Uryasev (2001). Credit risk optimization with conditional value at risk criterion. *Math. Programming Ser. B* 89, 273–291.

Armstrong, R.D., J.L. Balintfy (1975). A chance constrained multiple choice algorithm. *Oper. Res.* 23, 494–510.

Balintfy, J.L., R.D. Armstrong (1980). A chance constrained multiple choice programming algorithm with applications in: *Stochastic Programming* (Proc. Internat. Conf., Univ. Oxford, Oxford, England, 1974, M.A.H. Dempster). Academic Press, London, New York, pp. 301–325.

Balintfy, J.L., A. Prékopa (1996). Nature of random variation in the nutrient composition of meals. *Health Services Research* 1, 148–169.

Barndorff–Nielsen, O. (1973). Unimodality and exponential families. *Comm. Stat.* 1, 189–216.

Becker, T., V. Weispfenning (1993). *Gröbner Bases*, Springer, New York.
Beraldi, P., A. Ruszczyński (1999). The probabilistic set covering problem. *RUTCOR Research Reports* 88–99. *Op. Res.* 50, 956–967.
Beraldi, P., A. Ruszczyński (2002). A branch and bound method for stochastic integer problems under probabilistic constraints. *RUTCOR Research Reports* 16, 2001. *Optimization Method and Software* 17, 359–382.
Bisthoven, O.J., de P. Schuchewytsch, Y. Smeers (1988). Power generation planning with uncertain demand, in: Y. Ermeliev, R.J.-B. Wets (eds.), *Numerical Techniques for Stochastic Optimization*, Springer-Verlag, New York, pp. 465–480.
Bloom, J.A., M. Caramanis, L. Charny (1984). Long-range generation planning using generalized Benders' decomposition: implementation and experience. *Operations Research* 32, 290–313.
Bloom, J.A. (1988). Solving an electricity generating capacity expansion planning problem by generalized Benders' decomposition. *Operations Research* 31, 84–100.
Bogentoft, E., H.E. Romeijn, S. Uryasev (2001). Pension Funds Using CVaR constraints. *The Journal of Risk Finance*, 57–71.
Bonferroni, C.E. (1937). *Teoria Statistica Delle Classi e Calcolo Delle Probabilità*. Volume in onore di Riccardo Dalla Volta, Università di Firenze, 1–62.
Boole, G. (1854). *Laws of Thought*. Americam reprint of 1854 edition. Dover, New York.
Borell, C. (1975). Convex set functions in d-space. *Periodica Mathematica Hungarica* 6, 111–136.
Boros, E., A. Prékopa (1989). Closed form two-sided bounds for probabilities that exactly r and at least r out of n events occur. *Math. Opns. Res.* 14, 317–342.
Brascamp, H.J., E.H. Lieb (1976). On extensions of the Brunn-Minkowski and Prékopa-Leindler theorems, including inequalities for log concave functions, and with an application to the diffusion equations. *Journal of Functional Analysis* 22, 366–389.
Bukszár, J., A. Prékopa (2001). probability bounds with cherry trees. *Math. of Oper. Res.* 26, 174–192.
Bukszár, J., T. Szántai (2002). Probability bounds given by hypercherry trees. *Optimization Methods and Software* 17, 409–422.
Bukszár, J. (2001). Probability bounds with multitrees. *Adv. Appl. Prob.* 33, 437–452.
Burkauskas, A. (1986). On the convexity problem of probabilistic constrained stochastic programming problems (in Hungarian). *Alkalmazott Matematikai Lapok (Applied Mathematical Papers)* 12, 77–90.
Charnes, A., W.W. Cooper, G.H. Symonds (1958). Cost horizons and certainty equivalents; an approach to stochastic programming of heating oil. *Management Science* 4, 235–263.
Cooper, W.W., Z. Huang, S.X. Li, O.B. Olesen (1998). Chance constrained programming formulations for stochastic characterizations of efficiency and dominance in DEA. *J. of Productivity Anal.* 9, 53–79.
Coté, C., M. Laughton (1982). Stochastic production costing in generation planning: a large-scale mixed integer model. *Mathematical Programming Study*.
Davidovich, Yu.S., B.L. Korenblum, B.I. Hacet (1969). A property of logarithmically concave functions. *Soviet Mathematics, Doklady* 10, 477–480 (English Translation).
Dawson, D., A. Sankoff (1967). A inequality for probabilities. *Proc. Am. Math. Soc.* 18, 504–507.
Deák, I. (1980). Three digit Accurate Multiple Normal Probabilities. *Numerische Mathematik*, 35, 369–380.
Deák, I. (1990). *Random Number Generators and Simulation*, Akadémiai Kiadó, Budapest.
Deák, I. (1998). Linear regression estimators for multinormal distributions in optimization of stochastic programming problems. *EJOR* III, 555–568.
Deák, I. (1971). Computer evaluation of a stochastic programming model. Thesis, L. Eötvös Univ. of Budapest (in Hungarian).
Deák, I. (1986). Computing probabilities of rectangles in case of multidimensional distribution. *Journal of Statistical Computation and Simulation* 26, 101–114.
Deák, I. (1988). Multidimensional integration and stochastic programming, in: Yu Ermoliev, R.J.-B. Wets (eds.), *Numerical Techniques for Stochastic Optimization*, Springer-Verlag, Berlin, pp. 187–200.

Deák, I. (2000a). Solving stochastic programming problems by successive regression approximations. *Computational Optimization and Applications* (submitted).

Deák, I. (2000b). Subroutins for computing normal probabilities of sets – computer experinces. *Annals of Oper. Res.* 100, 103–122.

Deák, I. (2001). Two-stage stochastic problems with correlated normal variables: computational experiences. *Annals of Oper. Res.*, to appear.

Deák, I. (2002). Probabilities of simple *n*-dimensional sets for the normal distribution. *I.I.E. Transactions*, to appear.

Dentcheva, D., A. Prékopa, A. Ruszczyński (2000). Concavity and efficient points of discrete distributions in probabilistic programming. *Mathematical Programming Series A*. 89, 55–77.

Dert, C.L. (1995). *Asset-Liability Management for Pension Funds: A Multistage Chance-Constrained Programming Approach*, PhD Thesis, Erasmus University, Rotterdam.

Dodu, J.C., M. Goursat, A. Hertz, J.P. Quadrat, M. Viot (1981). Methode de gradient stochastique pour l'optimization des investissement dans un Réseau Électrique, Electricité de France. *Bulletin Dir. Etudes et Rech. Seriec. Math.* 2, 133–164 (in French).

Dupačová, J. (1991). Statistical security analysis in stochastic programming. *Annals of Operations Research* 30, 199–214.

Dupačová J., A. Gaivoronski, Z. Kos, T. Szántai (1991). Stochastic programming in water management: a case study and a comparsion of solution techniques. *Journal of Operations Research* 52, 28–44.

Fekete, M. and Gy. Pólya (1912). Über ein Problem von Laguerre. *Rediconti del Circolo Matematico di Palermo* 23, 89–120.

Fiacco, A.V., G.P. McCormick (1968). *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*, John Wiley & Sons, New York.

Galambos, J., J. Simonelli (1996). *Bonferroni-Type Inequalities with Applications*, Springer, New York.

Gao, L., A. Prékopa (2001). On performance prediction of cellular telephone networks. *RUTCOR Research Report*.

Gassmann, H.I. (1988). Conditional probability and conditional expectation of a random vector, in: Y. Ermeliev, R.J.-B. Wets (eds.), *Numerical Techniques for Stochastic Optimization*, Springer Verlag, Berlin, Heidelberg, New York, pp. 237–254.

Gassmann, H.I., A. Prékopa (2001). On stages and constraint replication in stochastic programming. Manuscript.

Gassmann, H.I., I. Deák, T. Szántai (2002). Computing multivariate normal probabilities. *J. of Computational and Graphical Stat.* 11, 920–949.

Genz, A. (1992). Numerical computation of the multivariate normal probabilities. *Journal of Computational and Graphical Statistics* 1, 141–150.

Gröwe, N. (1995). *Numerical Solution and Asymptotic Analysis of Stochastic Programs. Thesis*, Humboldt University, Berlin.

Gröwe, N. (1997). Estimated stochastic programs with chance constraint. *EJOR* 101, 285–305.

Hailperin, Th. (1965). Best possible inequalities for the probability of a logical functions of events, *Am. Math. Monthly* 72, 343–359.

Heikkinen, T., A. Prékopa (2002). Optimal power control under probabilistic quality-of-service constraints. Manuscript.

Heikkinen, T.M. (2001). On resource allocation in externality networks. Thesis, *RUTCOR, Rutgers Center for Operations Research*.

Henrion, R., A. Möller (2002). Optimization of a continuous distillation process under random inflow rate. *Computers and Math. with App.*, to appear.

Henrion, R., P. Li, A. Möller, M. Wendt, G. Wozny (2001). Optimal control of a continuous distillation process under probabilistic constraints, in: M. Grötschel, S. Krumke, J. Rambau (eds.) *Optimization of Large Scale Systems*, Springer, Berlin, pp. 499–517.

Henrion, R., W. Römisch (1998). Metric regularity and quantitative stability in stochastic programs with probabilistic constraints. *Mathematical Programming* 84, 55–88.

Henrion, R., W. Römisch (2000). Stability of solutions to chance constrained stochastic programs, in: J. Guddat, et al. (eds.), *Parametric Optimization and Related Topics V*, Peter Lang, Frankfurt a. M, pp. 95–114.

Henrion, R. (2000). Qualitative stability of convex programs with probabilistic constraints. *Lecture Notes in Economics and Math. Systems* 481, 164–180.

Hunter, D. (1976). Bounds for the probability of a union. *Journal of Applied Probability* 13, 597–603.

Kall, P. (1987). On approximations and stability in stochastic programming, in: *Parametric Optimization and Related Topics*, Akademie-Verlag, Berlin, pp. 387–407.

Kall, P., J. Mayer (1996). SLP–IOR: an interactive model management system for stochastic linear programs. *Mathematical Programming* 75, 221–240.

Kataoka, S. (1963). A stochastic progamming model. *Econometrica* 31, 181–196.

Kelle, P. (1984). On the safety stock problem for random delivery processes. *European Journal of Operational Research* 17, 191–200.

Kelle, P. (1985). Safety stock planning in a multi-stage production-inventory system. *Engineering Costs and Production Economics* 9, 231–237.

Kelman, J., J.M. Damazio, J.L. Mariën, J.P. da Costa (1989). The determination of flood control volumes in a multireservoir system. *Water Resources Research* 25, 337–344.

Kibzun, A.I., Y.S. Kan (1996). *Stochastic Programming Problems with Probability and Quantile Functions*, Wiley, New York.

Kibzun, A.I., V.Y. Kurbakovszkij (1991). Guaranteeing approach to solving quantile optimization problems. *Annals of Operations Research* 30, 81–99.

Klein Haneveld, W.K. (1986). *Duality in Stochastic Linear and Dynamic Programming. Lecture Notes in Economics and Mathematical Systems*, Springer-Verlag, New York.

Komáromi, É (1986a). Duality in probabilistic constrained programming, in: A. Prékopa et al. (eds.), System Modelling and Optimization. *Proceedings of the 12th IFIP Conference* 1986, pp. 423–429.

Komáromi, É. (1986b). A dual method of probabilistic constrained problem. *Mathematical Programming Study* 28, 94–112.

Komáromi, É. (1987). On properties of the probabilistic constrained linear programming problem and its dual. *Journal of Optimization Theory and Applications* 55, 337–390.

Kruskal, J. B. (1956). On the shortest spanning subtree of a graph and the traveling salesman problem. *Proceedings of the American Mathematical Society* 7, 48–50.

Kwerel, S.M. (1975a). Most stringent bounds on aggregated probabilities of partially specified dependent probability systems. *J. Am. Statist. Assoc*. 472–479.

Kwerel, S.M. (1975b). Bounds on the probability of a union and intersection of *m* events. *Adv. Appl. Prob*. 7, 431–448.

Lancaster, L.M. (1992). The evolution of the diet model in managing food systems. *Interfaces* 22, 59–68.

Luc, D.T. (1983). Duality in programming under probabilistic constrained problems. *Problems of Control and Information Theory* 12, 429–437.

Markowitz, H. (1952). Portfolio selection. *Journal of Finance* 7, 77–91.

Markowitz, H. (1959). *Portfolio Selection*, Wiley, New York.

Markowitz, H. (1987). *Mean-Variance Analysis in Portfolio Choice and Capital Markets*, Blackwell, New York.

Marti, K. (1988). Descent Directions and Efficient Solutions in Descretely Distributed Stochastic Programs, in: *Lecture Notes in Economics and Mathematical Systems* 299, Springer Verlag, Berlin, New York.

Mayer, J. (1980). A nonlinear programming method for the solution of a stochastic programming model of A. Prékopa, in: A. Prékopa (ed.), *Survey of Mathematical Programming*, Vol. 2, North-Holland, pp. 129–139.

Mayer, J. (1988). Probabilistic constrained programming: a reduced gradient algorithm implemented on PC. IIASA Working Paper, WP–88–39.

Mayer, J. (2000). On the numerical solution of jointly chance constrained problems. *Probabilistic Constrained Optimization: Methodology and Applications*, Kluwer Academic Publishers, pp. 220–233.

Mayer, J. (1992). Computational techniques for probabilistic constrained optimization problems). *Stochastic Optimization, Numerical Methods and Technical Applications*, Springer Verlag, Berlin.

Mayer, J. (1998). *Stochastic linear programming algorithms: a comparsion based on a model management system*, Gordon and Breach Science Publishers.

Medova, E.A., J.E. Scott (2000). Management of quality of service through chance constraint in multimedia networks, in: S.P. Uryasev (ed.), *Probabilistic Constrained Optimization*, Kluwer Academic Publishers, pp. 236–251.

Murr, M.R., A. Prékopa (2000). Solution of a product substitution problem using stochastic programming, in: Uryasev, S.P. (ed.), *Probabilistic Constrained Optimization*, Kluwer, Dordrecht.

Nash, S.G., A. Sofer (1993). A barrier-method for large-scale constrained optimization. *ORSA Journal on Computing* 5, 40–53.

Nash, S.G., A. Sofer (1998). Why extrapolation helps barrier methods. Working paper, Department of Operations Research and Engineering, George Mason University, Fairfax, Va.

Norkin V.I., N.V. Roenko (1991). α-concave functions and measures and their applications *Kibernet. Sistem. Anal.* 189, 77–88. (in Russian). Translation in: *Cybernet. Systems Anal.* 27 (1991) 860–869.

Ogryczak, W., A. Ruszczyński (2002). Dual stochastic dominance and related mean risk models. *SIAM Journal on Optimization* 13, 60–78.

Panne, van de C., W. Popp (1963). Minimum Cost Cattle Feed under Probabilistic Problem Constraints. *Management Science* 9, 405–430.

Pedersen, J.G. (1975). On strong unimodality and *m*-ancillarity with applications to contingency tables. *Scandinavian Journal of Statistics* 2, 127–137.

Pickens, J.B., J.G. Hof, B.M. Kent (1991). Use of chance-constrained programming to account for stochastic variation in the A-matrix of large scale linear programs: a forestry application. *Annals of Operations Research* 31, 511–526.

Pintér, J. (1991). Stochastic modelling and optimization for envirenmental management. *Annals of Operations Research* 31, 527–544.

Prékopa, A. (1999). The use of discrete moment bounds in probabilistic constrained stochastic programming models. *Annals of Operations Research* 85, 21–38.

Prékopa, A. (1972a). A class of stochastic programming decision problems. *Matematische Operations forschung und Statistik*, 349–354.

Prékopa, A. (1973b). On logarithmic concave measures and functions. *Acta Scientiarium Mathematicarum (Szeged)* 34, 335–343.

Prékopa, A. (2001a). On the concavity of multivariate probability distributions. *Operations Research Letters* 29, 1–4.

Prékopa, A., B. Vízvári, T. Badics (1998). Programming under probabilistic constraint with discrete random variable, in: Giannessi et al. (ed.), *New Trends in Mathematical Programming*, Kluwer, Dordrecht, Boston, pp. 235–255.

Prékopa, A., P. Kelle (1978). Reliability-type inventory models based on stochastic programming. *Math. Programming Study* 9, 43–58.

Prékopa, A., T. Szántai (1976). On Multi-Stage Stochastic Programming. *Colloquia Mathematica Societas János Bolyai*, North Holland Publishing Company, pp. 733–755.

Prékopa, A., T. Szántai (1978a). A new multivariate gamma distribution and its fitting to empirical streamflow data. *Water Resources Research* 14, 19–24.

Prékopa, A., T. Szántai (1978b). Flood control reservoir system design using stochastic programming. *Math. Programming Study* 9, 138–151.

Prékopa, A., T. Szántai (1979). On optimal regulation of a storage level with application to the water level regulation of a lake. *European Journal of Operations Research* 3, 175–189, also in: *Survey of Math Prog Publ. House of the H. A. S. Budapest*, Vol. II, 183–210.

Prékopa A., B. Vízvári, G. Regös, L. Gao (2001). Bounding the probability of the union of events by the use of aggregation and disaggregation of linear programs. *RUTCOR Research Report* 99–97. Submitted to Discrete Applied Math.

Prékopa, A., T. Rapcsák, I. Zsuffa (1978). Serially linked reservoir system design using stochastic programming. *Water Resources Research* 14, 672–678.

Prékopa, A., S. Ganczer, I. Deák, K. Patyi (1980). The STABIL stochastic programming model and its experimental application to the electrical energy sector of the Hungarian economy, in: M.A.H. Dempster (ed.), *Stochastic Programming*, Academic Press, London, pp. 369–385.

Prékopa, A. (1965). Reliability equation for an inventory problem and its asymptotic solutions. *Colloquium on Appl. of Math. to Economics, Publ. House. of the Hungarian Acad. Sci.*, Budapest, pp. 317–327.

Prékopa, A. (1970) *On Probabilistic Constrained Programming Proceedings of the Princeton Symposium on Math. Prog*, Princeton University Press, Princeton, NJ, pp. 113–138.

Prékopa, A. (1971). Logarithmic concave measures with applications to stochastic programming. *Acta Scientiarium Mathematicarum (Szeged)* 32, 301–316.

Prékopa, A. (1973a). Contributions to the theory of stochastic programming. *Math. Progr.* 4, 202–221.

Prékopa, A. (1973c). Stochastic programming models for inventory control and water storage. *Coll. Math. Soc. J. Bolyai* 7, North-Holland, Amsterdam, pp. 229–245.

Prékopa, A. (1974b). Programming under probabilistic constraints with a random technology matrix. *Matematische Operationsforschung und Statistik, Ser. Optimization* 5, 109–116.

Prékopa, A. (1980a). Logarithmic concave measures and related topics, in: M.A.H. Dempster (ed.), *Stochastic Programming*, Academic Press, London, pp. 63–82.

Prékopa, A. (1980b). Network planning using two-stage programming under uncertainty, in: *Recent Results in Stochastic Programming, Lecture Notes in Economics and Mathematical Systems*, Springer-Verlag, Berlin, pp. 216–237.

Prékopa, A. (1985). Recent results in optimization of electro-energetic systems. Proceedings of the Conference Applied Optimization Techniques in Energy Problems (June 25–29, 1984, Linz, Austrlia, Hj. Wacker, ed.), Teubner, Stuttgart, pp. 354–383.

Prékopa, A. (1988). Boole-Bonferoni inequalities and linear programming. *Operations Research* 36, 145–162.

Prékopa, A. (1988). Numerical solution of probabilistic constrained programming problems *Numerical Techniques for Stochastic Optimization*, Springer, New York, pp. 123–139.

Prékopa, A. (1990). Dual method for the solution of a one-stage stochastic programming problem with random RHS obeying a discrete probality distribution. *ZOR Methods and Models of Operations Research* 34, 441–461.

Prékopa, A. (1990a). Sharp bound on probabilities using linear programming. *Operations Research* 38, 227–239.

Prékopa, A. (1995). *Stochastic Programming*, Kluwer Academic Publishers, Dordrecht, Boston.

Prékopa, A. (2001b). Discrete higher order convex functions and their applications. Generalized convexity and monotonicity, in: N. Hadjisavvas et al. (eds.), *Lecture Notes in Economics and Mathematical Systems*, Springer, Berlin, pp. 21–47.

Raik, E. (1971). The quantile function in stochastic nonlinear programming (in Russian). *Eesti NSV Teaduste Akademia Toimetised (News of the Estonian Academy of Sciences) Füüs. Mat.* 20, 229–231.

Raik, E. (1972). On the stochastic programming problem with the probability and quantile functionals (in Russian). *Eesti NSV Teaduste Akademia Toimetised (News of the Estonian Academy of Sciences) Füüs. Mat.* 21, 142–148.

Raik, E. (1970). Inequalities in problems of stochastic programming (in Russian). *Eesti NSV Teaduste Akademia Toimetised (News of the Estonian Academy of Sciences) Füüs. Mat.* 19, 292–298.

Raik, E. (1971). Qualitative research into the stochastic nonlinear programming problems (in Russian). *Eesti NSV Teaduste Akademia Toimetised (News of the Estonian Academy of Sciences) Füüs. Mat.* 20, 8–14.

Rapcsák, T. (1974). *On the Numerical Solution of a Reservoir Model*. Ph.D. Thesis, University of Debrecen, Hungary (in Hungarian).

Rinott, Y. (1976). On convexity of measures. *Annals of Probability* 4, 1020–1026.

Rockafellar, R.T., S. Uryasev (2000). Optimization of conditional value at risk. *The Journal of Risk* 2, 21–41.

Römisch, W., R. Schultz (1991). Distribution sensitivity for certain classes of chance-constrained models with application to power dispatch. *J. Optimization Theory Appl.* 71, 569–588.

Roy, A.D. (1952). Safety first and the holding of assets. *Econometrica* 20, 431–449.

Salinetti, G. (1983). Approximations for chance constrained programming problems. *Stochastics* 10, 157–169.

Sathe, Y.S., M. Pradhan, S.P. Shah (1980). Inequalities for the probability of the occurence of at least *m* out of *n* events. *J. Appl. Prob.* 17, 1127–1132.

Sen, S. (1992). Relaxations for the probabilistically constrained programs with discrete random variables. *Operations Research Letters* 11, 81–86.

Shapiro, A.F. (1986). Applications of operations research techniques in insurance, in: Goovaerts et al. (eds.), *Insurance and Risk Theory*, pp. 129–143.

Singh, M.R., C.T. Abraham, R. Akella (1990). Wafer design problem in semiconductor manufacturing for reliable customer service. IEEE Transaction on Components. *Hybrids and Manufacturing Technology* 13, 103–108.

Slepian, D. (1962). On the one-sided barrier problem for Gaussian noise. *Bell System Technical Journal* 41, 463–501 .

Stancu-Minasian, I.M. (1984). *Stochastic Programming with Multiple Objective Functions*, Reidel Publishing Company, Boston.

Szántai, T. (1985). On the numerical computation on probabilities in connection with multivariate probability distributions. Thesis, Hung. Acad. Sci. (in Hungarian).

Szántai, T. (1988). A computer code for solution of probabilistic-constrained stochastic programming problems, Y. Ermoliev, R.J.-B. Wets (eds.), *Numerical Techniques for Stochastic Optimization* Springer-Verlag, Berlin, pp. 229–235.

Szántai, T. (1986). Evaluation of a special multivariate gamma distribution. *Mathematical Programming Study* 27, 1–16.

Szántai, T. (2000). Improved bounds and simulation procedures on the value of the multivariate normal probability distribution function. *Annals of Oper. Res.*100, 85–101.

Szántai, T. (2001). Approximation of multivariate probability integrals, in: P.M. Pardalos, C.A. Floudas (eds.) *Encyclopedia of Optimization*, Kluwer Academic Publishers.

Takács, L. (1967). On the method of inclusion and exclusion. *J. of the Amer. Math. Association* 62, 102–113.

Tamm, E. (1977). On *g*-concave functions and probability measures. *Eesti NSV Teaduste Akademia Toimetised (News of the Estonian Academy of Sciences) Füüs. Mat.* 26, 376–379.

Tayur S.L., R.L. Thomas, N.R. Natraj (1995). An algebraic geometry algorithm for scheduling in the presence of setups and correlated demands. *Mathematical Programming* 69, 369–401.

Thoft-Christensen, P., Y. Murotsu (1986). *Applications of Structural Systems Reliability Theory*, Springer-Verlag, New York.

Uryas'ev, S. (1989). A differentiation formula for integrals over sets given by inclusion. *Numerical Functional Analysis and Optimization* 10, 827–841.

Uryasev, S. (2001). Derivatives of probability and integral functions, in: P.M. Pardalos, C. M. Floudas (eds.), *Encyclopedia of Optimization*, Kluwer Academic Publishers.

Vanderbei R.J., D.F. Shanno (2000). Interior-point algorithms for nonconvex nonlinear programming: orderings and higher-order methods, *Math. Prog.* 87, 303–316.

Veinott, A.F. (1967). The supporting hyperplane method for unimodal programming. *Operations Research* 15, 147–152.

Vízvári, B. (1987). *Beiträge zum Frobenius Problem*, D. Sc. Nat. Dissertation, Technische Hochschule Carl Schorlemmer, Leuna-Merseburg, Germany.

Vízvári, B. (2002). The integer programming background of a stochastic integer programming algorithm of Dentcheva-Prékopa-Ruszczyński, *Optimization Methods and Software* 17, 543–559.

Worsley, K.J. (1982). An improved Bonferroni inequality and applications. *Biometrica* 69, 297–302.

Zoutendijk, G. (1960). *Methods of Feasible Directions*. Elsevier Publishing Co. Amsterdam, New York.

A. Ruszczyński and A. Shapiro, Eds., *Handbooks in OR & MS, Vol. 10*

Chapter 6

Monte Carlo Sampling Methods

Alexander Shapiro

School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA

Abstract

In this chapter we discuss Monte Carlo sampling methods for solving large scale stochastic programming problems. We concentrate on the "exterior" approach where a random sample is generated outside of an optimization procedure, and then the constructed, so-called sample average approximation (SAA), problem is solved by an appropriate deterministic algorithm. We study statistical properties of the obtained SAA estimators. The developed statistical inference is incorporated into validation analysis and error estimation. We describe some variance reduction techniques which may enhance convergence of sampling based estimates. We also discuss difficulties in extending this methodology to multistage stochastic programming. Finally, we briefly discuss the SAA method applied to stochastic generalized equations and variational inequalities.

Key words: Two-stage stochastic programming, Monte Carlo sampling, sample average approximation, consistency of estimators, Law of Large Numbers, exponential rates of convergence, conditioning of stochastic problems, validation analysis, variance reduction techniques, multistage stochastic programming, conditional sampling, stochastic generalized equations, variational inequalities.

1 Introduction

Let us consider a stochastic programming problem in the form

$$\operatorname*{Min}_{x \in X} \{f(x) := \mathbb{E}[F(x, \xi)]\}, \tag{1.1}$$

where $F(x, \xi)$ is a function of two vector variables $x \in \mathbb{R}^n$ and $\xi \in \mathbb{R}^d$, $X \subset \mathbb{R}^n$ is a given set and $\boldsymbol{\xi} = \xi(\omega)$ is a random vector. The expectation in (1.1) is taken with respect to the probability distribution of $\boldsymbol{\xi}$ which assumed to be known. In some applications considered in this chapter, the underlying probability space of the elementary events ω will be irrelevant. Therefore, in order to distinguish between random data and their numerical values we often use the bold script like $\boldsymbol{\xi}$ for the random vector $\xi(\omega)$, and ξ for its particular realization (numerical value). We denote by $\Xi \subset \mathbb{R}^d$ the support of the probability distribution of $\boldsymbol{\xi}$, that is, Ξ is the smallest closed set in $\mathbb{R}^d$ such that the probability of the event $\boldsymbol{\xi} \in \mathbb{R}^d \setminus \Xi$ is zero. We denote by $Prob(A)$ or $\mathbb{P}(A)$ the probability of an event A.

Often one can view the optimization problem (1.1) as a two-stage stochastic programming problem with $F(x, \xi)$ and ξ being the optimal value and data vector, respectively, of the corresponding second stage program. For example, in the case of two-stage linear stochastic programming with recourse, $F(x, \xi) := c^T x + Q(x, \xi)$, where $Q(x, \xi)$ is the optimal value of the following second stage problem

$$\operatorname*{Min}_{y \in \mathbb{R}^m} q^T y \quad \text{subject to} \quad Tx + Wy = h,\ y \geq 0, \tag{1.2}$$

with $\xi := (q, T, W, h)$. As such we need to consider situations where $F(x, \xi)$ can take values $+\infty$ or $-\infty$. That is, unless stated otherwise, we assume that $F(x, \xi)$ is an extended real valued function and the expected value $\mathbb{E}[F(x, \boldsymbol{\xi})]$ is *well defined* for every considered $x \in \mathbb{R}^n$ (see the Appendix of chapter "Stochastic Programming Models" for a discussion of the concept of "well defined" expected value). It is also important to notice that it is implicitly assumed in the above setting that for any $x \in X$ and $\xi \in \Xi$ the value $F(x, \xi)$ can be efficiently calculated. This generally holds true for two-stage programming. For multistage programming, however, the situation is more delicate, we will discuss this in Section 6.

If $\boldsymbol{\xi}$ has a finite number of possible realizations (called scenarios), say $\Xi = \{\xi_1, \ldots, \xi_K\}$ with respective (positive) probabilities p_k, $k = 1, \ldots, K$, then we can write the expected value function in the form

$$f(x) = \sum_{k=1}^{K} p_k F(x, \xi_k). \tag{1.3}$$

Note, however, that even a crude discretization of the probability distribution of $\boldsymbol{\xi}$ leads to an exponential growth of the number of scenarios. For example, if components of the random vector $\boldsymbol{\xi}$ are independent, each having just three possible realizations, then the total number of scenarios $K = 3^d$. No computer in a foreseeable future will be able to handle calculations involving 3^{100} scenarios. Therefore, that way or another, one needs to reduce the number of

scenarios to a manageable level. In this chapter we discuss an approach to solving the expected value problem (1.1), referred to as the *true* optimization problem, by using Monte Carlo sampling techniques.

Suppose that we can generate a sample of N replications of the random vector $\boldsymbol{\xi}$. In the Monte Carlo sampling method this is accomplished by generating a random (or rather pseudorandom) sequence $U^1, U^2, \ldots$, of numbers independent of each other and uniformly distributed on the interval $[0, 1]$, and then constructing a sample of $\boldsymbol{\xi}$ by an appropriate transformation. In that way we can consider the sequence $\omega := \{U^1, U^2, \ldots\}$ as an element of the probability space equipped with the corresponding (product) probability measure, and the sample $\boldsymbol{\xi}^i = \boldsymbol{\xi}^i(\omega)$, $i = 1, 2, \ldots$, as a function of ω. We can view the generated sample $\boldsymbol{\xi}^1, \boldsymbol{\xi}^2, \ldots$, as a sequence of random vectors, each having the same probability distribution as $\boldsymbol{\xi}$. If the generated random vectors are (stochastically) independent of each other, we say that the sample is independent identically distributed (iid). By $\xi^1, \xi^2, \ldots$, we denote a particular realization of the considered random sample.

With the generated sample $\xi^1, \ldots, \xi^N$, we associate the sample average function

$$\hat{f}_N(x) := \frac{1}{N} \sum_{i=1}^{N} F(x, \xi^i). \tag{1.4}$$

Note again that for any $x \in X$, the sample average $\hat{f}_N(x)$ can be viewed as a numerical value associated with the generated sample or as a random variable. Which one of these two meanings will be used in a particular situation will be clear from the context.

Since each $\boldsymbol{\xi}^i$ has the same probability distribution as $\boldsymbol{\xi}$, we have that for any $x \in X$, $\mathbb{E}[F(x, \boldsymbol{\xi}^i)] = f(x)$ and hence

$$\mathbb{E}\left[\hat{f}_N(x)\right] = f(x). \tag{1.5}$$

That is, $\hat{f}_N(x)$ is an *unbiased* estimator of $f(x)$. Moreover, under various conditions the Law of Large Numbers (LLN) can be applied[1] with the implication that $\hat{f}_N(x)$ converges with probability one (w.p.1) to $f(x)$ as $N \to \infty$. In that case we say that $\hat{f}_N(x)$ is a *consistent* estimator of $f(x)$. This certainly holds true if the sample is iid.

For the purpose of solving a particular stochastic programming problem, sampling techniques can be applied in different ways. One approach uses sampling in an "interior" fashion. Such algorithms aim at solving the considered problem by resorting to sampling whenever the procedure requires to compute (approximately) the value, and may be derivatives, of the expected

[1] Note again that we allow for $f(x)$ to take values $+\infty$ or $-\infty$.

value function at a current iteration point. Typically such an algorithm is tailored for a specific class of optimization problems and tries to mimic its deterministic counterpart. Often different samples are used each time the true function or its derivatives are estimated at different iteration points. Several such algorithms were suggested with a different level of statistical and convergence analysis.

In this chapter we mainly discuss an alternative approach, referred to as the "exterior" method. First, a sample $\xi^1, \ldots, \xi^N$ is generated, and then the true problem (1.1) is approximated by the optimization problem

$$\operatorname*{Min}_{x \in X} \left\{ \hat{f}_N(x) = \frac{1}{N} \sum_{i=1}^{N} F(x, \xi^i) \right\}. \tag{1.6}$$

Note that once the sample is generated, i.e., numerical values of vectors $\xi^1, \ldots, \xi^N$ are computed, $\hat{f}_N(x)$ becomes a deterministic function and its value can be calculated at any given point $x \in X$. From an optimization point of view, problem (1.6) can be considered as a stochastic programming problem with the finite set $\{\xi^1, \ldots, \xi^N\}$ of scenarios each with equal probability N^{-1}. Therefore, any numerical algorithm suitable for the considered class of problems can be applied to (1.6). The optimal value $\hat{v}_N$ and an optimal solution $\hat{x}_N$ of the problem (1.6) are considered as statistical estimators of their counterparts of the true problem (1.1).

The above approach is called "exterior" since the sample is generated outside of the considered optimization problem, and then the constructed problem (1.6) is solved by an appropriate deterministic algorithm. It should be noted that this method is not an algorithm, but rather a general approach to solving stochastic programs. One still needs to employ a particular (hopefully efficient) deterministic algorithm in order to solve the obtained problem (1.6). The basic idea of the "exterior" approach is simple indeed and the method was suggested by several authors in different contexts under various names. We refer to (1.6) as the *sample average approximation* (SAA) problem. The approach is also known as the sample path or the stochastic counterpart method.

Let us also remark that values of the sample average function $\hat{f}_N(x)$ can be computed in two somewhat different ways. The generated sample $\xi^1, \ldots, \xi^N$ can be stored in the computer memory, and called every time a new value of the sample average function should be computed. In another way the same sample is computed by using the common random number generator at every iteration of the numerical procedure. Which one to use depends on convenience of a particular application.

The idea of common random number generation is well known in simulation. That is, suppose that we want to compare values of the objective function at two points $x_1, x_2 \in X$. In that case we are interested in the

difference $f(x_1)-f(x_2)$ rather than in the individual values $f(x_1)$ and $f(x_2)$. If we use sample average estimates $\hat{f}_N(x_1)$ and $\hat{f}_N(x_2)$ based on *independent* samples, both of size N, then

$$\mathbb{V}\mathrm{ar}\left[\hat{f}_N(x_1)-\hat{f}_N(x_2)\right]=\mathbb{V}\mathrm{ar}\left[\hat{f}_N(x_1)\right]+\mathbb{V}\mathrm{ar}\left[\hat{f}_N(x_2)\right]. \tag{1.7}$$

On the other hand, if we use the *same* sample for the estimators $\hat{f}_N(x_1)$ and $\hat{f}_N(x_2)$, then

$$\mathbb{V}\mathrm{ar}\left[\hat{f}_N(x_1)-\hat{f}_N(x_2)\right]=\mathbb{V}\mathrm{ar}\left[\hat{f}_N(x_1)\right]+\mathbb{V}\mathrm{ar}\left[\hat{f}_N(x_2)\right]-2\mathbb{C}\mathrm{ov}\left(\hat{f}_N(x_1),\hat{f}_N(x_2)\right). \tag{1.8}$$

In both cases, $\hat{f}_N(x_1)-\hat{f}_N(x_2)$ is an unbiased estimator of $f(x_1)-f(x_2)$. However, in the case of the same sample the estimators $\hat{f}_N(x_1)$ and $\hat{f}_N(x_2)$ tend to be positively correlated with each other, in which case the variance in (1.8) is smaller than the one in (1.7). The difference between the independent and the common random number generated estimators of $f(x_1)-f(x_2)$ can be especially dramatic when the points x_1 and x_2 are close to each other and hence the common random number generated estimators are highly correlated.

There are several advantages in the exterior approach as compared with interior methods. One advantage is simplicity of the exterior method. In the following sections we discuss convergence properties of statistical estimators derived by the SAA method. For such statistical analysis a particular numerical algorithm applied to solve the SAA problem is irrelevant.

2 Statistical properties of SAA estimators

In this section we discuss statistical properties of the optimal value $\hat{v}_N$ and the set $\hat{S}_N$ of optimal solutions of the SAA problem (1.6). Unless stated otherwise we assume that the set X is nonempty and closed. We denote by v^* and S the optimal value and the set of optimal solutions, respectively, of the true problem (1.1). For sets $A, B\subset\mathbb{R}^n$ we denote by $\mathrm{dist}(x, A):=\inf_{x'\in A}\|x-x'\|$ the distance from $x\in\mathbb{R}^n$ to A, and by

$$\mathbb{D}(A, B):=\sup_{x\in A}\mathrm{dist}(x, B) \quad \text{and} \quad \mathbb{H}(A, B):=\max\{\mathbb{D}(A, B), \mathbb{D}(B, A)\} \tag{2.1}$$

the deviation of the set A from the set B and the Hausdorff distance between the sets A and B, respectively. By the definition, $\mathrm{dist}(x, A) = +\infty$ if A is empty, and $\mathbb{H}(A, B) = +\infty$ if A or B is empty.

As it was mentioned earlier, we assume that the expected value $f(x)$ is *well defined* for every considered point $x \in \mathbb{R}^n$. That is, $F(x, \cdot)$ is measurable, with respect to the Borel sigma algebra of $\mathbb{R}^d$, and either $\mathbb{E}[F(x, \xi)_+]$ or $\mathbb{E}[-F(x, \xi)_+]$ is finite. If we view the random vector $\xi = \xi(\omega)$ as a measurable mapping defined on a probability space $(\Omega, \mathcal{F}, P)$, then the above measurability requirement can be formulated as the $\mathcal{F}$-measurability of $F(x, \xi(\cdot))$. We also assume that the integrand function $F(x, \xi(\omega))$ is *random lower semicontinuous* (random lsc). Assuming that $\mathcal{F}$ is P-complete, this holds if $F(\cdot, \cdot)$ is measurable, with respect to the Borel sigma algebra of $\mathbb{R}^n \times \mathbb{R}^d$, and $F(\cdot, \xi)$ is lower semicontinuous for every $\xi \in \Xi$. We can also view $\hat{f}_N(x) = \hat{f}_N(x, \omega)$ as a sequence of random functions defined on the same probability space $(\Omega, \mathcal{F}, P)$. The above assumption that F is random lsc implies that the optimal value $\hat{v}_N = \hat{v}_N(\omega)$, of the SAA problem, and the multifunction $\omega \mapsto \hat{S}_N(\omega)$ are measurable. Consequently, $\mathbb{D}(\hat{S}_N, S)$ is measurable and there exists a measurable selection $\hat{x}_N \in \hat{S}_N$. This takes care of measurability of the considered statistical estimators. We refer to the Appendix of chapter "Stochastic Programming Models" for a more detail discussion of these concepts.

Let us observe that the feasible set X can be absorbed into the objective function. That is, define[2] $\overline{F}(x, \xi) := F(x, \xi) + i_X(x)$, i.e.,

$$\overline{F}(x, \xi) := \begin{cases} F(x, \xi), & \text{if } x \in X, \\ +\infty, & \text{if } x \notin X. \end{cases} \tag{2.2}$$

The true problem (1.1) can then be written in the form

$$\operatorname*{Min}_{x \in \mathbb{R}^n} \left\{ \overline{f}(x) := \mathbb{E}[\overline{F}(x, \xi)] \right\}. \tag{2.3}$$

Similarly, the SAA problem (1.6) can be writted as

$$\operatorname*{Min}_{x \in \mathbb{R}^n} \left\{ \tilde{f}_N(x) := \frac{1}{N} \sum_{i=1}^{N} \overline{F}(x, \xi^i) \right\}. \tag{2.4}$$

Note that $\overline{f}(x) = f(x) + i_X(x)$ and $\tilde{f}_N(x) = \hat{f}_N(x) + i_X(x)$.

[2] Recall that i_X denotes the indicator function of the set X, i.e., $i_X(x) = 0$ if $x \in X$ and $i_X(x) = +\infty$ if $x \notin X$.

2.1 Consistency of SAA estimators

We say that $\hat{v}_N$ is a *consistent* estimator of v^* if $\hat{v}_N$ converges w.p.1 to v^* as $N \to \infty$. Similarly, we say that an optimal solution $\hat{x}_N$ of the SAA problem is consistent if $\mathrm{dist}(\hat{x}_N, S)$ tends to zero w.p.1 as $N \to \infty$. If $S = \{x^*\}$ is a singleton, this means that $\hat{x}_N \to x^*$ w.p.1. In the case of nonunique optimal solutions, we would like to ensure that every $\hat{x}_N$ is consistent. That is, we would like to have that $\mathbb{D}(\hat{S}_N, S) \to 0$ w.p.1. We also consider the set S^ε of ε-optimal solutions of the true problem. That is, for $\varepsilon \geq 0$ we say that $\overline{x}$ is an ε-optimal solution of the true problem if $f(\overline{x})$ is finite[3] and $f(\overline{x}) \leq \inf_{x \in X} f(x) + \varepsilon$. Similarly is defined the set $\hat{S}_N^\varepsilon$ of ε-optimal solutions of the SAA problem.

Definition 1. We say that the LLN holds, for $\hat{f}_N(x)$, pointwise if $\hat{f}_N(x)$ converges w.p.1 to $f(x)$, as $N \to \infty$, for any fixed $x \in \mathbb{R}^n$.

If the sample is iid, then the LLN holds[4] pointwise provided that the expected value $f(x)$ is well defined. Unfortunately, such pointwise convergence of $\hat{f}_N(x)$ to $f(x)$ does not necessarily imply convergence of $\hat{v}_N$ to v^*. In order to ensure consistency of $\hat{v}_N$ we need a uniform, or more generally epiconvergence, type of convergence. The analysis is relatively simple in the convex case, which we consider first.

As it was mentioned earlier we can view $\hat{f}_N(x) = \hat{f}_N(x, \omega)$ as a sequence of random functions defined on a common probability space $(\Omega, \mathcal{F}, P)$. Another way of saying that an event happens w.p.1 is to say that it happens for almost every (a.e.) $\omega \in \Omega$, i.e., it happens for all $\omega \in \Omega \setminus \Upsilon$ where Υ is an $\mathcal{F}$-measurable subset of Ω such that $P(\Upsilon) = 0$. We also say that an event happens w.p.1 for N large enough if for a.e. $\omega \in \Omega$ there exists $M(\omega) \in \mathbb{R}$ such that the event happens for all $N \geq M(\omega)$. We say that the functions $\hat{f}_N$ epiconverge to f w.p.1, written $\hat{f}_N \xrightarrow{e} f$ w.p.1, if for a.e. $\omega \in \Omega$ the functions $\hat{f}_N(\cdot, \omega)$ epiconverge[5] to $f(\cdot)$. The following result is a simple consequence of Theorem 25 from Section 8.1 in the Appendix.

Proposition 2. *Suppose that for almost every $\xi \in \Xi$ the function $F(\cdot, \xi)$ is convex, the expected value function $f(\cdot)$ is lower semicontinuous and its domain,* $\mathrm{dom} f$*, has a nonempty interior, and the LLN holds pointwise. Then $\hat{f}_N \xrightarrow{e} f$ w.p.1.*

[3] Note that we assume here that if x^* is an optimal or ε-optimal solution of the true problem, then $f(x^*)$ is finite. Therefore, if the true problem possesses an optimal solution, then v^* is finite. Also v^* is finite iff for any $\varepsilon > 0$ the true problem has an ε-optimal solution.

[4] Often, LLN which ensures convergence w.p.1, is called strong Law of Large Numbers as opposed to weak LLN which ensures only convergence in probability. Since in this chapter we deal only with the strong LLN we omit the word "strong".

[5] See the Appendix for the definition of epiconvergence of a sequence of deterministic functions.

Proof. It follows from the assumed convexity of $F(\cdot, \xi)$ that the function $f(\cdot)$ is convex and that w.p.1 the functions $\hat{f}_N(\cdot)$ are convex. Let us choose a countable and dense[6] subset D of $\mathbb{R}^n$. By the pointwise LLN we have that for any $x \in D$, $\hat{f}_N(x)$ converges to $f(x)$ w.p.1 as $N \to \infty$. This means that there exists a set $\Upsilon_x \subset \Omega$ of P-measure zero such that for any $\omega \in \Omega \backslash \Upsilon_x$, $\hat{f}_N(x, \omega)$ tends to $f(x)$ as $N \to \infty$. Consider the set $\Upsilon := \cup_{x \in D} \Upsilon_x$. Since the set D is countable and $P(\Upsilon_x) = 0$ for every $x \in D$, we have that $P(\Upsilon) = 0$. We also have that for any $\omega \in \Omega \backslash \Upsilon$, $\hat{f}_N(x, \omega)$ converges to $f(x)$, as $N \to \infty$, pointwise on D. It follows then by Theorem 25 that $\hat{f}_N(\cdot, \omega) \overset{e}{\to} f(\cdot)$ for any $\omega \in \Omega \backslash \Upsilon$. That is, $\hat{f}_N(\cdot) \overset{e}{\to} f(\cdot)$ w.p.1. $\square$

As it was mentioned in the above proof, convexity of $f(\cdot)$ follows from convexity of $F(\cdot, \xi)$, $\xi \in \Xi$. Also, by Fatou's lemma, lower semicontinuity of $f(\cdot)$ is implied by lower semicontinuity of $F(\cdot, \xi)$, $\xi \in \Xi$, under the additional assumption that $F(x, \cdot)$ is bounded from below by an integrable function (see the Appendix of chapter "Stochastic Programming Models").

By the assertion (iii) of Theorem 25 we obtain the following corollary.

Corollary 3. *Suppose that for almost every $\xi \in \Xi$ the function $F(\cdot, \xi)$ is convex and the LLN holds pointwise. Let C be a compact subset of $\mathbb{R}^n$ such that $f(\cdot)$ is finite valued on a neighborhood of C. Then $\hat{f}_N \overset{e}{\to} f$ converges to f uniformly on C, that is*

$$\sup_{x \in C} |\hat{f}_N(x) - f(x)| \to 0 \quad \text{w.p.1 as } N \to \infty. \tag{2.5}$$

By the result (8.3) of Proposition 24 we have that $\hat{f}_N \overset{e}{\to} f$, w.p.1, implies that

$$\limsup_{N \to \infty} \hat{v}_N \leq v^*, \quad \text{w.p.1}. \tag{2.6}$$

Without an additional assumption the inequality in (2.6) can be strict. In addition to convexity we need a boundedness type condition to ensure consistency of $\hat{v}_N$.

Theorem 4. *Suppose that:* (i) *the integrand function F is random lower semicontinuous*, (ii) *for almost every $\xi \in \Xi$ the function $F(\cdot, \xi)$ is convex*, (iii) *the set X is closed and convex*, (iv) *the expected value function f is lower semicontinuous and there exists a point $\overline{x} \in X$ such that $f(x) < +\infty$ for all x in a neighborhood of $\overline{x}$*, (v) *the set S of optimal solutions of the true problem is nonempty and bounded*, (vi) *the LLN holds pointwise. Then $\hat{v}_N \to v^*$ and $\mathbb{D}(\hat{S}_N, S) \to 0$ w.p.1 as $N \to \infty$.*

[6] It is said that D is a dense subset of $\mathbb{R}^n$ if for any point $x \in \mathbb{R}^n$ and $\varepsilon > 0$ there exists a point $x' \in D$ such that $||x - x'|| < \varepsilon$.

Proof. Clearly we can restrict both the true and the SAA problems to the affine space generated by the convex set X. Relative to that affine space the set X has a nonempty interior. Therefore, without loss of generality we can assume that the set X has a nonempty interior. Since it is assumed that $f(x)$ possesses an optimal solution, we have that v^* is finite and hence $f(x) \geq v^* > -\infty$ for all $x \in X$. Since $f(x)$ is convex and is greater than $-\infty$ on an open set (e.g., interior of X), it follows that $f(x)$ is proper.

Now let $\overline{f}(x)$ and $\tilde{f}_N(x)$ be extended real valued functions defined in (2.3) and (2.4), respectively. Observe that the pointwise LLN for $F(x, \xi)$ (assumption (vi)) implies the corresponding pointwise LLN for $\overline{F}(x, \xi)$. Since X is convex and closed, it follows that $\overline{f}$ is convex and lower semicontinuous. Moreover, because of the assumption (iv) and since the interior of X is nonempty, we have that $\operatorname{dom} \overline{f}$ has a nonempty interior. By Proposition 2 it follows then that $\tilde{f}_N \overset{e}{\to} \overline{f}$ w.p.1. Consider a compact set K with a nonempty interior and such that it does not contain a boundary point of $\operatorname{dom} \overline{f}$, and $\overline{f}(x)$ is finite valued on K. Since $\operatorname{dom} \overline{f}$ has a nonempty interior such set exists. Then it follows from $\tilde{f}_N \overset{e}{\to} \overline{f}$, that $\tilde{f}_N(\cdot)$ converge to $\overline{f}(\cdot)$ uniformly on K, all w.p.1 (see Theorem 25). It follows that w.p.1 for N large enough the functions $\tilde{f}_N(x)$ are finite valued on K, and hence are proper.

Now let C be a compact subset of $\mathbb{R}^n$ such that the set S is contained in the interior of C. Such set exists since it is assumed that the set S is bounded. Consider the set $\tilde{S}_N$ of minimizers of $\tilde{f}_N(x)$ over C. Since C is nonempty and compact and $\tilde{f}_N(x)$ is lower semicontinuous and proper for N large enough, and because by the pointwise LLN we have that for any $x \in S$, $\tilde{f}_N(x)$ is finite w.p.1 for N large enough, the set $\tilde{S}_N$ is nonempty w.p.1 for N large enough. Let us show that $\mathbb{D}(\tilde{S}_N, S) \to 0$ w.p.1. Let $\omega \in \Omega$ be such that $\tilde{f}_N(\cdot, \omega) \overset{e}{\to} \overline{f}(\cdot)$. We have that this happens for a.e. $\omega \in \Omega$. We argue now by a contradiction. Suppose that there exists a minimizer $\tilde{x}_N = \tilde{x}_N(\omega)$ of $\tilde{f}_N(x, \omega)$ over C such that $\operatorname{dist}(\tilde{x}_N, S) \geq \varepsilon$ for some $\varepsilon > 0$. Since C is compact, by passing to a subsequence if necessary, we can assume that $\tilde{x}_N$ tends to a point $\overline{x} \in C$. It follows that $\overline{x} \notin S$. On the other hand, we have by Proposition 24 that $\overline{x} \in \arg\min_{x \in C} \overline{f}(x)$. Since $\arg\min_{x \in C} \overline{f}(x) = S$, we obtain a contradiction.

Now because of the convexity assumptions, any minimizer of $\tilde{f}_N(x)$ over C which lies inside the interior of C, is also an optimal solution of the SAA problem (2.4). Therefore, w.p.1 for N large enough we have that $\tilde{S}_N = \hat{S}_N$. Consequently, we can restrict both the true and the SAA optimization problems to the compact set C, and hence the assertions of the above proposition follow. □

Let us make the following observations. It was assumed in the above proposition that the LLN holds pointwise for all $x \in \mathbb{R}^n$. Actually it suffices to assume that this holds for all x in some neighborhood of the set S. Under the assumptions of the above theorem we have that $f(x) > -\infty$ for every $x \in \mathbb{R}^n$. The above assumptions do not prevent, however, for $f(x)$ to take value $+\infty$ at some points $x \in X$. Nevertheless, it was possible to push the proof through

because in the considered convex case local optimality implies global optimality. There are two possible reasons why $f(x)$ can be $+\infty$. Namely, it can be that $F(x, \cdot)$ is finite valued but grows sufficiently fast so that its integral is $+\infty$, or it can be that $F(x, \cdot)$ is equal $+\infty$ on a set of positive measure, and of course it can be both. For example, in the case of two-stage programming it may happen that for some $x \in X$ the corresponding second stage problem is infeasible with a positive probability p. Then w.p.1 for N large enough, for at least one of the sample points ξ^i the corresponding second stage problem will be infeasible, and hence $\hat{f}_N(x) = +\infty$. Of course, if the probability p is very small, then the required sample size for such event to happen could be very large.

Theorem 4 shows that in the convex case consistency of SAA estimators follows from the pointwise LLN and natural boundedness conditions. Without convexity the epiconvergence analysis becomes more involved. We give below convergence results based on uniform convergence which often are sufficient for practical applications.

Proposition 5. *Suppose that $\hat{f}_N(x)$ converges to $f(x)$ w.p.1, as $N \to \infty$, uniformly on X. Then $\hat{v}_N$ converges to v^* w.p.1 as $N \to \infty$.*

Proof. The uniform convergence of $\hat{f}_N(x)$ to $f(x)$ w.p.1 means that for any $\varepsilon > 0$ the following inequality holds w.p.1 for N large enough,

$$\sup_{x \in X} |\hat{f}_N(x) - f(x)| \leq \varepsilon. \tag{2.7}$$

It follows then that $|\hat{v}_N - v^*| \leq \varepsilon$ w.p.1 for N large enough, which completes the proof. □

Proposition 6. *Suppose that there exists a compact set $C \subset \mathbb{R}^n$ such that:* (i) *the set S of optimal solutions of the true problem is nonempty and is contained in C,* (ii) *the function $f(x)$ is finite valued and continuous on C,* (iii) *$\hat{f}_N(x)$ converges to $f(x)$ w.p.1, as $N \to \infty$, uniformly in $x \in C$,* (iv) *w.p.1 for N large enough the set $\hat{S}_N$ is nonempty and $\hat{S}_N \subset C$. Then $\hat{v}_N \to v^*$ and $\mathbb{D}(\hat{S}_N, S) \to 0$ w.p.1 as $N \to \infty$.*

Proof. Assumptions (i) and (iv) imply that both the true and SAA problems can be restricted to the set C. It can be easily verified that assumptions (ii) and (iii) imply that the functions $\hat{f}_N$ restricted to $X \cap C$ epiconverge to the function f restricted to $X \cap C$, i.e., $\hat{f}_N + i_{X \cap C} \overset{e}{\to} f + i_{X \cap C}$, w.p.1. Since C is compact, the proof then can be completed in a way similar to the proof of Theorem 4. □

The last assumption (iv) in the above proposition holds, in particular, if the feasible set X is closed, the functions $\hat{f}_N(x)$ are lower semicontinuous and for some $\alpha > v^*$ the level sets $\{x \in X : \hat{f}_N(x) \leq \alpha\}$ are uniformly bounded w.p.1.

There is a variety of results on uniform LLN (assumption (iii) of the above proposition). In the convex case this is ensured by the assumptions of Corollary 3. Following is a relatively simple uniform LLN without the convexity assumption. We say that $F(x, \xi)$, $x \in C$, is *dominated* by an integrable function if there exists a nonnegative valued measurable function $G(\xi)$ such that $\mathbb{E}[G(\boldsymbol{\xi})] < +\infty$ and for every $x \in C$ the inequality $|F(x, \boldsymbol{\xi})| \leq G(\boldsymbol{\xi})$ holds w.p.1.

Proposition 7. *Let C be a nonempty compact subset of $\mathbb{R}^n$ and suppose that:* (i) *for almost every $\xi \in \Xi$ the function $F(\cdot, \xi)$ is continuous on C*, (ii) *$F(x, \xi)$, $x \in C$, is dominated by an integrable function*, (iii) *the sample is* iid. *Then the expected value function $f(x)$ is finite valued and continuous on C, and $\hat{f}_N(x)$ converges to $f(x)$ w.p.1 uniformly on C.*

Proof. It follows from the assumption (ii) that $|f(x)| \leq \mathbb{E}[G(\boldsymbol{\xi})]$, and consequently $|f(x)| < +\infty$ for all $x \in C$. Consider a point $x \in C$ and let x_k be a sequence of points in C converging to x. By the Lebesgue Dominated Convergence Theorem assumption (ii) implies that

$$\lim_{k\to\infty} \mathbb{E}[F(x_k, \boldsymbol{\xi})] = \mathbb{E}\left[\lim_{k\to\infty} F(x_k, \boldsymbol{\xi})\right].$$

Since by (i), $F(x_k, \boldsymbol{\xi}) \to F(x, \boldsymbol{\xi})$ w.p.1, it follows that $f(x_k) \to f(x)$, and hence $f(x)$ is continuous.

Choose now a point $\overline{x} \in C$, a sequence γ_k of positive numbers converging to zero, and define $V_k := \{x \in C\colon \|x - \overline{x}\| \leq \gamma_k\}$ and

$$\delta_k(\xi) := \sup_{x \in V_k} |F(x, \xi) - F(\overline{x}, \xi)|. \tag{2.8}$$

By the assumption (i) we have that for any $\xi \in \Xi$, $\delta_k(\xi)$ tends to zero as $k \to \infty$. Moreover, by the assumption (ii) we have that $\delta_k(\xi)$, $k = 1, \ldots$, are dominated by an integrable function, and hence by the Lebesgue Dominated Convergence Theorem we have that

$$\lim_{k\to\infty} \mathbb{E}[\delta_k(\boldsymbol{\xi})] = \mathbb{E}\left[\lim_{k\to\infty} \delta_k(\boldsymbol{\xi})\right] = 0. \tag{2.9}$$

We also have that

$$|\hat{f}_N(x) - \hat{f}_N(\overline{x})| \leq \frac{1}{N} \sum_{i=1}^{N} |F(x, \boldsymbol{\xi}^i) - F(\overline{x}, \boldsymbol{\xi}^i)|,$$

and hence

$$\sup_{x\in V_k} |\hat{f}_N(x) - \hat{f}_N(\overline{x})| \leq \frac{1}{N}\sum_{i=1}^{N} \delta_k(\boldsymbol{\xi}^i). \tag{2.10}$$

Since the sample $\boldsymbol{\xi}^i$ is iid, it follows by the LLN that the right hand side of (2.10) converges w.p.1 to $\mathbb{E}[\delta_k(\boldsymbol{\xi})]$ as $N \to \infty$. Together with (2.9) this implies that for any given $\varepsilon > 0$ there exists a neighborhood W of $\overline{x}$ such that w.p.1 for sufficiently large N,

$$\sup_{x\in W\cap C} |\hat{f}_N(x) - \hat{f}_N(\overline{x})| < \varepsilon.$$

Since C is compact, there exists a finite number of points $x_1, \ldots, x_m \in C$ and corresponding neighborhoods $W_1, \ldots, W_m$ covering C such that w.p.1 for N large enough the following holds

$$\sup_{x\in W_j\cap C} |\hat{f}_N(x) - \hat{f}_N(x_j)| < \varepsilon, \quad j = 1, \ldots, m. \tag{2.11}$$

Furthermore, since $f(x)$ is continuous on C, these neighborhoods can be chosen in such a way that

$$\sup_{x\in W_j\cap C} |f(x) - f(x_j)| < \varepsilon, \quad j = 1, \ldots, m. \tag{2.12}$$

Again by the LLN we have that $\hat{f}_N(x)$ converges pointwise to $f(x)$ w.p.1. Therefore,

$$|\hat{f}_N(x_j) - f(x_j)| < \varepsilon, \quad j = 1, \ldots, m, \tag{2.13}$$

w.p.1 for N large enough. It follows from (2.11)–(2.13) that w.p.1 for N large enough

$$\sup_{x\in C} |\hat{f}_N(x) - f(x)| < 3\varepsilon. \tag{2.14}$$

Since $\varepsilon > 0$ was arbitrary, we obtain that (2.5) follows and hence the proof is complete. □

It is possible to extend the above result in various directions. For example, the assumption that the sample is iid was used in the proof in two places,

namely, to ensure pointwise convergence w.p.1 of $\hat{f}_N(x)$ to $f(x)$ (i.e., pointwise LLN) and applicability of the LLN to the right hand side of (2.10). As it was mentioned earlier, there is a variety of results, in the probability theory, on the (pointwise) LLN without the iid assumption, which can be applied to the present case.

Remark 8. We assumed that the feasible set X in the SAA problem (1.6) is fixed, i.e., independent of the sample. However, in some situations it also should be estimated. Then the corresponding SAA problem takes the form

$$\operatorname*{Min}_{x \in X_N} \hat{f}_N(x), \tag{2.15}$$

where X_N is a subset of $\mathbb{R}^n$ depending on the sample, and therefore is random. Suppose that in addition to the assumptions of Proposition 6 the following two conditions hold with probability one:

(a) If $x_N \in X_N$ and x_N converges w.p.1 to a point x, then $x \in X$.
(b) For some point $x \in S$ there exists a sequence $x_N \in X_N$ such that $x_N \to x$ w.p.1.

Under the assumption of the uniform convergence of $\hat{f}_N$ to f, the above conditions (a) and (b) ensure conditions (i) and (ii) of the (epiconvergence) Definition 23 as applied to the functions $\hat{f}_N + i_{X_N \cap C}$ and $f + i_{X \cap C}$ at the point x. Consistency of the SAA estimators then follows.

Suppose, for example, that the set X is defined by the constraints

$$X := \{x \in X_0 : g_j(x) \leq 0, j = 1, \ldots, p\}, \tag{2.16}$$

where X_0 is a nonempty closed subset of $\mathbb{R}^n$ and the constraint functions are given as the expected value functions

$$g_j(x) := \mathbb{E}[G_j(x, \xi)], \quad j = 1, \ldots, p, \tag{2.17}$$

with G_j being random lsc functions. Then the set X can be estimated by

$$X_N := \{x \in X_0 : \hat{g}_{jN}(x) \leq 0, j = 1, \ldots, p\}, \tag{2.18}$$

where $\hat{g}_{jN}(x) := N^{-1} \sum_{i=1}^{N} G_j(x, \xi^i)$. If for a given point $x \in X_0$ we have that $\hat{g}_{jN}$ converge uniformly to g_j w.p.1 on a neighborhood of x and the functions g_j are continuous, then the above condition (a) holds.

In order to ensure condition (b) one needs to impose a constraint qualification (on the true problem). Consider, for example, $X := \{x \in \mathbb{R} : g(x) \leq 0\}$ with $g(x) := x^2$. Clearly $X = \{0\}$, while an arbitrary small perturbation of the

function $g(\cdot)$ can result in the corresponding set X_N being empty. It is possible to show that if a constraint qualification for the true problem is satisfied at x, then condition (b) follows. For instance, if the set X_0 is convex and for every $\xi \in \Xi$ the functions $G_j(\cdot, \xi)$ are convex, and hence the corresponding expected value functions $g_j(\cdot)$, $j = 1, \ldots, p$, are also convex, then such a simple constraint qualification is the Slater condition. Recall that it is said that the *Slater condition* holds if there exists a point $\overline{x} \in X_0$ such that $g_j(\overline{x}) < 0$, $j = 1, \ldots, p$.

2.2 Asymptotics of the SAA optimal value

Consistency of the SAA estimators gives us a certain assurance that the error of the estimation approaches zero in the limit as the sample size grows to infinity. Although this is important conceptually, it does not give any indication of the magnitude of the error for a chosen sample size N. Suppose for a moment that the sample is iid and let us fix a point $x \in X$. Then we have that the sample average estimator $\hat{f}_N(x)$, of $f(x)$, is unbiased and has variance $\sigma^2(x)/N$, where $\sigma^2(x) := \mathbb{V}\mathrm{ar}[F(x, \xi)]$ is supposed to be finite. Moreover, by the Central Limit Theorem (CLT) we have that

$$N^{1/2}\left[\hat{f}_N(x) - f(x)\right] \Rightarrow Y(x), \tag{2.19}$$

where "$\Rightarrow$" denotes convergence in *distribution* and $Y(x)$ has a normal distribution with mean 0 and variance $\sigma^2(x)$, written $Y(x) \sim N(0, \sigma^2(x))$. That is, for large N, $\hat{f}_N(x)$ has approximately normal distribution with mean $f(x)$ and variance $\sigma^2(x)/N$.

This leads to the following (approximate) $100(1-\alpha)\%$ confidence interval for $f(x)$:

$$\left[\hat{f}_N(x) - \frac{z_{\alpha/2}\hat{\sigma}(x)}{\sqrt{N}}, \hat{f}_N(x) + \frac{z_{\alpha/2}\hat{\sigma}(x)}{\sqrt{N}}\right], \tag{2.20}$$

where $z_{\alpha/2} := \Phi^{-1}(1-\alpha/2)$ and[7]

$$\hat{\sigma}^2(x) := \frac{1}{N-1}\sum_{i=1}^{N}\left[F(x, \xi^i) - \hat{f}_N(x)\right]^2 \tag{2.21}$$

is the sample variance estimate of $\sigma^2(x)$. That is, the error of estimation of $f(x)$ is (stochastically) of order $O_p(N^{-1/2})$. The involved constant, which is proportional to $\sigma(x)$, can be reduced (sometimes significantly) by variance

[7] Here $\Phi(\cdot)$ denotes the cdf of the standard normal distribution. For example, to 95% confidence intervals corresponds $z_{0.025} = 1.96$.

reduction techniques. However, the basic rate of $O_p(N^{-1/2})$ is characteristic for Monte Carlo sampling and cannot be changed.

Consider now the optimal value $\hat{v}_N$ of the SAA problem (1.6). Clearly we have that for any $x' \in X$ the inequality $\hat{f}_N(x') \geq \inf_{x \in X} \hat{f}_N(x)$ holds. By taking the expected value of both sides of this inequality and minimizing the left hand side over all $x' \in X$ we obtain

$$\inf_{x \in X} \mathbb{E}\left[\hat{f}_N(x)\right] \geq \mathbb{E}\left[\inf_{x \in X} \hat{f}_N(x)\right]. \tag{2.22}$$

Since $\mathbb{E}[\hat{f}_N(x)] = f(x)$, it follows that $v^* \geq \mathbb{E}[\hat{v}_N]$. In fact, typically, $\mathbb{E}[\hat{v}_N]$ is strictly less than v^*, i.e., $\hat{v}_N$ is a downwards *biased* estimator of v^*. Note that the inequality (2.22) holds even if $f(x) = +\infty$ or $f(x) = -\infty$ for some $x \in X$.

Let us discuss now the following simple example which demonstrates various basic properties of the SAA estimators.

Example 9. Consider the function $F(x, \xi) := |x - \xi|$, with $x, \xi \in \mathbb{R}$, and $X := \mathbb{R}$. In that case optimal solutions of the true and SAA problems are given by the true (population) and the sample medians, respectively. Suppose, first, that the corresponding random variable $\boldsymbol{\xi}$ has a discrete distribution with $\mathbb{P}(\boldsymbol{\xi} = -1) = \mathbb{P}(\boldsymbol{\xi} = 1) = 1/2$. Then

$$f(x) = \frac{1}{2}(|x - 1| + |x + 1|),$$

and hence $v^* = 1$ and $S = [-1, 1]$. Let $\xi^1, \ldots, \xi^N$ be an iid sample of $\boldsymbol{\xi}$, i.e., each ξ^i can take value -1 or 1 with probability half and they are independent of each other. Then

$$\hat{f}_N(x) = p_N|x - 1| + (1 - p_N)|x + 1|,$$

where p_N is the proportion of times that $\xi^i = 1$. It follows that if $p_N < 1/2$, then $\hat{S}_N = \{-1\}$ and $\hat{v}_N = 2p_N$. If $p_N > 1/2$, then $\hat{S}_N = \{1\}$ and $\hat{v}_N = 2(1 - p_N)$, and finally if $p_N = 1/2$, then $\hat{S}_N = [-1, 1]$ and $\hat{v}_N = 1$. That is, $\hat{v}_N = \min\{2p_N, 2(1 - p_N)\}$. Now by the CLT we have that $N^{1/2}(2p_N - 1)$ converges in distribution to $N(0, 1)$. It follows that $N^{1/2}(\hat{v}_N - 1)$ converges in distribution to $W := -|Z|$, where $Z \sim N(0, 1)$. In general, convergence in distribution does not imply convergence of the corresponding expected values. However, in the present case the uniform integrability condition[8] can be verified, and hence we obtain that

$$\mathbb{E}[\hat{v}_N] = v^* - N^{-1/2}C + o(N^{-1/2}), \tag{2.23}$$

[8] See the Appendix, (8.5) in particular, for a definition of the uniform integrability condition.

where $C := \mathbb{E}[Z] = \sqrt{2/\pi}$. That is, the estimator $\hat{v}_N$ has a negative bias of order $O(N^{-1/2})$.

Suppose now that ξ can take three values -1, 0 and 1 with equal probabilities 1/3. In that case $v^* = 2/3$ and $S = \{0\}$, i.e., the true optimization problem has unique optimal solution $x^* = 0$. The SAA estimator $\hat{x}_N$ can be equal to -1, 0, or 1. Moreover, the event $\{\hat{x}_N = 1\}$ happens if more than half of the sample points are equal to one. Probability of that is given by $\mathbb{P}(W > N/2)$, where W has a binomial distribution $B(N, 1/3)$. If exactly half of the sample points are equal to one, then the sample estimate can be any number in the interval $[0, 1]$. Similar conclusions hold for the event $\{\hat{x}_N = -1\}$. Therefore, the probability that $\hat{x}_N = 0$ is at least $1 - 2\mathbb{P}(W \geq N/2)$. By the Large Deviations theory (see Section 8.3 in the Appendix) we have that $\mathbb{P}(W > N/2) \leq e^{-\beta N}$, with the (best) exponential constant β can be calculated to be $\beta = 0.059$. It follows that

$$1 - \mathbb{P}(\hat{x}_N = 0) \leq 2e^{-\beta N}. \tag{2.24}$$

That is, the probability that the SAA estimator $\hat{x}_N$ is equal exactly to the true optimal solution x^* approaches one exponentially fast. Consequently, it happens with probability approaching one exponentially fast that $\hat{v}_N = \hat{f}_N(x^*)$. Since $\hat{f}_N(x^*)$ is an unbiased estimator of v^*, it follows then that the bias $\mathbb{E}[\hat{v}_N] - v^*$, although is negative, approaches zero very fast as N increases.

Finally, suppose that the probability distribution of ξ is continuous, i.e., its cumulative distribution function (cdf) $G(\cdot)$ is continuous. Then the expected value function $f(x)$ is continuously differentiable with $\mathrm{d}f(x)/\mathrm{d}x = 2G(x) - 1$. Suppose further that $G(\cdot)$ is differentiable at x^* with $g(x^*) := \mathrm{d}G(x^*)/\mathrm{d}x > 0$. Then $\mathrm{d}^2 f(x^*)/\mathrm{d}x^2 = 2g(x^*)$ is positive, and hence, since $f(x)$ is convex, it follows that the minimizer x^* is unique. Also, it is well known that in this case $N^{1/2}(\hat{x}_N - x^*)$ converges in distribution to normal with zero mean and variance $[2g(x^*)]^{-2}$. That is, $\hat{x}_N$ tends to x^* at a stochastic rate of $O_p(N^{-1/2})$. By the Mean Value Theorem we have that

$$\hat{f}_N(\hat{x}_N) - \hat{f}_N(x^*) = \gamma_N(\hat{x}_N - x^*),$$

where γ_N is a subgradient of $\hat{f}_N$ calculated at some point $\tilde{x}_N \in [x^*, \hat{x}_N]$. By the uniform LLN we also have that $|\gamma_N - \mathrm{d}f(\tilde{x}_N)/\mathrm{d}x|$ tends w.p.1, and hence in probability, to zero. Moreover, since x^* is a minimizer of $f(x)$, we have that $\mathrm{d}f(x^*)/\mathrm{d}x = 0$, and hence it follows that $\gamma_N = o_p(1)$. Consequently, we obtain that $\hat{f}_N(\hat{x}_N) - \hat{f}_N(x^*) = o_p(N^{-1/2})$, and hence

$$\hat{v}_N = \hat{f}_N(\hat{x}_N) = \hat{f}_N(x^*) + o_p(N^{-1/2}). \tag{2.25}$$

Together with (2.19) this implies that $N^{1/2}(\hat{v}_N - v^*) \Rightarrow N(0, \sigma^2(x^*))$. Typically in such cases the bias $\mathbb{E}[\hat{v}_N] - v^*$ tends to zero at a rate of $O(N^{-1})$.

Let us consider now the general case. Let $Y(x)$ be random variables defined in (2.19). Recall that $Y(x) \sim N(0, \sigma^2(x))$ and that the covariance between $Y(x_1)$ and $Y(x_2)$, for some $x_1, x_2 \in X$, is the same as the covariance between $F(x_1, \xi)$ and $F(x_2, \xi)$. We use the following assumptions about the integrand F:

(A1) For some point $x \in X$ the expectation $\mathbb{E}[F(x, \xi)^2]$ is finite.

(A2) There exists a measurable function $K: \Xi \to \mathbb{R}_+$ such that $\mathbb{E}[K(\xi)^2]$ is finite and

$$|F(x_1, \xi) - F(x_2, \xi)| \leq K(\xi)\|x_1 - x_2\|, \tag{2.26}$$

for all $x_1, x_2 \in X$ and $\xi \in \Xi$.

The above assumptions imply that the expected value $f(x)$ and variance $\sigma^2(x)$ are finite valued for all $x \in X$. Moreover, it follows from (2.26) that $|f(x_1) - f(x_2)| \leq \kappa\|x_1 - x_2\|$, where $\kappa := \mathbb{E}[K(\xi)]$, and hence $f(x)$ is Lipschitz continuous on X. If X is compact, we have then that the set S, of minimizers of $f(x)$ over X, is nonempty.

Theorem 10. *Suppose that the sample is* iid, *the set X is compact and assumptions* (A1) *and* (A2) *are satisfied. Then the following holds*

$$\hat{v}_N = \inf_{x \in S} \hat{f}_N(x) + o_p(N^{-1/2}), \tag{2.27}$$

$$N^{1/2}(\hat{v}_N - v^*) \Rightarrow \inf_{x \in S} Y(x). \tag{2.28}$$

Proof of (2.27) is based on upper and lower estimates, while (2.28) follows from (2.27). It is clear that $\hat{v}_N \leq \inf_{x \in S} \hat{f}_N(x)$, which leads to the required upper bound. Derivation of the lower bound is more involved and is based on a functional CLT, and will be not given here.

Under mild additional conditions (see Section 8.2 in the Appendix) it follows from (2.28) that

$$\mathbb{E}[\hat{v}_N] - v^* = N^{-1/2}\mathbb{E}\left[\inf_{x \in S} Y(x)\right] + o(N^{-1/2}). \tag{2.29}$$

In particular, it follows from (2.28) that if $S = \{x^*\}$ is a singleton, then

$$N^{1/2}(\hat{v}_N - v^*) \Rightarrow N(0, \sigma^2(x^*)). \tag{2.30}$$

Moreover, since $\mathbb{E}[Y(x^*)] = 0$, we obtain that in this case the bias $\mathbb{E}[\hat{v}_N] - v^*$ is of order $o(N^{-1/2})$. On the other hand, if the true problem has more than one optimal solution, then the right hand side of (2.28) is given by the minimum of a number of random variables. Although each $Y(x)$ has mean zero, their minimum typically has a negative mean. Therefore, if S is not a singleton, then the bias $\mathbb{E}[\hat{v}_N] - v^*$ typically is negative and of order $O(N^{-1/2})$. Moreover, the bias tends to be bigger the larger the set S is.

Suppose now that the feasible set X is defined by constraints in the form (2.16). The Lagrangian function of the true problem is $L(x, \lambda) := f(x) + \sum_{j=1}^{p} \lambda_j g_j(x)$. Suppose also that the problem is *convex*, that is the set X_0 is convex and for all $\xi \in \Xi$ the functions $F(\cdot, \xi)$ and $G_j(\cdot, \xi)$, $j = 1, \ldots, p$, are convex. Suppose, further, that the functions $f(x)$ and $g_j(x)$ are finite valued on a neighborhood of S and the Slater condition holds. Then with every optimal solution $x^* \in S$ is associated a nonempty and bounded set Λ of Lagrange multipliers vectors $\lambda = (\lambda_1, \ldots, \lambda_p)$ satisfying the optimality conditions:

$$x^* \in \arg\min_{x \in X_0} L(x, \lambda), \quad \lambda_j \geq 0 \quad \text{and} \quad \lambda_j g_j(x^*) = 0, j = 1, \ldots, p. \tag{2.31}$$

The set Λ coincides with the set of optimal solutions of the dual of the true problem, and therefore is the same for any optimal solution $x^* \in S$.

Let $\hat{v}_N$ be the optimal value of the SAA problem (2.15) with X_N given in the form (2.18). We use the assumptions (A1) and (A2), given before Theorem 10, applied to the integrands G_j as well as to the integrand F. It follows then that the functions $f(x)$ and $g_j(x)$ are finite valued and continuous on X. As in Theorem 10, we denote by $Y(x)$ random variables which are normally distributed and have the same covariance structure as $F(x, \xi)$. We also denote by $Y_j(x)$ random variables which are normally distributed and have the same covariance structure as $G_j(x, \xi)$, $j = 1, \ldots, p$.

Theorem 11. *Suppose that the sample is* iid, *the problem is convex and the following conditions are satisfied:* (i) *the set X is compact*, (ii) *the functions $f(x)$ and $g_j(x)$ are finite valued on a neighborhood of S*, (iii) *the Slater condition for the true problem holds*, (iv) *the assumptions* (A1) *and* (A2) *hold for the integrands F and G_j, $j = 1, \ldots, p$. Then*

$$N^{1/2}(\hat{v}_N - v^*) \Rightarrow \inf_{x \in S} \sup_{\lambda \in \Lambda} \left[Y(x) + \sum_{j=1}^{p} \lambda_j Y_j(x) \right]. \tag{2.32}$$

Proof of the above theorem is based on duality theory of convex programming and a functional CLT, and will be not given here. It follows

from (2.32) that if $S=\{x^*\}$ and $\Lambda=\{\overline{\lambda}\}$ are singletons, i.e., both the true and its dual problems have unique optimal solutions, then

$$N^{1/2}(\hat{v}_N - v^*) \Rightarrow N(0, \sigma_0^2), \tag{2.33}$$

with $\sigma_0^2 := \mathbb{V}\mathrm{ar}[F(x^*, \xi) + \sum_{j=1}^{p} \overline{\lambda}_j G_j(x^*, \xi)]$.

3 Exponential rates of convergence

In this section we discuss exponential rates of convergence of optimal and nearly optimal SAA solutions. This allows to give an estimate of the sample size which is required to solve the true problem with a given accuracy by solving the SAA problem. Although such estimates of the sample size typically are too conservative for a practical use, they give an insight into the complexity of solving the true (expected value) problem. The analysis is based on the Large Deviations theory.[9] For the sake of simplicity, unless stated otherwise, we assume in this section that the random sample is iid.

3.1 The case of finite feasible set

In this section we consider cases where[10] $|X| < \infty$, i.e., the set X is finite, although may be very large. For the sake of simplicity we assume that $F(x, \xi)$ is finite for all $x \in X$ and $\xi \in \Xi$ and $f(x)$ is finite for every $x \in X$. Then the sets $\hat{S}_N^{\varepsilon}$ and S^{ε} of ε-optimal solutions of the SAA and true problems, respectively, are nonempty and, of course, finite for any $\varepsilon \geq 0$. Consider numbers $\varepsilon \geq 0$, $\delta \geq 0$ with $\delta \leq \varepsilon$, and the event $\{\hat{S}_N^{\delta} \subset S^{\varepsilon}\}$ which means that any δ-optimal solution of the SAA problem provides an ε-optimal solution of the true problem. We estimate now the probability of that event.

We can write

$$\left\{\hat{S}_N^{\delta} \not\subset S^{\varepsilon}\right\} = \bigcup_{x \in X \setminus S^{\varepsilon}} \bigcap_{y \in X} \left\{\hat{f}_N(x) \leq \hat{f}_N(y) + \delta\right\}, \tag{3.1}$$

and hence

$$\mathbb{P}\left(\hat{S}_N^{\delta} \not\subset S^{\varepsilon}\right) \leq \sum_{x \in X \setminus S^{\varepsilon}} \mathbb{P}\left(\bigcap_{y \in X} \left\{\hat{f}_N(x) \leq \hat{f}_N(y) + \delta\right\}\right). \tag{3.2}$$

[9] See Section 8.3 of the Appendix for a brief discussion of the LD theory.

[10] By $|X|$ we denote the cardinality of the set X.

Consider a mapping $u: X\backslash S^{\varepsilon} \to X$. It follows from (3.2) that

$$\mathbb{P}\left(\hat{S}_N^{\delta} \not\subset S^{\varepsilon}\right) \leq \sum_{x \in X\backslash S^{\varepsilon}} \mathbb{P}\left\{\hat{f}_N(x) - \hat{f}_N(u(x)) \leq \delta\right\}. \tag{3.3}$$

We assume that the mapping $u(x)$ is chosen in such a way that for some $\varepsilon^* > \varepsilon$,

$$f(u(x)) \leq f(x) - \varepsilon^* \quad \text{for all } x \in X\backslash S^{\varepsilon}. \tag{3.4}$$

Note that if $u(\cdot)$ is a mapping from $X\backslash S^{\varepsilon}$ into the set S, i.e., $u(x) \in S$ for all $x \in X\backslash S^{\varepsilon}$, then (3.4) holds with

$$\varepsilon^* := \min_{x \in X\backslash S^{\varepsilon}} f(x) - v^*, \tag{3.5}$$

and such that ε^* is greater than ε since the set X is finite. Therefore, a mapping $u(\cdot)$ satisfying condition (3.4) always exists. Of course, there are many possible choices of the mapping $u(x)$ with different values of the constant ε^*.

For each $x \in X\backslash S^{\varepsilon}$, define

$$Y(x, \xi) := F(u(x), \xi) - F(x, \xi).$$

Note that $\mathbb{E}[Y(x, \boldsymbol{\xi})] = f(u(x)) - f(x)$, and hence $\mathbb{E}[Y(x, \boldsymbol{\xi})] \leq -\varepsilon^*$ for all $x \in X\backslash S^{\varepsilon}$. The corresponding sample average is

$$\hat{Y}_N(x) := \frac{1}{N} \sum_{i=1}^{N} Y(x, \xi^i) = \hat{f}_N(u(x)) - \hat{f}_N(x).$$

By (3.3) we have

$$\mathbb{P}\left(\hat{S}_N^{\delta} \not\subset S^{\varepsilon}\right) \leq \sum_{x \in X\backslash S^{\varepsilon}} \mathbb{P}\left\{\hat{Y}_N(x) \geq -\delta\right\}. \tag{3.6}$$

Let $I_x(\cdot)$ denote the (large deviations) rate function[11] of the random variable $Y(x, \boldsymbol{\xi})$. The inequality (3.6) together with the LD upper bound (8.6) implies

$$1 - \mathbb{P}\left(\hat{S}_N^{\delta} \subset S^{\varepsilon}\right) \leq \sum_{x \in X\backslash S^{\varepsilon}} e^{-NI_x(-\delta)}. \tag{3.7}$$

[11] See Section 8.3 of the Appendix for the definition of the rate function.

Note that the above inequality (3.7) is not asymptotic and is valid for any random sample of size N.

Assumption (B). For every $x \in X \backslash S^{\varepsilon}$ the moment generating function of the random variable $Y(x, \xi)$ is finite valued in a neighborhood of $t = 0$.

The above assumption (B) holds, for example, if the support Ξ of ξ is a bounded subset of $\mathbb{R}^d$, or if $Y(x, \cdot)$ grows at most linearly and ξ has a distribution from the exponential family.

Theorem 12. *Let ε and δ be nonnegative numbers such that $\delta \leq \varepsilon$. Then*

$$1 - \mathbb{P}\left(\hat{S}_N^{\delta} \subset S^{\varepsilon}\right) \leq |X| e^{-N\eta(\delta, \varepsilon)}, \tag{3.8}$$

where

$$\eta(\delta, \varepsilon) := \min_{x \in X \backslash S^{\varepsilon}} I_x(-\delta). \tag{3.9}$$

Moreover, if assumption (B) holds, then $\eta(\delta, \varepsilon) > 0$.

Proof. The inequality (3.8) is an immediate consequence of the inequality (3.7). We have that $-\delta > -\varepsilon^* \geq \mathbb{E}[Y(x, \xi)]$, and hence it follows by assumption (B) that $I_x(-\delta) > 0$ for every $x \in X \backslash S^{\varepsilon}$. This implies that $\eta(\delta, \varepsilon) > 0$. □

The following asymptotic result is an immediate consequence of inequality (3.8),

$$\limsup_{N \to \infty} \frac{1}{N} \log\left[1 - \mathbb{P}\left(\hat{S}_N^{\delta} \subset S^{\varepsilon}\right)\right] \leq -\eta(\delta, \varepsilon). \tag{3.10}$$

It means that the probability of the event that any δ-optimal solution of the SAA problem provides an ε-optimal solution of the true problem approaches one *exponentially fast* as $N \to \infty$. This suggests that the SAA method can efficiently find an ε-optimal solution of the true problem by solving the SAA problem with accuracy δ, provided that the constant $\eta(\delta, \varepsilon)$ is not "too small".

For δ close to $-\mathbb{E}[Y(x, \xi)]$, we can write by (8.9) that

$$I_x(-\delta) \approx \frac{(-\delta - \mathbb{E}[Y(x, \xi)])^2}{2\sigma_x^2} \geq \frac{(\varepsilon^* - \delta)^2}{2\sigma_x^2}, \tag{3.11}$$

where

$$\sigma_x^2 := \mathbb{V}\mathrm{ar}[Y(x, \xi)] = \mathbb{V}\mathrm{ar}[F(u(x), \xi) - F(x, \xi)].$$

Recall that $I_x(\cdot)$ is a convex function attaining its minimum at $\mathbb{E}[Y(x, \xi)]$, and hence $I_x(z)$ is monotonically increasing for $z \leq \mathbb{E}[Y(x, \xi)]$ as $z \to -\infty$. Therefore, for all $\varepsilon^* \geq 0$ sufficiently small and $\delta \in [0, \varepsilon]$, the constant $\eta(\delta, \varepsilon)$, given in (3.9), can be estimated (see (8.10)) as

$$\eta(\delta, \varepsilon) \geq \frac{(\varepsilon^* - \delta)^2}{3\sigma_{\max}^2} \geq \frac{(\varepsilon - \delta)^2}{3\sigma_{\max}^2}, \tag{3.12}$$

where

$$\sigma_{\max}^2 := \max_{x \in X \setminus S^\varepsilon} \mathbb{V}\mathrm{ar}[F(u(x), \xi) - F(x, \xi)]. \tag{3.13}$$

3.2 *Estimates of the sample size*

Let us fix a significance level $\alpha \in (0, 1)$, and estimate the sample size N which is needed for the probability $\mathbb{P}(\hat{S}_N^\delta \subset S^\varepsilon)$ to be at least $1-\alpha$. By requiring the right-hand side of (3.8) to be less than or equal to α, we obtain that

$$N \geq \frac{1}{\eta(\delta, \varepsilon)} \log\left(\frac{|X|}{\alpha}\right). \tag{3.14}$$

Moreover, we have by (3.12) that $\eta(\delta, \varepsilon) \geq (\varepsilon - \delta)^2/(3\sigma_{\max}^2)$ for all $\varepsilon \geq 0$ sufficiently small and $\delta \in [0, \varepsilon]$. It follows that for all $\varepsilon > 0$ small enough and $0 \leq \delta < \varepsilon$, the sample size which is required for $\mathbb{P}(\hat{S}_N^\delta \subset S^\varepsilon) \geq 1 - \alpha$ to hold can be estimated as

$$N \geq \frac{3\sigma_{\max}^2}{(\varepsilon - \delta)^2} \log\left(\frac{|X|}{\alpha}\right). \tag{3.15}$$

Of course, the constant $\sigma_{\max}^2$, defined in (3.13), depends on the choice of the mapping $u(x)$ and could be difficult to estimate. Moreover, as it was mentioned earlier, the bound (3.15) typically is too conservative for practical estimates of the required sample sizes. However, the estimate (3.15) has interesting consequences for complexity issues. A key characteristic of (3.15) is that N depends only *logarithmically* both on the size of the feasible set X and on the tolerance probability α.

Suppose now that X is a bounded, not necessarily finite, subset of $\mathbb{R}^n$, and that $f(x)$ is finite valued for all $x \in X$. Then we can apply the sample size estimate (3.15) as follows. Recall that any two norms on the finite dimensional

space $\mathbb{R}^n$ are equivalent. For technical reasons it will be convenient to use here the max-norm $||x|| := \max\{|x_1|, \ldots, |x_n|\}$. For a given $\nu > 0$, consider a finite subset X_ν of X such that for any $x \in X$ there is $x' \in X_\nu$ satisfying $||x - x'|| \leq \nu$. Since X is bounded, its diameter $D := \sup_{x, y \in X} ||x - y||$ is finite. Then such set X_ν can be constructed with $|X_\nu| \leq (D/\nu)^n$. By reducing the feasible set X to its subset X_ν, as a consequence of (3.15) we obtain the following estimate of the sample size, required to solve the reduced problem with an accuracy $\varepsilon' > \delta$:

$$N \geq \frac{3\sigma^2_{\max}}{(\varepsilon' - \delta)^2}\left(n \log \frac{D}{\nu} - \log \alpha\right). \tag{3.16}$$

Suppose, further, that the expectation function $f(x)$ is Lipschitz continuous on X modulus L, that is, $|f(x) - f(y)| \leq L\|x - y\|$ for all $x, y \in X$. Then an ε'-optimal solution of the reduced problem is an ε-optimal solution of the true problem with $\varepsilon = \varepsilon' + L\nu$. Let us set $\nu := (\varepsilon - \delta)/(2L)$ and $\varepsilon' := \varepsilon - L\nu = \varepsilon - (\varepsilon - \delta)/2$. By employing (3.16) we obtain the following estimate of the sample size N required to solve the true problem:

$$N \geq \frac{12\sigma^2_{\max}}{(\varepsilon - \delta)^2}\left(n \log \frac{2DL}{\varepsilon - \delta} - \log \alpha\right). \tag{3.17}$$

The above estimate (3.17) shows that the sample size which is required to solve the true problem with probability $1 - \alpha$ and accuracy $\varepsilon > 0$ by solving the SAA problem with accuracy $\delta < \varepsilon$, grows linearly in dimension n of the decision vector x. The estimate (3.17) also involves constants D, L and $\sigma^2_{\max}$ which should be estimated for the considered class of problems. Note that the variance constant $\sigma^2_{\max}$ appears linearly in the estimate (3.17). This should not be surprising since for deciding (comparing), with a given accuracy, which one of just two values $f(u(x))$ and $f(x)$, of the true objective function, is smaller we need a sample of a size proportional to σ^2_x.

3.3 Piecewise linear case

The estimate (3.17) is general and can be applied to any stochastic problem of the form (1.1). In the case of two-stage linear stochastic programming it is possible to say more. In that case the feasible set X and the optimal value $F(x, \xi)$ of the second stage problem satisfy the following assumptions.

(C1) The set X is a convex closed polyhedron.
(C2) For every $\xi \in \Xi$ the function $F(\cdot, \xi)$ is polyhedral.[12]

[12] Recall that a function $g : \mathbb{R}^n \to \overline{\mathbb{R}}$ is said to be polyhedral if it is proper convex and lower semicontinuous, its domain is a convex closed polyhedron and $g(x)$ is piecewise linear on its domain.

We also assume in this section that the support of ξ is finite.

(C3) The support Ξ of ξ is finite, $\Xi = \{\xi_1, \ldots, \xi_K\}$, with respective (positive) probabilities p_k, $k = 1, \ldots, K$.

We can absorb the feasible set X into the objective function by formulating the true and SAA problems in the forms (2.3) and (2.4), respectively. Denote $\overline{F}_k(\cdot) := \overline{F}(\cdot, \xi_k)$. We have then that

$$\overline{f}(\cdot) = \sum_{k=1}^{K} p_k \overline{F}_k(\cdot),$$

and $\operatorname{dom} \overline{f} = \cap_{k=1}^{K} \operatorname{dom} \overline{F}_k$. Note also that $\operatorname{dom} \overline{F}_k = X \cap \operatorname{dom} F(\cdot, \xi_k)$. It follows from the above assumptions (C1)–(C3) that the function $\overline{f}$ is polyhedral provided that its domain is nonempty.

Suppose, further, for the sake of simplicity that the true problem has unique optimal solution, i.e., $S = \{x^*\}$. Consider the following event.

($\mathcal{E}_N$) The SAA problem (1.6) has unique optimal solution $\hat{x}_N$ and $\hat{x}_N = x^*$.

The above event can also be formulated in the form: "the point x^* is the unique optimal solution of the SAA problem".

Theorem 13. *Suppose that the assumptions* (C1)–(C3) *hold and the true problem has unique optimal solution* x^*. *Then the event* ($\mathcal{E}_N$) *happens w.p.1 for* N *large enough and, moreover, there exists a constant* $\eta^* > 0$ *such that*

$$\lim_{N \to \infty} \frac{1}{N} \log[1 - \mathbb{P}(\mathcal{E}_N)] = -\eta^*. \tag{3.18}$$

Proof of the above theorem is based on the following lemma which is implied by the polyhedral structure of the considered problem. Denote by

$$\overline{X} := \operatorname{dom} \overline{f} = X \cap \operatorname{dom} f$$

the domain of the function $\overline{f}$, by $T_{\overline{X}}(x^*)$ the tangent cone to $\overline{X}$ at x^*, and by $\overline{f}'(x^*, d)$ the directional derivative of $\overline{f}$ at the point x^* in the direction $d \in \mathbb{R}^n$. Note that since $\overline{f}(x^*)$ is finite and the function $\overline{f}$ is polyhedral, the directional derivative $\overline{f}'(x^*, d)$ is finite if $d \in T_{\overline{X}}(x^*)$, and $\overline{f}'(x^*, d) = +\infty$ otherwise.

Lemma 14. *Suppose that the assumptions* (C1)–(C3) *hold and the true problem has unique optimal solution* x^*. *Then there exists a finite number of directions* $d_1, \ldots, d_J \in \mathbb{R}^n \setminus \{0\}$ *such that* $\overline{f}'(x^*, d_j) > 0$, $j = 1, \ldots, J$, *and the event* ($\mathcal{E}_N$) *happens if and only if* $\tilde{f}'_N(x^*, d_j) > 0$, $j = 1, \ldots, J$.

Proof. For the sake of notational convenience we can assume without loss of generality that $x^* = 0$. We have that for every $k \in \{1, \ldots, K\}$, the function $\overline{F}_k(\cdot)$ is polyhedral and $\overline{F}_k(x^*)$ is finite. Because of the polyhedral structure of $\overline{F}_k$ we have that the tangent cone to the domain of $\overline{F}_k$ at the point $x^* = 0$ can be represented as a union of a finite number of closed convex polyhedral cones such that $\overline{F}_k$ is linear[13] on each cone in a neighborhood of $x^* = 0$. By taking all possible intersections of these cones for $k = 1, \ldots, K$, we can construct a finite number of closed convex polyhedral cones $T_1, \ldots, T_M$ with the following properties. (i) The union of the tangent cones $T_{\mathrm{dom}\, F_k}(x^*)_{k \in \{1, \ldots, K\}}$ coincides with $\cup_{m=1}^M T_m$. (ii) To every $k \in \{1, \ldots, K\}$ corresponds a subset M_k of $\{1, \ldots, M\}$ such that the tangent cone $T_{\mathrm{dom}\, F_k}(x^*)$ coincides with $\cup_{m \in M_k} T_m$ and $\overline{F}_k$ is linear on each cone T_m, $m \in M_k$, in a neighborhood of $x^* = 0$. Now let $d_1, \ldots, d_J$ be the set of vectors of length one generating all extreme rays of all cones T_m, $m = 1, \ldots, M$.

For any generated sample ξ^i, $i = 1, \ldots, N$, we have that the corresponding function $\tilde{f}_N$ is polyhedral and $\mathrm{dom}\, \tilde{f}_N = \cap_{i=1}^N \mathrm{dom}\, \overline{F}(\cdot, \xi^i)$. Moreover, by the above construction we have that there exists a set $\mathcal{M} \subset \{1, \ldots, M\}$ such that the tangent cone to $\mathrm{dom}\, \tilde{f}_N$ at x^* coincides with $\cup_{m \in \mathcal{M}} T_m$ and $\tilde{f}_N(\cdot)$ is linear on each cone T_m, $m \in \mathcal{M}$, in a neighborhood of $x^* = 0$. Let $\mathcal{J}$ be a subset of $\{1, \ldots, J\}$ such that the vectors d_j, $j \in \mathcal{J}$, generate the set of all extreme rays of the cones T_m, $m \in \mathcal{M}$. We have then that if $\tilde{f}'_N(x^*, d_j) > 0$ for all $j \in \mathcal{J}$, then $x^* = 0$ is the minimizer of $\tilde{f}_N$ over every cone T_m, $m \in \mathcal{M}$. By convexity this implies that x^* is the unique minimizer of $\tilde{f}_N$ over $\mathbb{R}^n$. Conversely, if $\tilde{f}'_N(x^*, d_j)$ is less than or equal to zero for some $j \in \{1, \ldots, J\}$, then either x^* is not a minimizer or x^* is not a unique minimizer of $\tilde{f}_N$ over $\mathbb{R}^n$. The last assertion follows since $\tilde{f}_N$ is polyhedral.

Finally, the expected value function $\overline{f}$ has the same type of structure with the corresponding extreme rays forming a subset of the set $\{d_1, \ldots, d_J\}$. We have then that $\overline{f}'(x^*, d_j)$ is positive and finite for $d_j \in T_{\overline{X}}(x^*)$, and $\overline{f}'(x^*, d_j) = +\infty$ otherwise. This completes the proof of the lemma. □

Proof (of Theorem 13). Since $\overline{f}(x^*)$ is finite we have that for any $d \in \mathbb{R}^n$,

$$\overline{f}'(x^*, d) = \sum_{k=1}^K p_k \overline{F}'_k(x^*, d) = \mathbb{E}\left[\overline{F}'_\xi(x^*, d)\right], \tag{3.19}$$

where $\overline{F}_\xi(\cdot) := \overline{F}(\cdot, \xi)$ and $\overline{F}'_\xi(x^*, d)$ denotes the directional derivative of $\overline{F}_\xi$ at x^* in the direction d. We also have that

$$\tilde{f}'_N(x^*, d) = \frac{1}{N} \sum_{i=1}^N \overline{F}'_{\xi^i}(x^*, d). \tag{3.20}$$

[13] We say here that a real valued function $g(x)$ is linear if $g(x) = a^T x + b$ for some $a, b \in \mathbb{R}^n$. It is more accurate to call such a function affine.

For a given $d \in \mathbb{R}^n$, by applying the LLN to the random variable $\overline{F}'_\xi(x^*, d)$ we obtain that

$$\tilde{f}'_N(x^*, d) \to \overline{f}'(x^*, d) \quad \text{w.p.1 as } N \to \infty. \tag{3.21}$$

Note that (3.21) holds if $\overline{f}'(x^*, d) = +\infty$ as well. Now let $d_1, \ldots, d_J$ be vectors specified in Lemma 14. Since $\overline{f}'(x^*, d_j) > 0$, $j = 1, \ldots, J$, it follows by (3.21) that $\tilde{f}'_N(x^*, d_j) > 0$, $j = 1, \ldots, J$, and hence the event $(\mathcal{E}_N)$ happens, w.p.1 for N large enough.

Now by the upper LD bound (see (8.8) and (8.13)) we have that for any $d \in \mathbb{R}^n$ the inequality

$$\mathbb{P}\left\{\tilde{f}'_N(x^*, d) \le 0\right\} \le e^{-N\eta_d}, \tag{3.22}$$

holds with

$$\eta_d := \log(1 - p_d)^{-1} + \varrho_d I_d(0).$$

Here p_d is the probability that $\overline{F}'_\xi(x^*, d) = +\infty$, $I_d(\cdot)$ is the rate function of $\overline{F}'_\xi(x^*, d)$ conditional on the event $\{\overline{F}'_\xi(x^*, d) < +\infty\}$, and $\varrho_d := 1$ if the mean of $\overline{F}'_\xi(x^*, d)$ conditional on $\{\overline{F}'_\xi(x^*, d) < +\infty\}$ is positive, and $\varrho_d := 0$ otherwise. If $d \notin T_{\overline{X}}(x^*)$, then p_d is positive, and hence $\eta_d > 0$. If $d \in T_{\overline{X}}(x^*)$, then $p_d = 0$ and $\overline{f}'(x^*, d)$ is finite. Since ξ has a finite support we have that in this case $I_d(0) > 0$ if $\overline{f}'(x^*, d) > 0$. Note also that if $p_d = 0$ and $\overline{f}'(x^*, d) > 0$, then $\eta_d = I_d(0)$ with $I_d(\cdot)$ being the rate function of $\overline{F}'_\xi(x^*, d)$.

Now by Lemma 14 we have that the complement of the event $(\mathcal{E}_N)$ coincides with the union of the events $\{\tilde{f}'_N(x^*, d_j) \le 0\}_{j \in \{1, \ldots, J\}}$. Therefore, it follows from (3.22) that

$$1 - \mathbb{P}(\mathcal{E}_N) \le \sum_{j=1}^{J} e^{-N\eta_{d_j}}, \tag{3.23}$$

which in turn implies

$$\limsup_{N \to \infty} \frac{1}{N} \log[1 - \mathbb{P}(\mathcal{E}_N)] \le -\min_{1 \le j \le J} \eta_{d_j}. \tag{3.24}$$

We also have that

$$1 - \mathbb{P}(\mathcal{E}_N) \ge \mathbb{P}\left\{\tilde{f}'_N(x^*, d_j) \le 0\right\}, \quad j = 1, \ldots, J. \tag{3.25}$$

Since ξ has a finite support, the lower bound of Cramér's LD Theorem can be applied to the right hand side of (3.25) to obtain

$$\liminf_{N\to\infty} \frac{1}{N} \log[1 - \mathbb{P}(\mathcal{E}_N)] \geq -\eta_{d_j}, \quad j = 1, \ldots, J. \tag{3.26}$$

The inequalities (3.25) and (3.26) imply (3.18) with

$$\eta^* := \min_{1\leq j\leq J} \eta_{d_j}. \tag{3.27}$$

By the above discussion we have that every η_{d_j} is positive and hence $\eta^* > 0$. This completes the proof. □

Theorem 13 shows that, under the assumed polyhedral structure, by solving the SAA problem one obtains the exact optimal solution of the true problem for sufficiently large sample size, and moreover, the probability of that event approaches one exponentially fast with increase of the sample size. Note that even so, the corresponding optimal value $\hat{v}_N$ is not exact and may have a relatively large variability. Of course, although (3.18) ensures an exponential convergence of the probability $\mathbb{P}(\mathcal{E}_N)$ to one, the convergence can be very slow if the constant η^* is small. Consider, for example, a situation where for some direction $d \in \mathbb{R}^n$ the directional derivative $\overline{F}'_k(x^*, d) = +\infty$ for just one $k \in \{1, \ldots, K\}$ with the corresponding probability p_k being very small, while $\overline{F}'_\ell(x^*, d) < 0$ for all $\ell \neq k$. Then $\tilde{f}'(x^*, d) = +\infty$, while $\tilde{f}'_N(x^*, d)$ is negative unless the corresponding sample includes the point ξ_k in which case $\tilde{f}'_N(x^*, d) = +\infty$, and that happens with probability $1-(1-p_k)^N$. Therefore, in such situation one will need a very large sample to ensure that $\tilde{f}'_N(x^*, d) = +\infty$ with a reasonably large probability.

Suppose now that the assumptions (C1)–(C3) hold, but the set S of optimal solutions of the true problem is not necessarily a singleton. Because of the assumed polyhedral structure we have here that the set S is a convex closed polyhedron. The set $\hat{S}_N$ of optimal solutions of the SAA problem is also a convex closed polyhedron. Consider the following event.

($\mathcal{G}_N$) The set $\hat{S}_N$ is nonempty and forms a face of the set S.

Proof of the following result is similar to the proof of Theorem 13.

Theorem 15. *Suppose that the assumptions* (C1)–(C3) *hold and the set S is nonempty and bounded. Then the event* ($\mathcal{G}_N$) *happens w.p.1 for N large enough and, moreover,*

$$\limsup_{N\to\infty} \frac{1}{N} \log[1 - \mathbb{P}(\mathcal{G}_N)] < 0. \tag{3.28}$$

The above analysis is based on the polyhedral structure of the functions $\overline{F}(\cdot, \xi)$, $\xi \in \Xi$, which is inherited by the expected value function $\overline{f}(\cdot)$ because Ξ is finite. Since $\overline{f}(\cdot)$ is *polyhedral*, it follows that the set $S = \arg\min \overline{f}$ is nonempty, provided that $\overline{f}(\cdot)$ is bounded from below, and moreover,

$$\overline{f}(x) \geq v^* + c \operatorname{dist}(x, S) \tag{3.29}$$

for all $x \in \mathbb{R}^n$ and some constant $c > 0$. In particular, if $S = \{x^*\}$ is a singleton, then (3.29) takes the form

$$\overline{f}(x) \geq \overline{f}(x^*) + c\|x - x^*\|. \tag{3.30}$$

Optimal solution x^* is said to be *sharp* if the above condition (3.30) holds. Since $\overline{f}(\cdot)$ is polyhedral, x^* is always sharp provided it is unique.

If $\boldsymbol{\xi}$ has a continuous distribution, and hence its support Ξ is not a finite set, then condition (3.30) still may hold in some particular situations. It is possible to show that in such cases, under some mild additional conditions, the event $(\mathcal{E}_N)$ happens w.p.1 for N large enough and the probability $\mathbb{P}(\mathcal{E}_N)$ approaches one at an exponential rate. By convexity of $\overline{f}(\cdot)$, condition (3.30) is equivalent to the condition that $\overline{f}'(x^*, d) \geq c\|d\|$ for all $d \in \mathbb{R}^n$. If $\boldsymbol{\xi}$ has a continuous distribution, then it may happen that $\overline{f}'(x^*, d) = 0$ for some $d \neq 0$ even if the optimal solution x^* is unique. In such cases an SAA optimal solution $\hat{x}_N$ converges to x^* typically at a rate of $O_p(N^{-1/2})$.

3.4 Conditioning of piecewise linear stochastic programs

It was shown in Section 3.3 that, under the assumed polyhedral structure of the problem, one can find an exact optimal solution of the true problem by solving the SAA problem with sufficiently large sample size. Of course, the sample size N which is required for that to happen, with a reasonably large probability, is problem dependent. It turns out that in some cases the true problem possesses unique optimal solution x^* which can be computed exactly by solving the corresponding SAA problem based on a relatively small sample. This usually is indicated by the fact that while solving several SAA problems, based on independent samples of $\boldsymbol{\xi}$ of a relatively small size, one obtains the same optimal solution. It is natural to call such problems well conditioned. On the other hand, it may happen that either the set S is not a singleton or the sample size which is needed for the event $(\mathcal{E}_N)$ to happen, with a reasonably large probability, is very large. We call such problems *ill conditioned*.

Suppose, for the sake of simplicity, that in addition to the assumptions (C1)–(C3) the following holds.

(C4) The expected value function $f(x)$ is finite valued for all $x \in X$ in a neighborhood of the point x^*.

In the case of two-stage stochastic programming the above assumption holds, for example, if the recourse is relatively complete. We have by Lemma 14 that, under the assumptions (C1)–(C4), there exists a finite number of nonzero vectors $d_1, \ldots, d_J \in T_X(x^*)$ such that the event $(\mathcal{E}_N)$ happens iff $\tilde{f}'_N(x^*, d_j) > 0$, $j = 1, \ldots, J$. Furthermore, the (3.18) holds with the exponential constant η^* given by

$$\eta^* = \min_{1 \le j \le J} I_{d_j}(0), \tag{3.31}$$

where $I_d(\cdot)$ denotes the rate function of the random variable $F'_\xi(x^*, d)$.

We have that $\mathbb{E}[F'_\xi(x^*, d)] = f'(x^*, d)$. Moreover, because of the assumption (C4) and the assumed polyhedral structure, we have here that if $d \in T_X(x^*)$, then $f'(x^*, d)$ is finite, and hence $F'_\xi(x^*, d)$ is finite for any $\xi \in \Xi = \{\xi_1, \ldots, \xi_K\}$. Now, by (8.9), for $f'(x^*, d)$ close to zero we can approximate $I_d(0)$ as follows

$$I_d(0) \approx \frac{[f'(x^*, d)]^2}{2\mathbb{V}\mathrm{ar}[F'_\xi(x^*, d)]}. \tag{3.32}$$

This leads to the following concept. We call

$$\kappa := \max_{1 \le j \le J} \frac{\mathbb{V}\mathrm{ar}[F'_\xi(x^*, d_j)]}{[f'(x^*, d_j)]^2} \tag{3.33}$$

the *condition number* of the true problem. By (3.21) and (3.22) we have that $\eta^* \approx (2\kappa)^{-1}$. That is, for large κ the true problem is viewed as ill conditioned since one needs a very large sample in order to solve it exactly. Note that $f'(x^*, d) = 0$ for some $d \in T_X(x^*) \setminus \{0\}$ iff the true problem has another optimal solution apart from x^*, i.e., the set S is not a singleton. In that case $\kappa = +\infty$.

Of course, it is impractical to try to compute the above condition number. For one thing it depends on the optimal solution x^* which is not known a priori. Therefore, the above analysis has rather a conceptual value. Note also that we approach here the question of conditioning of the true problem from the point of view of SAAs. That is, problems which are "flat" in some feasible directions near the optimal solution x^* are difficult to solve exactly. This is measured by the squared coefficient of variation of the corresponding directional derivative.

As it was mentioned in the previous section, in case ξ has a continuous distribution it can happen that $f'(x^*, d) = 0$ for some $d \in T_X(x^*) \setminus \{0\}$ even if the optimal solution x^* is unique. Typically in such cases $\hat{x}_N$ converges to x^* at a rate of $O_p(N^{-1/2})$, i.e., accuracy of the estimator $\hat{x}_N$ improves slowly with increase of the sample size N. From the above point of view any such problem is ill conditioned.

Let us also remark that it is possible to introduce a similar concept of condition number for problems with finite feasible set X by replacing the directional derivatives with the corresponding finite differences (see the discussion of Section 3.1).

4 Validation analysis

Suppose that we are given a feasible point $\overline{x} \in X$ as a candidate for an optimal solution of the true problem. For example, $\overline{x}$ can be an output of a run of the corresponding SAA problem. In this section we discuss ways to evaluate quality of this candidate solution. This is important, in particular, for a choice of the sample size and stopping criteria in simulation based optimization. There are basically two approaches to such validation analysis. We can either try to estimate the optimality gap $f(\overline{x}) - v^*$, or to evaluate first order (KKT) optimality conditions at the considered point $\overline{x}$.

Let us emphasize that the following analysis is designed for the situations where the value $f(\overline{x})$, of the true objective function at the considered point, is *finite*. In the case of two-stage programming this requires, in particular, that the second stage problem, associated with first stage decision variables $\overline{x}$, is feasible for almost every realization of the random data.

4.1 Estimation of the optimality gap

In this section we consider the problem of estimating the optimality gap

$$\mathrm{gap}(\overline{x}) := f(\overline{x}) - v^* \tag{4.1}$$

associated with the candidate solution $\overline{x}$. Clearly, for any feasible $\overline{x} \in X$, $\mathrm{gap}(\overline{x})$ is nonnegative and $\mathrm{gap}(\overline{x}) = 0$ iff $\overline{x}$ is an optimal solution of the true problem.

Consider the optimal value $\hat{v}_N$ of the SAA problem (1.6). As it was discussed in Section 2.2 we have that $v^* \geq \mathbb{E}[\hat{v}_N]$. This means that $\hat{v}_N$ provides a valid *statistical lower bound* for the optimal value v^* of the true problem. The expectation $\mathbb{E}[\hat{v}_N]$ can be estimated by averaging. That is, one can solve M times SAA problems based on independently generated samples each of size N. Let $\hat{v}_N^1, \ldots, \hat{v}_N^M$ be the computed optimal values of these SAA problems. Then

$$\overline{v}_{N,M} := \frac{1}{M} \sum_{j=1}^{M} \hat{v}_N^j \tag{4.2}$$

is an unbiased estimator of $\mathbb{E}[\hat{v}_N]$. Since the samples, and hence $\hat{v}_N^1, \ldots, \hat{v}_N^M$, are independent, we can estimate the variance of $\overline{v}_{N,M}$ by

$$\hat{\sigma}^2_{N,\,M} := \frac{1}{M}\left[\frac{1}{M-1}\sum_{j=1}^{M}\left(\hat{v}_N^j - \overline{v}_{N,\,M}\right)^2\right]. \tag{4.3}$$

Note that the above make sense only if the optimal value v^* of the true problem is finite. Note also that the inequality $v^* \geq \mathbb{E}[\hat{v}_N]$ holds and $\hat{v}_N$ gives a valid statistical lower bound even if $f(x) = +\infty$ for some $x \in X$.

In general, the random variable $\hat{v}_N$, and hence its replications $\hat{v}_N^j$, does not have a normal distribution, even approximately (see Theorem 10 and the following up discussion). However, by the CLT, the probability distribution of the average $\overline{v}_{N,M}$ becomes approximately normal as M increases. Therefore, we can use

$$L_{N,M} := \overline{v}_{N,M} - t_{\alpha,M-1}\hat{\sigma}_{N,M} \tag{4.4}$$

as an approximate $100(1-\alpha)\%$ confidence[14] lower bound for the expectation $\mathbb{E}[\hat{v}_N]$.

We can also estimate $f(\overline{x})$ by sampling.[15] That is, let $\hat{f}_{N'}(\overline{x})$ be the sample average estimate of $f(\overline{x})$, based on a sample of size N' generated independently of samples involved in computing $\overline{x}$. Let $\hat{\sigma}^2_{N'}(\overline{x})$ be an estimate of the variance of $\hat{f}_{N'}(\overline{x})$. In the case of iid sample, one can use the sample variance estimate

$$\hat{\sigma}^2_{N'}(\overline{x}) := \frac{1}{N'(N'-1)}\sum_{i=1}^{N'}\left[F(\overline{x}, \xi^i) - \hat{f}_{N'}(\overline{x})\right]^2. \tag{4.5}$$

Then

$$U_{N'}(\overline{x}) := \hat{f}_{N'}(\overline{x}) + z_\alpha\hat{\sigma}_{N'}(\overline{x}) \tag{4.6}$$

gives an approximate $100(1-\alpha)\%$ confidence upper bound for $f(\overline{x})$. Note that since N' typically is large, we use here the standard normal critical value z_α.

We have that

$$\mathbb{E}\left[\hat{f}_{N'}(\overline{x}) - \overline{v}_{N,M}\right] = f(\overline{x}) - \mathbb{E}[\hat{v}_N] = \mathrm{gap}(\overline{x}) + v^* - \mathbb{E}[\hat{v}_N] \geq \mathrm{gap}(\overline{x}),$$

[14] Here $t_{\alpha,\nu}$ is the α-critical value of t-distribution with ν degrees of freedom. This critical value is slightly bigger than the corresponding standard normal critical value z_α, and $t_{\alpha,\nu}$ quickly approaches z_α as ν increases.

[15] Compare with the discussion of Section 2.2, see (2.20) in particular.

i.e., $\hat{f}_{N'}(\overline{x}) - \overline{v}_{N,M}$ is a biased estimator of the gap($\overline{x}$). Also the variance of this estimator is equal to the sum of the variances of $\hat{f}_{N'}(\overline{x})$ and $\overline{v}_{N,M}$, and hence

$$\hat{f}_{N'}(\overline{x}) - \overline{v}_{N,M} + z_\alpha \sqrt{\hat{\sigma}^2_{N'}(\overline{x}) + \hat{\sigma}^2_{N,M}} \tag{4.7}$$

provides a conservative $100(1-\alpha)\%$ confidence upper bound for the gap($\overline{x}$). We say that this upper bound is "conservative" since in fact it gives a $100(1-\alpha)\%$ confidence upper bound for the gap($\overline{x}$) $+ v^* - \mathbb{E}[\hat{v}_N]$, and we have that $v^* - \mathbb{E}[\hat{v}_N] \geq 0$.

In order to calculate the estimate $\hat{f}_{N'}(\overline{x})$ one needs to compute the value $F(\overline{x}, \xi^i)$ of the objective function for every generated sample realization ξ^i, $i = 1, \ldots, N'$. Typically it is much easier to compute $F(\overline{x}, \xi)$, for a given $\xi \in \Xi$, than to solve the corresponding SAA problem. Therefore, often one can use a relatively large sample size N', and hence to estimate $f(\overline{x})$ quite accurately. Evaluation of the optimal value v^* by employing the estimator $\overline{v}_{N,M}$ is a more delicate problem.

There are two types of errors in using $\overline{v}_{N,M}$ as an estimator of v^*, namely the bias $v^* - \mathbb{E}[\hat{v}_N]$ and variability of $\overline{v}_{N,M}$ measured by its variance. Both errors can be reduced by increasing N, and the variance can be reduced by increasing N and M. Note, however, that the computational effort in computing $\overline{v}_{N,M}$ is proportional to M, since the corresponding SAA problems should be solved M times, and to the computational time for solving a single SAA problem based on a sample of size N. Naturally one may ask what is the best way of distributing computational resources between increasing the sample size N and the number of repetitions M. This question is, of course, problem dependent. Let p_N be the probability that an optimal solution of an SAA problem, based on a sample of size N, is an optimal solution of the true problem. Then the probability that in M repetitions of the SAA problems at least one optimal solution is an optimal solution of the true problem is $1-(1-p_N)^M$. Now if we use the asymptotic $p_N \approx 1-e^{-N\eta}$ for some positive constant η, then

$$1 - (1 - p_N)^M \approx 1 - e^{-NM\eta},$$

which is the same as the asymptotic $p_{NM} \approx 1-e^{-NM\eta}$. Of course, the above are asymptotics and for finite sample sizes the situation can be different. In any case it appears that in cases where complexity of SAA problems grows fast with increase of the sample size, it is more advantageous to use a larger number of repetitions M. On the other hand, in some cases the computational effort in solving an SAA problem grows only linearly with the sample size N. In such cases one can use a larger N and make only a few repetitions M in order to estimate the variance of $\overline{v}_{N,M}$.

The bias $v^* - \mathbb{E}[\hat{v}_N]$ does not depend on M, of course. The following result shows that it decreases monotonically with increase of the sample size N.

Proposition 16. *Suppose that the sample is* iid. *Then* $\mathbb{E}[\hat{v}_N] \leq \mathbb{E}[\hat{v}_{N+1}]$ *for any* $N \in \mathbb{N}$.

Proof. We can write

$$\hat{f}_{N+1}(x) = \frac{1}{N+1} \sum_{i=1}^{N+1} \left[\frac{1}{N} \sum_{j \neq i} F(x, \xi^j) \right].$$

Moreover,

$$\inf_{x \in X} \mathbb{E}\left[\frac{1}{N} \sum_{j \neq i} F(x, \xi^j) \right] \geq \mathbb{E}\left[\inf_{x \in X} \frac{1}{N} \sum_{j \neq i} F(x, \xi^j) \right], \tag{4.8}$$

and since the sample is iid the right hand side of (4.8) is equal to $\mathbb{E}[\hat{v}_N]$. It follows that

$$\mathbb{E}[\hat{v}_{N+1}] = \inf_{x \in X} \mathbb{E}\left[\hat{f}_{N+1}(x) \right] \geq \frac{1}{N+1} \sum_{i=1}^{N+1} \mathbb{E}\left[\inf_{x \in X} \frac{1}{N} \sum_{j \neq i} F(x, \xi^j) \right] = \mathbb{E}[\hat{v}_N],$$

which completes the proof. □

As it was discussed in Section 2.2 (see (2.29) in particular) we have that if the set S is not a singleton, then the bias $v^* - \mathbb{E}[\hat{v}_N]$ typically converges to zero, as N increases, at a rate of $O(N^{-1/2})$, and tends to be bigger for a larger set S. Ill conditioned problems with a large set of optimal or nearly optimal solutions show a similar behavior of the bias. On the other hand, in well conditioned problems, the event $\hat{x}_N = x^*$, and hence $\hat{v}_N = \hat{f}_N(x^*)$, happens with probability approaching one exponentially fast. Since $\mathbb{E}[\hat{f}_N(x^*)] = f(x^*)$, in such cases the bias $v^* - \mathbb{E}[\hat{v}_N] = f(x^*) - \mathbb{E}[\hat{v}_N]$ tends to be much smaller.

In the above approach the upper and lower statistical bounds were computed independently of each other. Alternatively, it is possible to use the same sample for estimating $f(\bar{x})$ and $\mathbb{E}[\hat{v}_N]$. That is, for M generated samples each of size N, the gap is estimated by

$$\widehat{\mathrm{gap}}_{N,M}(\bar{x}) := \frac{1}{M} \sum_{j=1}^{M} \left[\hat{f}_N^j(\bar{x}) - \hat{v}_N^j \right], \tag{4.9}$$

where $\hat{f}_N^j(\bar{x})$ and $\hat{v}_N^j$ are computed from the *same* sample $j=1,\ldots,M$. We have that the expected value of $\widehat{\mathrm{gap}}_{N,M}(\bar{x})$ is $f(\bar{x})-\mathbb{E}[\hat{v}_N]$, i.e., the estimator $\widehat{\mathrm{gap}}_{N,M}(\bar{x})$ has the same bias as $f_{N'}(\bar{x})-\bar{v}_{N,M}$. On the other hand, for well conditioned problems it happens with high probability that $\hat{v}_N^j = \hat{f}_N^j(x^*)$, and as a consequence $\hat{f}_N^j(\bar{x})$ tends to be highly positively correlated with $\hat{v}_N^j$, provided that $\bar{x}$ is close to x^*. In such cases variability of $\widehat{\mathrm{gap}}_{N,M}(\bar{x})$ can be considerably smaller than variability of $\hat{f}_{N'}(\bar{x})-\bar{v}_{N,M}$. This is the idea of common random number generated estimators which was discussed at the end of Section 1 (see (1.7) and (1.8), in particular).

In order to reduce the bias of the above estimators of the optimality gap let us consider the following approach. Let $\lambda: \Xi \to \mathbb{R}^n$ be a measurable mapping such that $\mathbb{E}[\lambda(\xi)]=0$. We have then that $\mathbb{E}[\bar{F}(x,\xi)+\lambda(\xi)^{\mathrm{T}}x] = \mathbb{E}[\bar{F}(x,\xi)]$ for any $x\in\mathbb{R}^n$, where $\bar{F}(\cdot,\cdot)$ is defined in (2.2). Consequently, v^* is equal to the optimal value of the problem

$$\operatorname*{Min}_{x\in\mathbb{R}^n} \mathbb{E}\big[\,\bar{F}(x,\xi)+\lambda(\xi)^T x\big]. \tag{4.10}$$

It follows that $v^* \geq \mathbb{E}[\tilde{v}_N^{\lambda}]$, where $\tilde{v}_N^{\lambda}$ is the optimal value of the SAA problem corresponding to the problem (4.10). That is, $\tilde{v}_N^{\lambda}$ provides a valid statistical lower bound for the true optimal value v^*. Of course, quality of this lower bound depends on a choice of the mapping $\lambda(\cdot)$. This is closely related to a variance reduction technique called "Linear Control Random Variables" method. We discuss that later (see Section 5.2).

From a theoretical point of view there is an optimal choice of the mapping $\lambda(\cdot)$. As it was discussed in Chapter "Optimality and Duality in Stochastic Programming", with the problem (4.10) is associated a dual problem such that if the function $\bar{F}(\cdot,\xi)$ is convex w.p.1 and $\bar{\lambda}(\cdot)$ is an optimal solution of that dual problem, then under mild regularity conditions,

$$v^* = \mathbb{E}\left[\inf_{x\in\mathbb{R}^n}\left\{\bar{F}(x,\xi)+\bar{\lambda}(\xi)^T x\right\}\right], \tag{4.11}$$

and for a.e. ξ the set of minimizers of $\bar{F}(x,\xi)+\bar{\lambda}(\xi)^T x$ contains S. In particular, if the set $\Xi=\{\xi_1,\ldots,\xi_K\}$ is finite, then this dual problem can be written in the form

$$\operatorname*{Max}_{\lambda_1,\ldots,\lambda_K} \; -\sum_{k=1}^{K} p_k \bar{F}_k^*(-\lambda_k) \quad \text{subject to} \quad \sum_{k=1}^{K} p_k\lambda_k = 0, \tag{4.12}$$

where $\bar{F}_k^*(\cdot)$ is the conjugate of the function $\bar{F}_k(\cdot) := \bar{F}(\cdot,\xi_k)$. Of course, if $\bar{\lambda}(\cdot)$ is known, i.e., the dual problem is solved, then the expected value problem (1.1) can be easily solved and one does not need sampling.

4.2 KKT statistical test

Suppose that the feasible set X is defined by constraints in the form

$$X := \{x \in \mathbb{R}^n : g_j(x) = 0, j = 1, \ldots, q;\ g_j(x) \leq 0, j = q+1, \ldots, p\}, \tag{4.13}$$

where $g_j(x)$ are smooth (at least continuously differentiable) deterministic functions. Let $x^* \in X$ be an optimal solution of the true problem and suppose that the expected value function $f(x^*)$ is differentiable at x^*. Then, under a constraint qualification, first order (KKT) optimality conditions hold at x^*. That is, there exist Lagrange multipliers λ_j such that $\lambda_j \geq 0$, $j \in \mathcal{I}(x^*)$ and

$$\nabla f(x^*) + \sum_{j \in \mathcal{J}(x^*)} \lambda_j \nabla g_j(x^*) = 0, \tag{4.14}$$

where $\mathcal{I}(x) := \{j : g_j(x) = 0, j = q+1, \ldots, p\}$ denotes the index set of inequality constraints active at a point $x \in \mathbb{R}^n$, and $\mathcal{J}(x) := \{1, \ldots, q\} \cup \mathcal{I}(x)$. Note that if the constraint functions are linear, say $g_j(x) := a_j^T x + b_j$, then $\nabla g_j(x) = a_j$ and the above KKT conditions hold without a constraint qualification. Consider the (polyhedral) cone

$$K(x) := \left\{ z \in \mathbb{R}^n : z = \sum_{j \in \mathcal{J}(x)} \alpha_j \nabla g_j(x),\ \alpha_j \leq 0,\ j \in \mathcal{I}(x) \right\}. \tag{4.15}$$

Then the KKT optimality conditions can be written in the form $\nabla f(x^*) \in K(x^*)$.

Suppose now that $f(\cdot)$ is differentiable at the candidate point $\overline{x} \in X$, and that the gradient $\nabla f(\overline{x})$ can be estimated by a (random) vector $\gamma_N(\overline{x})$. In particular, if $F(\cdot, \boldsymbol{\xi})$ is differentiable at $\overline{x}$ w.p.1, then we can use the estimator

$$\gamma_N(\overline{x}) := \frac{1}{N} \sum_{i=1}^{N} \nabla_x F(\overline{x}, \xi^i) = \nabla \hat{f}_N(\overline{x}) \tag{4.16}$$

associated with the generated[16] random sample. Note that if, moreover, the derivatives can be taken inside the expectation, that is,

$$\nabla f(\overline{x}) = \mathbb{E}[\nabla_x F(\overline{x}, \boldsymbol{\xi})], \tag{4.17}$$

[16] We emphasize that the random sample in (4.16) is generated independently of the sample used to compute the candidate point $\overline{x}$.

then the above estimator is unbiased, i.e., $\mathbb{E}[\gamma_N(\overline{x})] = \nabla f(\overline{x})$. Regularity conditions which are required for the interchangeability formula (4.17) to hold are discussed in section "Expectation Functions" of the chapter "Optimality and Duality in Stochastic Programming". In the case of two-stage linear stochastic programming with recourse, formula (4.17) typically holds if the corresponding random data have a continuous distribution. On the other hand, if the random data have a discrete distribution with a finite support, then the expected value function $f(x)$ is piecewise linear and typically is nondifferentiable at an optimal solution.

Suppose, further, that $V_N := N^{1/2}[\gamma_N(\overline{x}) - \nabla f(\overline{x})]$ converges in distribution, as N tends to infinity, to multivariate normal with zero mean vector and covariance matrix Σ, written $V_N \Rightarrow N(0, \Sigma)$. For the estimator $\gamma_N(\overline{x})$ defined in (4.16), this holds by the CLT if the interchangeability formula (4.17) holds, the sample is iid, and $\nabla_x F(\overline{x}, \xi)$ has finite second order moments. Moreover, in that case the covariance matrix Σ can be estimated by the corresponding sample covariance matrix

$$\hat{\Sigma}_N := \frac{1}{N-1} \sum_{i=1}^{N} \left[\nabla_x F(\overline{x}, \xi^i) - \nabla \hat{f}_N(\overline{x})\right]\left[\nabla_x F(\overline{x}, \xi^i) - \nabla \hat{f}_N(\overline{x})\right]^T. \quad (4.18)$$

Under the above assumptions, the sample covariance matrix $\hat{\Sigma}_N$ is an unbiased and consistent estimator of Σ.

We have that if $V_N \Rightarrow N(0, \Sigma)$ and the covariance matrix Σ is nonsingular, then given a consistent estimator $\hat{\Sigma}_N$ of Σ, the following holds

$$N(\gamma_N(\overline{x}) - \nabla f(\overline{x}))^T \hat{\Sigma}_N^{-1}(\gamma_N(\overline{x}) - \nabla f(\overline{x})) \Rightarrow \chi_n^2, \quad (4.19)$$

where χ_n^2 denotes the chi-square distribution with n degrees of freedom. This allows to construct the following (approximate) $100(1-\alpha)\%$ confidence region[17] for $\nabla f(\overline{x})$:

$$\left\{ z \in \mathbb{R}^n : (\gamma_N(\overline{x}) - z))^T \hat{\Sigma}_N^{-1}(\gamma_N(\overline{x}) - z)) \leq \frac{\chi^2_{\alpha,n}}{N} \right\}. \quad (4.20)$$

Consider the statistic

$$T_N := N \inf_{z \in K(\overline{x})} (\gamma_N(\overline{x}) - z)^T \hat{\Sigma}_N^{-1}(\gamma_N(\overline{x}) - z). \quad (4.21)$$

[17] Here $\chi^2_{\alpha,n}$ denotes the α-critical value of chi-square distribution with n degrees of freedom. That is, if $Y \sim \chi_n^2$, then $\text{Prob}\{Y \geq \chi^2_{\alpha,n}\} = \alpha$.

Note that since the cone $K(\bar{x})$ is polyhedral and $\hat{\Sigma}_N^{-1}$ is positive definite, the minimization in the right hand side of (4.21) can be formulated as a quadratic programming problem, and hence can be solved by standard quadratic programming algorithms. We have that the confidence region, defined in (4.20), does not have common points with the cone $K(\bar{x})$ iff $T_N > \chi^2_{\alpha,n}$. We can also use the statistic T_N for testing the hypothesis:

$$H_0: \nabla f(\bar{x}) \in K(\bar{x}) \quad \text{against the alternative} \quad H_1: \nabla f(\bar{x}) \notin K(\bar{x}). \tag{4.22}$$

The T_N statistic represents the squared distance, with respect to the norm[18] $\|\cdot\|_{\hat{\Sigma}_N^{-1}}$, from $N^{1/2}\gamma_N(\bar{x})$ to the cone $K(\bar{x})$. Suppose for the moment that only equality constraints are present in the definition (4.13) of the feasible set, and that the gradient vectors $\nabla g_j(\bar{x})$, $j=1,\ldots,q$, are linearly independent. Then the set $K(\bar{x})$ forms a linear subspace of $\mathbb{R}^n$ of dimension q, and the optimal value of the right hand side of (4.21) can be written in a closed form. Consequently, it is possible to show that T_N has asymptotically noncentral chi-square distribution with $n-q$ degrees of freedom and the noncentrality parameter[19]

$$\delta := N \inf_{z \in K(\bar{x})} (\nabla f(\bar{x}) - z)^T \Sigma^{-1} (\nabla f(\bar{x}) - z). \tag{4.23}$$

In particular, under H_0 we have that $\delta = 0$, and hence the null distribution of T_N is asymptotically central chi-square with $n-q$ degrees of freedom.

Consider now the general case where the feasible set is defined by equality and inequality constraints as in (4.13). Suppose that the gradient vectors $\nabla g_j(\bar{x})$, $j \in \mathcal{J}(\bar{x})$, are linearly independent and that the *strict complementarity* condition holds at $\bar{x}$, that is, the Lagrange multipliers λ_j, $j \in \mathcal{I}(\bar{x})$, corresponding to the active at $\bar{x}$ inequality constraints, are positive. Then for $\gamma_N(\bar{x})$ sufficiently close to $\nabla f(\bar{x})$ the minimizer in the right hand side of (4.21) will be lying in the linear space generated by vectors $\nabla g_j(\bar{x})$, $j \in \mathcal{J}(\bar{x})$. Therefore, in such case the null distribution of T_N is asymptotically central chi-square with $\nu := n - |\mathcal{J}(\bar{x})|$ degrees of freedom. Consequently, for a computed value T_N^* of the statistic T_N we can calculate (approximately) the corresponding p-value, which is equal to $\text{Prob}\{Y \geq T_N^*\}$, where $Y \sim \chi^2_\nu$. This p-value gives an indication of the quality of the candidate solution $\bar{x}$ with respect to the stochastic precision.

It should be understood that by accepting (i.e., failing to reject) H_0, we do not claim that the KKT conditions hold exactly at $\bar{x}$. By accepting H_0 we rather assert that we cannot separate $\nabla f(\bar{x})$ from $K(\bar{x})$, given precision of the

[18] For a positive definite matrix A, the norm $\|\cdot\|_A$ is defined as $\|z\|_A := (z^T A z)^{1/2}$.

[19] Note that under the alternative (i.e., if $\nabla f(\bar{x}) \notin K(\bar{x})$), the noncentrality parameter δ tends to infinity as $N \to \infty$. Therefore, in order to justify the above asymptotics one needs a technical assumption known as Pitman's parameter drift.

generated sample. That is, statistical error of the estimator $\gamma_N(\overline{x})$ is bigger than the squared $\|\cdot\|_{\Sigma^{-1}}$-norm distance between $\nabla f(\overline{x})$ and $K(\overline{x})$. Also rejecting H_0 does not necessarily mean that $\overline{x}$ is a poor candidate for an optimal solution of the true problem. The calculated value of T_N statistic can be large, i.e., the p-value can be small, simply because the estimated covariance matrix $N^{-1}\hat{\Sigma}_N$ of $\gamma_N(\overline{x})$ is "small". In such cases, $\gamma_N(\overline{x})$ provides an accurate estimator of $\nabla f(\overline{x})$ with the corresponding confidence region (4.20) being "small". Therefore, the above p-value should be compared with the size of the confidence region (4.20), which in turn is defined by the size of the matrix $N^{-1}\hat{\Sigma}_N$ measured, for example, by its eigenvalues. Note also that it may happen that $|\mathcal{J}(\overline{x})| = n$, and hence $\nu = 0$. Under the strict complementarity condition, this means that $\nabla f(\overline{x})$ lies in the interior of the cone $K(\overline{x})$, which in turn is equivalent to the condition that $f'(\overline{x}, d) \geq c\|d\|$ for some $c > 0$ and all $d \in \mathbb{R}^n$. Then, by the LD principle (see (4.15) in particular), the event $\gamma_N(\overline{x}) \in K(\overline{x})$ happens with probability approaching one exponentially fast.

Let us remark again that the above testing procedure is applicable if $F(\cdot, \xi)$ is differentiable at $\overline{x}$ w.p.1 and the interchangeability formula (4.17) holds. This typically happens in cases where the corresponding random data have a continuous distribution. On the other hand, under the conditions (C1)–(C2), of Section 3.3, and condition (C3) which assumes that the distribution of ξ has a finite support, the expectation function $f(\cdot)$ typically is nondifferentiable at an optimal solution, in which case the above statistical KKT test is not applicable. We discuss such cases in the next section.

4.3 Testing optimality conditions in nondifferentiable cases

Suppose that the feasible set X is defined by constraints in the form (4.13) with the index sets $\mathcal{I}(x)$ and $\mathcal{J}(x)$ defined in the same way as in the previous section. Suppose further that for every $\xi \in \Xi$ the function $F(\cdot, \xi)$ is convex. Then the function $f(\cdot)$ is convex and first order optimality conditions, at a point $x^* \in X$, can be written in the form: there exist Lagrange multipliers λ_j such that $\lambda_j \geq 0$, $j \in \mathcal{I}(x^*)$, and

$$0 \in \partial f(x^*) + \sum_{j \in \mathcal{J}(x^*)} \lambda_j \nabla g_j(x^*). \tag{4.24}$$

If the true problem is convex, i.e., the functions g_j, $j = 1, \ldots, q$, are affine and g_j, $j = q+1, \ldots, p$, are convex, the above conditions are sufficient for x^* to be an optimal solution of the true problem. Under a constraint qualification, conditions (4.24) are also necessary optimality conditions. If the function $f(x)$ is differentiable at x^*, i.e., the subdifferential $\partial f(x^*)$ is a singleton, $\partial f(x^*) = \{\nabla f(x^*)\}$, then conditions (4.24) coincide with the KKT optimality conditions (4.14). In this section we discuss cases where $\partial f(x^*)$ possibly is not a singleton, i.e., $f(x)$ is not necessarily differentiable at x^*.

For a point $\bar{x} \in X$ define

$$\Delta(\bar{x}) := \inf_{\gamma \in \partial f(\bar{x})} \operatorname{dist}(\gamma, K(\bar{x})), \tag{4.25}$$

where $K(x)$ is the cone defined in (4.15). We have that the optimality conditions (4.24) hold at the point $\bar{x}$ iff $\Delta(\bar{x}) = 0$.

It is natural to try to test optimality conditions (4.24) by replacing the subdifferential $\partial f(x^*)$ with its sample average estimate $\partial \hat{f}_N(x^*)$, and consequently estimating $\Delta(\bar{x})$ by

$$\hat{\Delta}_N(\bar{x}) := \inf_{\gamma \in \partial \hat{f}_N(\bar{x})} \operatorname{dist}(\gamma, K(\bar{x})). \tag{4.26}$$

There are, however, two basic problems with such an approach. First, in order to calculate the estimate $\hat{\Delta}_N(\bar{x})$ one needs to compute the whole subdifferential $\partial \hat{f}_N(\bar{x})$. Second, the mapping (multifunction) $x \mapsto \partial f(x)$ is not continuous unless $\partial f(x)$ is a singleton. Take, for example, $f(x) := |x|$, $x \in \mathbb{R}$. This function has unique minimizer $x^* = 0$ with $\partial f(x^*) = [-1, 1]$. On the other hand for $\bar{x} \neq 0$, the subdifferential $\partial f(\bar{x})$ is either $\{1\}$ or $\{-1\}$, and hence $\Delta(\bar{x}) = 1$, does not matter how close $\bar{x}$ to x^*. This makes testing optimality conditions (4.24) in nondifferentiable cases really difficult.

We give below some results about convergence of the subdifferentials $\partial \hat{f}_N(x)$ which have an independent interest.

Proposition 17. *Suppose that the sample is* iid, *for every* $\xi \in \Xi$ *the function* $F(\cdot, \xi)$ *is convex, and the expected value function* $f(x)$ *is well defined and finite valued in a neighborhood of a point* $\bar{x} \in \mathbb{R}^n$. *Then*

$$\lim_{N \to \infty} \mathbb{H}\Big(\partial \hat{f}_N(\bar{x}), \partial f(\bar{x})\Big) = 0 \quad \text{w.p.1.} \tag{4.27}$$

Proof. As it was shown in section "Expectation Functions" of chapter "Optimality and Duality in Stochastic Programming", we have here that $f(x)$ is directionally differentiable at $\bar{x}$ and

$$f'(\bar{x}, d) = \mathbb{E}\Big[F'_\xi(\bar{x}, d)\Big]. \tag{4.28}$$

We also have that

$$\hat{f}'_N(\bar{x}, d) = \frac{1}{N} \sum_{i=1}^{N} F'_{\xi^i}(\bar{x}, d). \tag{4.29}$$

Therefore, by the LLN it follows from (4.28) that for any $d \in \mathbb{R}^n$, $\hat{f}'_N(\bar{x}, d)$ converges to $f'(\bar{x}, d)$ w.p.1 as $N \to \infty$. Consequently, for any countable set $D \subset \mathbb{R}^n$ we have that the event

$$\lim_{N\to\infty} \hat{f}'_N(\bar{x}, d) = f'(\bar{x}, d), \quad \forall\, d \in D,$$

happens w.p.1. Let us take a countable and dense subset D of $\mathbb{R}^n$. Since the functions $\hat{f}'_N(\bar{x}, \cdot)$ are convex, it follows by Theorem 25 that $\hat{f}'_N(\bar{x}, \cdot)$ converges to $f'(\bar{x}, \cdot)$ w.p.1 uniformly on the unit ball $\{d\colon \|d\| \leq 1\}$. Furthermore, we have

$$\mathbb{H}\Big(\partial \hat{f}_N(\bar{x}),\, \partial f(\bar{x})\Big) = \sup_{\|d\|\leq 1} |\hat{f}'_N(\bar{x}, d) - f'(\bar{x}, d)|, \tag{4.30}$$

and hence (4.27) follows. □

Note that the assumption that the sample is iid in the above proposition was used only to ensure that the LLN, for the random variable $F'_\xi(\bar{x}, d)$, holds pointwise, i.e., for any (fixed) $d \in \mathbb{R}^n$.

Under the polyhedricity assumptions (C2) and (C3) of Section 3.3, it is possible to say more. In that case the functions $f(\cdot)$ and $\hat{f}_N(\cdot)$ are polyhedral and hence their subdifferentials are polyhedrons. Suppose, further, that $f(x)$ is finite valued in a neighborhood of a point $\bar{x}$. Then the convergence in (4.27) is uniform on a neighborhood of $\bar{x}$. Also because of the polyhedral structure of the functions $F(\cdot, \xi)$, $\xi \in \Xi$, we have here that the space $\mathbb{R}^n$ can be partitioned into a union of polyhedral cones T_m, $m = 1, \ldots, M$, such that $f(\cdot)$ is linear on each $\bar{x} + T_m$ in a neighborhood of $\bar{x}$, and there is a one-to-one correspondence between vertices (extreme points) of $\partial f(\bar{x})$ and $\nabla f(x)$ for $x \in \bar{x} + T_m \setminus \{0\}$ sufficiently close to $\bar{x}$. Similarly, every sample average function $\hat{f}_N(\cdot)$ is linear on unions of $\bar{x} + T_m$ in a neighborhood of $\bar{x}$. Consequently, the number of vertices of $\partial \hat{f}_N(\bar{x})$ is always less than or equal to the number of vertices of $\partial f(\bar{x})$. Moreover, because of the convergence (4.27) it follows that w.p.1 for N large enough there is a one-to-one correspondence between vertices of $\partial \hat{f}_N(\bar{x})$ and $\partial f(\bar{x})$, and vertices of $\partial \hat{f}_N(\bar{x})$ converge to the corresponding vertices of $\partial f(\bar{x})$ w.p.1 as $N \to \infty$.

This suggests the following procedure for estimating $\partial \hat{f}_N(\bar{x})$, and hence the number $\hat{\Delta}_N(\bar{x})$. Generate L points $x_1, \ldots, x_L$ randomly, say uniformly or normally, distributed in a small neighborhood of $\bar{x}$. At each point calculate the gradient (subgradient) of $\hat{f}_N$, and consequently estimate $\hat{\Delta}_N(\bar{x})$ by replacing $\partial \hat{f}_N(\bar{x})$ with the convex hull of $\nabla \hat{f}_N(x_\ell)$, $\ell = 1, \ldots, L$. It is not difficult to show that the convex hull of $\nabla \hat{f}_N(x_\ell)$, $\ell = 1, \ldots, L$, will converge w.p.1 to $\partial \hat{f}_N(\bar{x})$ as $L \to \infty$. However, numerical experiments indicate that the convergence is slow and one needs a large number L for a test, based on such procedure, to be reasonably accurate. Therefore, even for moderate

values of the dimension n such procedure appears to be not practical. For numerical experiments and an additional discussion see Linderoth et al. (2002).

Remark 18. It is also possible to derive the Large Deviations principle for the Hausdorff distance between $\partial\hat{f}_N(\bar{x})$ and $\partial f(\bar{x})$. Let us consider for the sake of simplicity the polyhedral case, i.e., suppose that the assumptions (C2) and (C3) of Section 3.3 hold and the expected value function $f(x)$ is finite valued in a neighborhood of a point $\bar{x} \in \mathbb{R}^n$. Then there exist a finite number of directions $d_1, \ldots, d_J$, independent of the sample, such that the supremum in the right hand side of (4.30) is attained at one of d_j. This can be shown essentially in the same way as in the proof of Lemma 14. Now let us choose a constant $\varepsilon > 0$. By Cramér's LD Theorem we have that for every d_j, $j = 1, \ldots, J$, there exists a constant $\beta_j > 0$ such that

$$\lim_{N\to\infty} \frac{1}{N} \log\Big[\mathbb{P}\Big(|\hat{f}'_N(\bar{x}, d_j) - f'(\bar{x}, d_j)| \geq \varepsilon\Big)\Big] = -\beta_j. \tag{4.31}$$

Because of (4.30) we have that the event $\{\mathbb{H}(\partial\hat{f}_N(\bar{x}), \partial f(\bar{x})) \geq \varepsilon\}$ coincides with the union of the events $\{|\hat{f}'_N(\bar{x}, d_j) - f'(\bar{x}, d_j)| \geq \varepsilon\}$, $j = 1, \ldots, J$. It follows then that

$$\lim_{N\to\infty} \frac{1}{N} \log\Big[\mathbb{P}\Big\{\mathbb{H}\Big(\partial\hat{f}_N(\bar{x}), \partial f(\bar{x})\Big) \geq \varepsilon\Big\}\Big] = -\beta, \tag{4.32}$$

where $\beta := \min_{1\leq j\leq J} \beta_j$. In a general, not necessarily polyhedral, case the estimate (4.32) can be obtained, under certain regularity conditions, by applying an infinite dimensional form of the LD principle in the functional space of continuous real valued functions defined on the unit sphere of $\mathbb{R}^n$.

5 Variance reduction techniques

Consider the sample average estimators $\hat{f}_N(x)$. We have that if the sample is iid, then the variance of $\hat{f}_N(x)$ is equal to $\sigma^2(x)/N$, where $\sigma^2(x) := \mathbb{V}\text{ar}[F(x, \xi)]$. In some cases it is possible to reduce the variance of generated sample averages, which in turn enhances convergence of the corresponding SAA estimators. It is beyond the scope of this chapter to give a complete survey of such variance reduction techniques. Therefore, we briefly discuss in this section a few variance reduction approaches which seem to be useful in the SAA method. For an interval $[a, b] \subset \mathbb{R}$, we denote by $U[a, b]$ the uniform probability distribution on that interval.

5.1 Latin hypercube sampling

Suppose that the random data vector $\xi = \xi(\omega)$ is one dimensional with the corresponding cdf $G(\cdot)$. We can then write

$$\mathbb{E}[F(x, \xi)] = \int_{-\infty}^{+\infty} F(x, \xi)\, \mathrm{d}G(\xi). \tag{5.1}$$

In order to evaluate the above integral numerically it will be much better to generate sample points evenly distributed than to use an iid sample. That is, we can generate independent random points

$$U^i \sim U[(i-1)/N, i/N], \quad i = 1, \ldots, N, \tag{5.2}$$

and then to construct the random sample of ξ by the inverse[20] transformation $\xi^i := G^{-1}(U^i)$, $i = 1, \ldots, N$.

Now suppose that i is chosen at random from the set $\{1, \ldots, N\}$ (with equal probability for each element of that set). Then conditional on i the corresponding random variable U^i is uniformly distributed on the interval $[(i-1)/N, i/N]$, and the unconditional distribution of U^i is uniform on the interval $[0, 1]$. Consequently, let $\{i_1, \ldots, i_N\}$ be a random permutation of the set $\{1, \ldots, N\}$. Then the random variables $\xi^{i_1}, \ldots, \xi^{i_N}$ have the same marginal distribution, with the same cdf $G(\cdot)$, and are negatively correlated with each other. Therefore, the expected value of

$$\hat{f}_N(x) = \frac{1}{N} \sum_{i=1}^{N} F(x, \xi^i) = \frac{1}{N} \sum_{s=1}^{N} F(x, \xi^{i_s}) \tag{5.3}$$

is $f(x)$, while[21]

$$\mathbb{V}\mathrm{ar}\left[\hat{f}_N(x)\right] = N^{-1}\sigma^2(x) + 2N^{-2} \sum_{s<t} \mathbb{C}\mathrm{ov}\big(F(x, \xi^{i_s}), F(x, \xi^{i_t})\big). \tag{5.4}$$

If the function $F(x, \cdot)$ is monotonically increasing or decreasing, than the random variables $F(x, \xi^{i_s})$ and $F(x, \xi^{i_t})$, $s \neq t$, are also negatively correlated. Therefore, the variance of $\hat{f}_N(x)$ tends to be smaller, and in some cases much smaller, than $\sigma^2(x)/N$.

[20] From the theoretical point of view such inverse transformation always exists although it can be difficult to calculate numerically.

[21] By $\mathbb{C}\mathrm{ov}(X, Y)$ and $\mathbb{C}\mathrm{orr}(X, Y)$ we denote the covariance and correlation, respectively, between random variables X and Y.

Suppose now that the random vector $\xi = (\xi_1, \ldots, \xi_d)$ is d-dimensional, and that its components ξ_j, $j = 1, \ldots, d$, are distributed independently of each other. Then we can use the above procedure for each component ξ_j. That is, a random sample U^i of the form (5.2) is generated, and consequently N replications of the first component of ξ are computed by the corresponding inverse transformation applied to randomly permuted U^{i_s}. The same procedure is applied to every component of ξ with the corresponding random samples of the form (5.2) and random permutations generated independently of each other. This sampling scheme is called the *Latin Hypercube* (LH) sampling.

If the function $F(x, \cdot)$ is decomposable, i.e., $F(x, \xi) := F_1(x, \xi_1) + \cdots + F_d(x, \xi_d)$, then $\mathbb{E}[F(x, \xi)] = \mathbb{E}[F_1(x, \xi_1)] + \cdots + \mathbb{E}[F_d(x, \xi_d)]$, where each expectation is calculated with respect to a one dimensional distribution. In that case the LH sampling ensures that each expectation $\mathbb{E}[F_j(x, \xi_j)]$ is estimated in nearly optimal way. Therefore, the LH sampling works especially well in cases where the function $F(x, \cdot)$ tends to have a somewhat decomposable structure. In any case the LH sampling procedure is easy to implement and can be applied to SAA optimization procedures in a straightforward way. Since in LH sampling the random replications of $F(x, \xi)$ are correlated with each other, one cannot use variance estimates like (2.21) or (4.5). Therefore, the LH method usually is applied in several independent batches in order to estimate variance of the corresponding estimators.

5.2 Linear control random variables method

Suppose that we have a measurable function $A(x, \xi)$ such that $\mathbb{E}[A(x, \xi)] = 0$ for all $x \in X$. Then, for any $t \in \mathbb{R}$, the expected value of $F(x, \xi) + tA(x, \xi)$ is $f(x)$, while

$$\begin{aligned}&\mathbb{V}\mathrm{ar}[F(x, \xi) + tA(x, \xi)]\\ &\quad = \mathbb{V}\mathrm{ar}[F(x, \xi)] + t^2\, \mathbb{V}\mathrm{ar}[A(x, \xi)] + 2t\, \mathbb{C}\mathrm{ov}(F(x, \xi), A(x, \xi)).\end{aligned}$$

It follows that the above variance attains its minimum, with respect to t, for

$$t^* := -\rho_{F,A}(x)\left[\frac{\mathbb{V}\mathrm{ar}(F(x, \xi))}{\mathbb{V}\mathrm{ar}(A(x, \xi))}\right]^{1/2}, \tag{5.5}$$

where $\rho_{F,A}(x) := \mathbb{C}\mathrm{orr}(F(x, \xi), A(x, \xi))$, and with

$$\mathbb{V}\mathrm{ar}[F(x, \xi) + t^* A(x, \xi)] = \mathbb{V}\mathrm{ar}[F(x, \xi)][1 - \rho_{F,A}(x)^2]. \tag{5.6}$$

For a given $x \in X$ and generated sample $\xi^1, \ldots, \xi^N$, one can estimate, in the standard way, the covariance and variances appearing in the right hand side of (5.5), and hence to construct an estimate $\hat{t}$ of t^*. Then $f(x)$ can be estimated by

$$\hat{f}_N^A(x) := \frac{1}{N} \sum_{i=1}^{N} \left[F(x, \xi^i) + \hat{t} A(x, \xi^i) \right]. \tag{5.7}$$

By (5.6), the *linear control* estimator $\hat{f}_N^A(x)$ has a smaller variance than $\hat{f}_N(x)$ if $F(x, \boldsymbol{\xi})$ and $A(x, \boldsymbol{\xi})$ are highly correlated with each other.

Let us make the following observations. The estimator $\hat{t}$, of the optimal value t^*, depends on x and the generated sample. Therefore, it is difficult to apply linear control estimators in an SAA optimization procedure. That is, linear control estimators are mainly suitable for estimating expectations at a fixed point. Also if the same sample is used in estimating $\hat{t}$ and $\hat{f}_N^A(x)$, then $\hat{f}_N^A(x)$ can be a slightly biased estimator of $f(x)$.

Of course, the above Linear Control procedure can be successful only if a function $A(x, \xi)$, with mean zero and highly correlated with $F(x, \xi)$, is available. Choice of such a function is problem dependent. For instance, one can use a linear function $A(x, \xi) := \lambda(\xi)^T x$ (compare with the discussion at the end of Section 4.1). Consider, for example, two-stage stochastic programming with recourse. Suppose that the random vector $\boldsymbol{h} = h(\omega)$ and matrix $\boldsymbol{T} = T(\omega)$, in the second stage problem (1.2), are independently distributed, and let $\mu := \mathbb{E}(\boldsymbol{h})$. Then

$$\mathbb{E}[(\boldsymbol{h} - \mu)^T \boldsymbol{T}] = \mathbb{E}[(\boldsymbol{h} - \mu)]^T \mathbb{E}[\boldsymbol{T}] = 0,$$

and hence one can use $A(x, \xi) := (h - \mu)^T T x$ as the control variable.

Let us finally remark that the above procedure can be extended in a straightforward way to a case where several functions $A_1(x, \xi), \ldots, A_m(x, \xi)$, each with zero mean and highly correlated with $F(x, \xi)$, are available.

5.3 *Importance sampling and likelihood ratio methods*

Suppose that $\boldsymbol{\xi}$ has a continuous distribution with probability density function (pdf) $g(\cdot)$. Let $\psi(\cdot)$ be another pdf such that the so-called *likelihood ratio* function $L(\cdot) := g(\cdot)/\psi(\cdot)$ is well defined. That is, if $\psi(z) = 0$ for some $z \in \mathbb{R}^d$, then $g(z) = 0$, and by the definition, $0/0 = 0$, i.e., we do not divide a positive number by zero. Then we can write

$$f(x) = \int F(x, \xi) g(\xi) \, \mathrm{d}\xi = \int F(x, \zeta) L(\zeta) \psi(\zeta) \, \mathrm{d}\zeta = \mathbb{E}_\psi[F(x, \boldsymbol{\zeta}) L(\boldsymbol{\zeta})], \tag{5.8}$$

where the integration is performed over the space $\mathbb{R}^d$ and the notation $\mathbb{E}_\psi$ emphasizes that the expectation is taken with respect to the random vector ζ having pdf $\psi(\cdot)$.

Let us show that, for a fixed x, the variance of $F(x,\zeta)L(\zeta)$ attains its minimal value for $\psi(\cdot)$ proportional to $|F(x,\cdot)g(\cdot)|$, i.e., for

$$\psi^*(\cdot) := \frac{|F(x,\cdot)g(\cdot)|}{\int |F(x,\zeta)g(\zeta)|\,\mathrm{d}\zeta}. \tag{5.9}$$

Since $\mathbb{E}_\psi[F(x,\zeta)L(\zeta)] = f(x)$ and does not depend on $\psi(\cdot)$, we have that the variance of $F(x,\zeta)L(\zeta)$ is minimized if

$$\mathbb{E}_\psi[F(x,\zeta)^2 L(\zeta)^2] = \int \frac{F(x,\zeta)^2 g(\zeta)^2}{\psi(\zeta)}\,\mathrm{d}\zeta, \tag{5.10}$$

is minimized. Furthermore, by Cauchy inequality we have

$$\left(\int |F(x,\zeta)g(\zeta)|\,\mathrm{d}\zeta\right)^2 \le \left(\int \frac{F(x,\zeta)^2 g(\zeta)^2}{\psi(\zeta)}\,\mathrm{d}\zeta\right)\left(\int \psi(\zeta)\,\mathrm{d}\zeta\right). \tag{5.11}$$

It remains to note that $\int \psi(\zeta)\,\mathrm{d}\zeta = 1$ and the left hand side of (5.11) is equal to the expected value of squared $F(x,\zeta)L(\zeta)$ for $\psi(\cdot) = \psi^*(\cdot)$.

Note that if $F(x,\cdot)$ is nonnegative valued, then $\psi^*(\cdot) = F(x,\cdot)g(\cdot)/f(x)$ and for that choice of the pdf $\psi(\cdot)$, the function $F(x,\cdot)L(\cdot)$ is identically equal to $f(x)$. Of course, in order to achieve such absolute variance reduction to zero we need to know the expectation $f(x)$ which was our goal in the first place. Nevertheless, it gives the idea that if we can construct a pdf $\psi(\cdot)$ roughly proportional to $|F(x,\cdot)g(\cdot)|$, then we may achieve a considerable variance reduction by generating a random sample $\zeta^1,\ldots,\zeta^N$ from the pdf $\psi(\cdot)$, and then estimating $f(x)$ by

$$\tilde{f}_N^{\psi}(x) := \frac{1}{N}\sum_{i=1}^{N} F(x,\zeta^i)L(\zeta^i). \tag{5.12}$$

The estimator $\tilde{f}_N^{\psi}(x)$ is an unbiased estimator of $f(x)$ and may have significantly smaller variance than $\hat{f}_N(x)$ depending on a successful choice of the pdf $\psi(\cdot)$.

Similar analysis can be performed in cases where ξ has a discrete distribution by replacing the integrals with the corresponding summations.

Let us remark that the above approach, called *importance sampling*, is extremely sensitive to a choice of the pdf $\psi(\cdot)$ and is notorious for its instability. This is understandable since the likelihood ratio function in the tail

is the ratio of two very small numbers. For a successful choice of $\psi(\cdot)$, the method may work very well while even a small perturbation of $\psi(\cdot)$ may be disastrous. This is why a single choice of $\psi(\cdot)$ usually does not work for different points x, and consequently cannot be used for a whole optimization procedure. Note also that $\mathbb{E}_{\psi}[L(\zeta)] = 1$. Therefore, $L(\zeta) - 1$ can be used as a linear control variable for the likelihood ratio estimator $\tilde{f}_N^{\psi}(x)$.

In some cases it is also possible to use the likelihood ratio method for estimating first and higher order derivatives of $f(x)$. Consider, for example, the optimal value $Q(x, \xi)$ of the second stage linear program (1.2). Suppose that the vector q and matrix W are fixed, i.e., not stochastic, and for the sake of simplicity that $\boldsymbol{h} = h(\omega)$ and $\boldsymbol{T} = T(\omega)$ are distributed independently of each other. We have then that $Q(x, \xi) = \mathcal{Q}(h - Tx)$, where

$$\mathcal{Q}(z) := \inf\{q^T y : Wy = z, y \geq 0\}.$$

Suppose, further, that $\boldsymbol{h}$ has a continuous distribution with pdf $g(\cdot)$. We have that

$$\mathbb{E}[Q(x, \boldsymbol{\xi})] = \mathbb{E}_T\{\mathbb{E}_{h|T}[Q(x, \boldsymbol{\xi})]\},$$

and by using the transformation $z = h - Tx$, since $\boldsymbol{h}$ and $\boldsymbol{T}$ are independent we obtain

$$\begin{aligned} \mathbb{E}_{h|T}[Q(x, \boldsymbol{\xi})] = \mathbb{E}_h[Q(x, \boldsymbol{\xi})] &= \int \mathcal{Q}(h - Tx) g(h)\, \mathrm{d}h = \int \mathcal{Q}(z) g(z + Tx)\, \mathrm{d}z \\ &= \int \mathcal{Q}(\zeta) L(x, \zeta) \psi(\zeta)\, \mathrm{d}\zeta = \mathbb{E}_{\psi}[L(x, \boldsymbol{\zeta}) \mathcal{Q}(\boldsymbol{\zeta})], \end{aligned} \tag{5.13}$$

where $\psi(\cdot)$ is a chosen pdf and $L(x, \zeta) := g(\zeta + Tx)/\psi(\zeta)$. If the function $g(\cdot)$ is smooth, then the likelihood ratio function $L(\cdot, \zeta)$ is also smooth. In that case, under mild additional conditions, first and higher order derivatives can be taken inside the expected value in the right hand side of (5.13), and consequently can be estimated by sampling. Note that the first order derivatives of $Q(\cdot, \xi)$ are piecewise constant, and hence its second order derivatives are zeros whenever defined. Therefore, second order derivatives cannot be taken inside the expectation $\mathbb{E}[Q(x, \boldsymbol{\xi})]$ even if $\boldsymbol{\xi}$ has a continuous distribution.

6 Multistage stochastic programming

Analysis developed in the previous sections can be applied to two-stage stochastic programming in a straightforward way. However, for multistage programming the situation is more subtle. In this section we discuss T-stage linear stochastic programming problems with recourse of the form[22]

$$\underset{\substack{A_{11}x_1=b_1\\ x_1\geq 0}}{\operatorname{Min}} c_1x_1+\mathbb{E}\left[\min_{\substack{A_{21}x_1+A_{22}x_2=b_2\\ x_2\geq 0}} c_2x_2+\mathbb{E}\left[\cdots+\mathbb{E}\left[\min_{\substack{A_{T,T-1}x_{T-1}+A_{TT}x_T=b_T\\ x_T\geq 0}} c_Tx_T\right]\right]\right], \tag{6.1}$$

driven by the random data process $\xi_2,\ldots,\xi_T$. Here $x_t\in\mathbb{R}^{n_t}$, $t=1,\ldots,T$, are decision variables, $\xi_1:=(c_1,A_{11},b_1)$ is known at the first stage (and hence is nonrandom), and $\xi_t:=(c_t,A_{t,t-1},A_{tt},b_t)\in\mathbb{R}^{d_t}$, $t=2,\ldots,T$, are data vectors some (all) elements of which can be random. Such multistage problems were discussed in section "Multistage Models" of chapter "Stochastic Programming Models", notation and terminology of which we follow.

If we denote by $Q_2(x_1,\xi_2)$ the optimal value of the $(T-1)$-stage problem

$$\underset{\substack{A_{21}x_1+A_{22}x_2=b_2\\ x_2\geq 0}}{\operatorname{Min}} c_2x_2+\mathbb{E}\left[\cdots+\mathbb{E}\left[\min_{\substack{A_{T,T-1}x_{T-1}+A_{TT}x_T=b_T\\ x_T\geq 0}} c_Tx_T\right]\right], \tag{6.2}$$

then we can write the T-stage problem (6.1) in the following form of two-stage programming problem

$$\operatorname{Min}_{x_1} c_1x_1+\mathbb{E}[Q_2(x_1,\xi_2)] \quad \text{subject to} \quad A_{11}x_1=b_1,\ x_1\geq 0. \tag{6.3}$$

Note, however, that if $T\geq 3$, then problem (6.2) in itself is a stochastic programming problem. Consequently, if the number of scenarios involved in (6.2) is very large, or infinite, then the optimal value $Q_2(x_1,\xi_2)$ may be calculated only approximately, say by sampling.

There are several ways how sampling can be applied to multistage programming. Sampling strategies are closely related to derivations of upper and lower statistical bounds for the optimal value of the true problem. We discuss that in the next section.

[22] For the sake of notational convenience we write in this section the scalar product between two vectors $c,x\in\mathbb{R}^n$ as cx instead of c^Tx.

6.1 Statistical bounds

For given x_1 and ξ_2, the corresponding expected value(s) can be estimated by generating random samples and solving the obtained SAA problems. Let us observe that in case $T \geq 3$, it follows from (2.22) that for any estimator $\hat{Q}_2(x_1, \xi_2)$ of $Q_2(x_1, \xi_2)$ obtained in that way the following relation holds[23]

$$Q_2(x_1, \xi_2) \geq \mathbb{E}\Big[\hat{Q}_2(x_1, \boldsymbol{\xi}_2)|\boldsymbol{\xi}_2 = \xi_2\Big] \tag{6.4}$$

for every feasible x_1 and ξ_2. That is, for $T \geq 3$ any SAA estimator of $Q_2(x_1, \xi_2)$ is biased downwards.

In order to get a better insight into the above problem of bias let us assume for the sake of simplicity that $T = 3$. In that case problem (6.2) becomes the two-stage program

$$\operatorname*{Min}_{x_2} c_2 x_2 + \mathbb{E}[Q_3(x_2, \boldsymbol{\xi}_3)|\boldsymbol{\xi}_2] \quad \text{subject to} \quad A_{21}x_1 + A_{22}x_2 = b_2,\ x_2 \geq 0, \tag{6.5}$$

where $Q_3(x_2, \xi_3)$ is the optimal value of the problem

$$\operatorname*{Min}_{x_3} c_3 x_3 \quad \text{subject to} \quad A_{32}x_2 + A_{33}x_3 = b_3,\ x_3 \geq 0. \tag{6.6}$$

The corresponding first stage problem is then (6.3) with $Q_2(x_1, \xi_2)$ given by the optimal value of (6.5). The functions $Q_2(\cdot\,, \xi_2)$ and $Q_3(\cdot\,, \xi_3)$ are extended real valued convex functions for any ξ_2 and ξ_3.

Let us note that if we relax the nonanticipativity constraints at the second stage of the above three-stage problem we obtain the following two-stage program

$$\operatorname*{Min}_{x_1 \in X_1} c_1 x_1 + \mathbb{E}[Q(x_1, \boldsymbol{\xi}_2, \boldsymbol{\xi}_3)], \tag{6.7}$$

where

$$X_1 := \{x_1 \in \mathbb{R}^{n_1} : A_{11}x_1 = b_1,\ x_1 \geq 0\},$$

[23] The notation $\mathbb{E}[\cdot\,|\boldsymbol{\xi} = \xi]$ denotes the conditional expectation given the event "$\boldsymbol{\xi} = \xi$".

and $Q(x_1, \xi_2, \xi_3)$ is the optimal value of the following problem

$$\begin{aligned} &\operatorname*{Min}_{x_2, x_3} c_2 x_2 + c_3 x_3 \\ &\text{subject to} \quad A_{21} x_1 + A_{22} x_2 = b_2, \\ &\qquad\qquad A_{32} x_2 + A_{33} x_3 = b_3, \\ &\qquad\qquad x_2 \geq 0, \, x_3 \geq 0. \end{aligned} \tag{6.8}$$

Since (6.7)–(6.8) is obtained by a relaxation of the nonanticipativity constraints, its optimal value is smaller than the optimal value of the corresponding three-stage problem.

There are several ways how one can sample from the random data $\boldsymbol{\xi}_2$, $\boldsymbol{\xi}_3$ (recall that ξ_1 is not random). Let P be the probability distribution of the random vector $(\boldsymbol{\xi}_2, \boldsymbol{\xi}_3)$. Suppose that a random sample

$$\left(\xi_2^i, \xi_3^i\right) = \left((c_2^i, A_{21}^i, A_{22}^i, b_2^i), (c_3^i, A_{32}^i, A_{33}^i, b_3^i)\right) \sim P, \quad i = 1, \ldots, N, \tag{6.9}$$

of N replications of the random data is generated. Suppose, further, that for each (ξ_2^i, ξ_3^i) the corresponding linear programming problem

$$\begin{aligned} &\operatorname*{Min}_{x_1, x_2, x_3} c_1 x_1 + c_2^i x_2 + c_3^i x_3 \\ &\text{subject to} \quad A_{11} x_1 = b_1, \\ &\qquad\qquad A_{21}^i x_1 + A_{22}^i x_2 = b_2^i, \\ &\qquad\qquad A_{32}^i x_2 + A_{33}^i x_3 = b_3^i, \\ &\qquad\qquad x_1 \geq 0, \, x_2 \geq 0, \, x_3 \geq 0. \end{aligned} \tag{6.10}$$

is solved, and the obtained optimal values are averaged.

For each $i \in \{1, \ldots, N\}$, the above problem (6.10) is equivalent to the problem

$$\operatorname*{Min}_{x_1 \in X_1} c_1 x_1 + Q(x_1, \xi_2^i, \xi_3^i). \tag{6.11}$$

Similar to (2.22), we have that

$$\inf_{x_1 \in X_1} \mathbb{E}[c_1 x_1 + Q(x_1, \boldsymbol{\xi}_2, \boldsymbol{\xi}_3)] \geq \mathbb{E}\left[\inf_{x_1 \in X_1} \{c_1 x_1 + Q(x_1, \boldsymbol{\xi}_2, \boldsymbol{\xi}_3)\}\right]. \tag{6.12}$$

We also have that the average of the optimal values of (6.10) is an unbiased and consistent estimator of the right hand side of (6.12). Recall that the left

hand side of (6.12) is the optimal value of two-stage relaxation (6.7)–(6.8) of the considered three-stage problem. It follows that the average of the optimal values of (6.10) provides a valid, but *not* consistent, statistical lower bound for the optimal value of (6.7)–(6.8), and hence for the considered three-stage problem.

Suppose that $\mathbb{E}[Q(x_1, \xi_2, \xi_3)]$ is finite. Then by the LLN we have that

$$c_1 x_1 + \frac{1}{N}\sum_{i=1}^{N} Q(x_1, \xi_2^i, \xi_3^i) \to c_1 x_1 + \mathbb{E}[Q(x_1, \xi_2, \xi_3)] \quad \text{w.p.1 as } N \to \infty, \tag{6.13}$$

and the expected value of the left hand side of (6.13) is equal to the right hand side of (6.13). Therefore, for any feasible $x_1 \in X_1$ of the first stage problem, the left hand side of (6.13) provides a valid upper statistical bound for the optimal value of the two-stage problem (6.7)–(6.8). However, as we discussed earlier, the optimal value of (6.7)–(6.8) is smaller than the optimal value of the corresponding three-stage problem. Therefore, there is no guarantee that the left hand side of (6.13) gives a *valid* upper statistical bound for the optimal value of the three-stage problem.

In order to improve the above statistical lower bound let us consider the following optimization problem

$$\operatorname*{Min}_{x_1 \in X_1} c_1 x_1 + \frac{1}{N}\sum_{i=1}^{N} Q(x_1, \xi_2^i, \xi_3^i). \tag{6.14}$$

Problem (6.14) can be considered as a two-stage program with scenarios (ξ_2^i, ξ_3^i), $i = 1, \ldots, N$, having equal probabilities N^{-1}. We have that the optimal value of (6.14) gives a valid and consistent statistical lower bound for the two-stage problem (6.7)–(6.8). Yet for the three-stage problem it gives a valid, but not consistent[24] lower statistical bound.

Let us observe that if the number of scenarios of the considered (true) three-stage problem is finite, then some of the generated second stage vectors ξ_2^i can be equal to each other. In that case we can view the generated sample as a scenario tree and associate with it a (sample) three-stage stochastic programming problem. The program (6.14) becomes then a two-stage relaxation of the obtained three-stage sample program. Note, however, that if the number of scenarios K_1 at the second stage of the considered (true) three-stage problem is very large, then the probability that some of ξ_2^i are equal to each other is very small unless the sample size N is comparable with K_1. For example,

[24] Unless the optimal value of the three-stage problem coincides with the optimal value of its two-stage relaxation (6.7) and (6.8).

if each scenario at the second stage can happen with equal probability K_1^{-1}, then the probability that at least two of $\boldsymbol{\xi}_2^i$ are equal to each other is

$$\rho_N = 1 - \prod_{i=1}^{N-1} \left(1 - \frac{i}{K_1}\right) \approx 1 - e^{-N(N-1)/(2K_1)}.$$

In order to attain ρ_N at a given level $\alpha \in (0, 1)$, one needs then a sample of size $N \approx \sqrt{2K_1 \log[(1-\alpha)^{-1}]}$. This shows that if K_1 is very large, then trying to construct a scenario tree by sampling from the distribution of the random vector $(\boldsymbol{\xi}_2, \boldsymbol{\xi}_3)$ does not help much in improving the statistical lower bound derived from the program (6.14). Moreover, if $\boldsymbol{\xi}_2$ has a continuous distribution, then the probability that some of $\boldsymbol{\xi}_2^i$ are equal to each other is zero for any sample size. Therefore, in that case such sampling will never produce a scenario tree structure.

The above discussion shows that a valid upper statistical bound cannot be obtained by a straightforward sampling. In order to compute such an upper bound one needs to construct an implementable and feasible policy. Recall that a sequence of mappings $x_t(\cdot)$, $t = 1, \ldots, T$, is called an *implementable* policy if $x_1(\cdot) = x_1 \in \mathbb{R}^{n_1}$ and each $x_t(\cdot) \in \mathbb{R}^{n_t}$, $t = 2, \ldots, T$, is a function of x_1 and the history $\xi_{[1,t]} := (\xi_1, \ldots, \xi_t)$ of the process up to time t. An implementable policy is feasible if it satisfies, w.p.1., the corresponding feasibility constraints at each stage $t = 1, \ldots, T$. Given any implementable and feasible[25] policy $x_1, x_2(x_1, \xi_{[1,2]}), \ldots, x_T(x_1, \xi_{[1,T]})$, the expectation

$$\mathbb{E}[c_1 x_1 + c_2 x_2(x_1, \boldsymbol{\xi}_{[1,2]}) + \cdots + c_T x_T(x_1, \boldsymbol{\xi}_{[1,T]})] \tag{6.15}$$

provides an upper bound for the optimal value of the true multistage problem. The above expectation can be estimated by the average

$$\frac{1}{N} \sum_{i=1}^{N} [c_1 x_1 + c_2^i x_2(x_1, \xi_{[1,2]}^i) + \cdots + c_T^i x_T(x_1, \xi_{[1,T]}^i)] \tag{6.16}$$

for a generated sample $\xi_2^i, \ldots, \xi_T^i$ of N realizations of the random process $\boldsymbol{\xi}_2, \ldots, \boldsymbol{\xi}_T$.

In the case of two-stage problem (i.e., for $T = 2$), one can construct an implementable and feasible policy by choosing a feasible first stage point $x_1 \in X_1$ and computing $x_2(x_1, \xi_2)$ as an optimal solution of the corresponding second stage problem[26]

$$\operatorname*{Min}_{x_2} c_2 x_2 \quad \text{subject to} \quad A_{21} x_1 + A_{22} x_2 = b_2,\ x_2 \geq 0. \tag{6.17}$$

[25] Recall that ξ_1 is not random. Therefore, the decision variables actually do not depend on ξ_1, we write ξ_1 in $\xi_{[1,t]}$ for uniformity of notation. In particular, $x_2(x_1, \xi_{[1,2]}) = x_2(x_1, \xi_2)$.

[26] Recall that $\xi_2 = (c_2, A_{21}, A_{22}, b_2)$.

For such choice of the policy, we have that $c_2x_2(x_1, \xi_2)$ is equal to the optimal value of the above problem (6.17), and hence the statistical upper bound (6.16) is the same as the one used in Section 4.1. For $T > 2$ one would like to construct $x_2(x_1, \xi_2)$ as an optimal solution of the problem (6.2), and so on for $x_3(x_1, \xi_{[1,3]})$, etc. Note, however, that this would require solving multistage programming problems and numerically may be infeasible.

To summarize the discussion of this section we can say the following:

- Any SAA method provides a valid statistical lower bound for the optimal value of the true stochastic program. However, direct sampling from the distribution of the random data vector $\boldsymbol{\xi} = (\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \ldots, \boldsymbol{\xi}_T)$ does not give a consistent statistical lower bound if $T \geq 3$.
- If the (total) number K of scenarios is finite, then by direct sampling from the scenario population, one can eventually reconstruct the scenario tree of the true problem. However, for $T \geq 3$ and K very large, the sample size which will be required for a reasonable approximation of the true scenario tree would be comparable with K.
- For $T \geq 3$, by taking a feasible point $x_1 \in X_1$ of the first stage program and then applying the SAA procedure in a straightforward way, as it was discussed above, to a generated random sample of $\boldsymbol{\xi}$, does not give a valid statistical upper bound for the optimal value of the corresponding multistage program.
- In order to compute a valid statistical upper bound one needs to construct an implementable and feasible policy. However, for $T \geq 3$, it could be difficult to construct such a policy which will provide a tight and numerically feasible upper bound.

In order to improve these bounds one needs to increase the sample size at every stage conditionally on the scenarios generated at the previous stage. We discuss this in the next section.

6.2 *Conditional sampling of multistage programs*

For the sake of simplicity we discuss in this section the linear multistage program (6.1) with $T = 3$. Let us generate a random sample in the following way. First we generate a random sample

$$\xi_2^i = (c_2^i, A_{21}^i, A_{22}^i, b_2^i), \quad i = 1, \ldots, N_1,$$

of N_1 replications of the random vector $\boldsymbol{\xi}_2$. Then for every $i \in \{1, \ldots, N_1\}$, we generate a random sample

$$\xi_3^{ij} = (c_3^{ij}, A_{32}^{ij}, A_{33}^{ij}, b_3^{ij}), \quad j = 1, \ldots, N_2,$$

from the conditional distribution of $\boldsymbol{\xi}_3$ given the event $\{\boldsymbol{\xi}_2 = \xi_2^i\}$. In that way we obtain the following three-stage stochastic program

$$\operatorname*{Min}_{x_1} c_1 x_1 + \frac{1}{N_1}\sum_{i=1}^{N_1} \hat{Q}_{2,N_2}(x_1, \xi_2^i) \quad \text{subject to} \quad A_{11}x_1 = b_1,\ x_1 \geq 0, \tag{6.18}$$

where $\hat{Q}_{2,N_2}(x_1, \xi_2^i)$ is the optimal value of

$$\operatorname*{Min}_{x_2} c_2^i x_2 + \frac{1}{N_2}\sum_{j=1}^{N_2} Q_3(x_2, \xi_3^{ij}) \quad \text{subject to} \quad A_{21}^i x_1 + A_{22}^i x_2 = b_2^i,\ x_2 \geq 0, \tag{6.19}$$

with $Q_3(x_2, \xi_3)$ being the optimal value of the problem (6.6).

We refer to the above sampling scheme as the *conditional* sampling. The sample size of third stage scenarios, associated with each second stage scenario, does not need to be the same, we assumed it to be constant for the sake of simplicity. The constructed three-stage stochastic programming problem (6.18)–(6.19) has $N = N_1 N_2$ scenarios, each with equal probability $1/N$. It can be noted that for any fixed $j \in \{1, \ldots, N_2\}$ in the above conditional sampling, the corresponding sample (ξ_2^i, ξ_3^{ij}), $i = 1, \ldots, N_1$, is a random sample of the type (6.9), i.e., is derived from the distribution of the random vector $(\boldsymbol{\xi}_2, \boldsymbol{\xi}_3)$. Therefore, if $N_2 = 1$, then the above conditional sampling becomes the same as the sampling (6.9). Note also that at this stage we do not specify how the conditional samples ξ_3^{ij} are generated. For example, we do not necessarily assume that for different i, $k \in \{1, \ldots, N_1\}$ the corresponding random samples $\boldsymbol{\xi}_3^{ij}$ and $\boldsymbol{\xi}_3^{kj}$, $j = 1, \ldots, N_2$, are independent of each other conditional on $\boldsymbol{\xi}_2^i$ and $\boldsymbol{\xi}_2^k$, respectively.

As it was discussed in the previous section (see (6.4) in particular), we have that

$$\begin{aligned} Q_2(x_1, \xi_2^i) &= \inf_{\substack{A_{21}^i x_1 + A_{22}^i x_2 = b_2^i \\ x_2 \geq 0}} \left\{ c_2^i x_2 + \mathbb{E}\left[\frac{1}{N_2}\sum_{j=1}^{N_2} Q_3(x_2, \boldsymbol{\xi}_3^{ij}) \middle| \boldsymbol{\xi}_2 = \xi_2^i \right] \right\} \\ &\geq \mathbb{E}\left[\inf_{\substack{A_{21}^i x_1 + A_{22}^i x_2 = b_2^i \\ x_2 \geq 0}} \left\{ c_2^i x_2 + \frac{1}{N_2}\sum_{j=1}^{N_2} Q_3(x_2, \boldsymbol{\xi}_3^{ij}) \right\} \middle| \boldsymbol{\xi}_2 = \xi_2^i \right] \\ &= \mathbb{E}\left[\hat{Q}_{2,N_2}(x_1, \boldsymbol{\xi}_2) | \boldsymbol{\xi}_2 = \xi_2^i \right]. \end{aligned} \tag{6.20}$$

We also have that

$$\inf_{x_1 \in X_1} \{c_1 x_1 + \mathbb{E}[Q_2(x_1, \xi_2)]\} \geq \mathbb{E}\left[\inf_{x_1 \in X_1} \left\{c_1 x_1 + \frac{1}{N_1} \sum_{i=1}^{N_1} Q_2(x_1, \xi_2^i)\right\}\right]. \tag{6.21}$$

It follows from (6.20) and (6.21) that the optimal value $\hat{v}_{N_1,N_2}$ of the first stage (6.18), of the problem (6.18)–(6.19), gives a valid statistical lower bound for the optimal value v^* of the corresponding (true) three-stage stochastic programming problem.

As it was shown in Section 2.1 (see Theorem 4 in particular), under mild boundedness conditions, for any fixed x_1 and ξ_2^i the estimator $\hat{Q}_{2,N_2}(x_1, \xi_2^i)$ converges to $Q_2(x_1, \xi_2^i)$ w.p.1 as N_2 tends to infinity. Therefore, it is natural to expect that $\hat{v}_{N_1,N_2} \to v^*$ w.p.1 as $N_1 \to \infty$ and $N_2 \to \infty$, i.e., that $\hat{v}_{N_1,N_2}$ is a consistent estimator of v^*. And, indeed, it is possible to show that this holds true under some regularity conditions. Note, however, that although conceptually important such consistency result in itself is insufficient for a justification of the conditional sampling. If, in the case of a finite number of scenarios, the sample size, which is required for a reasonably accurate approximation of the true problem, is comparable with the total number of scenarios, one does not need sampling. Unfortunately, at this moment we do not have a useful theory or numerical evidence for rates of convergence of Monte Carlo sampling methods in multistage programming. Note also that for T-stage problems the above conditional sampling, with the corresponding branching of sizes $N_1, N_2, \ldots, N_{T-1}$, results in $N = N_1 N_2 \cdots N_{T-1}$ number of scenarios. That is, in conditional sampling the total number of scenarios of SAA problems grows fast with increase of the number of stages.

6.3 An example of financial planning

In this section we discuss the example of "Financial Planning" described in section "Examples of Multistage Models" of the chapter "Stochastic Programming Models", to which we refer for details. We now briefly recall the basic model. Let $\boldsymbol{R}_t = (\boldsymbol{R}_{1t}, \ldots, \boldsymbol{R}_{nt})$, $t = 1, \ldots, T$, be a random process representing returns of n investment opportunities, and $U(W)$ be a chosen utility function. For the sake of simplicity we assume that the process $\boldsymbol{R}_t$ is *Markovian*, i.e., for all $t = 1, \ldots, T-1$, the conditional distribution of $\boldsymbol{R}_{t+1}$ given $(\boldsymbol{R}_1, \ldots, \boldsymbol{R}_t)$ is the same as the conditional distribution of $\boldsymbol{R}_{t+1}$ given $\boldsymbol{R}_t$. The associated (multistage) stochastic programming problem is defined by the cost-to-go functions $Q_t(x_{t-1}, R_t)$, $t = 1, \ldots, T-1$, and the following first stage

problem at $t=0$,

$$\begin{aligned} &\underset{x_0}{\text{Max}}\ \mathbb{E}[Q_1(x_0, \boldsymbol{R}_1)] \\ &\text{s.t.}\quad \sum_{i=1}^{n} x_{i0} = W_0, \\ &\qquad x_{i0} \geq 0,\ i = 1, \ldots, n. \end{aligned} \tag{6.22}$$

Here W_0 is the initial wealth, the function $Q_{T-1}(x_{T-2}, R_{T-1})$ is the optimal value of the problem

$$\begin{aligned} &\underset{x_{T-1}}{\text{Max}}\ \mathbb{E}\left\{ U\left[\sum_{i=1}^{n} (1+\boldsymbol{R}_{iT})x_{i,T-1}\right] \middle| \boldsymbol{R}_{T-1} = R_{T-1} \right\} \\ &\text{s.t.}\quad \sum_{i=1}^{n} x_{i,T-1} = \sum_{i=1}^{n} (1+R_{i,T-1})x_{i,T-2}, \\ &\qquad x_{i,T-1} \geq 0,\ i = 1, \ldots, n, \end{aligned} \tag{6.23}$$

and $Q_t(x_{t-1}, R_t)$ is the optimal value of the problem

$$\begin{aligned} &\underset{x_t}{\text{Max}}\ \mathbb{E}[Q_{t+1}(x_t, \boldsymbol{R}_{t+1}) | \boldsymbol{R}_t = R_t] \\ &\text{s.t.}\quad \sum_{i=1}^{n} x_{it} = \sum_{i=1}^{n} (1+R_{it})x_{i,t-1}, \\ &\qquad x_{it} \geq 0,\ i = 1, \ldots, n, \end{aligned} \tag{6.24}$$

for $t=T-2, \ldots, 1$. Note that the above is a maximization rather than minimization problem.

As it was mentioned in the chapter "Stochastic Programming Models", the cost-to-go function $Q_t(x_{t-1}, R_t)$ depends on x_{t-1} through $W_t = \sum_{i=1}^{n} (1+R_{it})x_{i,t-1}$ only. That is, if $\tilde{Q}_t(W_t, R_t)$ is the optimal value of the problem

$$\begin{aligned} &\underset{x_t}{\text{Max}}\ \mathbb{E}[Q_{t+1}(x_t, \boldsymbol{R}_{t+1}) | \boldsymbol{R}_t = R_t] \\ &\text{s.t.}\quad \sum_{i=1}^{n} x_{it} = W_t, \\ &\qquad x_{it} \geq 0,\ i = 1, \ldots, n, \end{aligned} \tag{6.25}$$

then $Q_t(x_{t-1}, R_t) = \tilde{Q}_t(\sum_{i=1}^{n} (1+R_{it})x_{i,t-1}, R_t)$.

We can define the following implementable and feasible policy for the above problem. Choose a feasible vector $x_0 \in \mathbb{R}^n$ for the problem (6.22), and define

$$x_{it} := (1 + R_{it})x_{i,t-1} = \left[\prod_{\tau=1}^{t} (1 + R_{i\tau})\right] x_{i0}, \quad t = 1, \ldots, T-1.$$

We have that x_t is a function of $R_{[1,t]} := (R_1, \ldots, R_t)$ and x_0, and hence defines an implementable policy, and satisfies the constraints of (6.24) and hence is feasible. This policy corresponds to the initial investment x_0 without rebalancing the portfolio at the later stages. It gives the following lower bound (remember that this is a maximization problem) for the multistage problem (6.22)–(6.24):

$$\mathbb{E}\left\{U\left(\sum_{i=1}^{n}\left[\prod_{t=1}^{T}(1 + \boldsymbol{R}_{it})\right]x_{i0}\right)\right\}. \tag{6.26}$$

In order to investigate the value of this lower bound suppose further that $U(W) \equiv W$. We have then that

$$Q_{T-1}(x_{T-2}, R_{T-1}) = M_{T-1}(R_{T-1}) \sum_{i=1}^{n} (1 + R_{i,T-1})x_{i,T-2}, \tag{6.27}$$

where $M_{T-1}(R_{T-1}) := \max_{1\le i\le n} \mathbb{E}[1 + \boldsymbol{R}_{iT} | \boldsymbol{R}_{T-1} = R_{T-1}]$,

$$Q_{T-2}(x_{T-3}, R_{T-2}) = M_{T-2}(R_{T-2}) \sum_{i=1}^{n} (1 + R_{i,T-2})x_{i,T-3}, \tag{6.28}$$

where

$$\begin{aligned} M_{T-2}(R_{T-2}) &:= \max_{1\le i\le n} \mathbb{E}\left\{\left(\max_{1\le i\le n} \mathbb{E}[1+\boldsymbol{R}_{iT}|\boldsymbol{R}_{T-1}]\right)(1+\boldsymbol{R}_{i,T-1})|\boldsymbol{R}_{T-2}=R_{T-2}\right\} \\ &= \max_{1\le i\le n} \mathbb{E}\left\{M_{T-1}(\boldsymbol{R}_{T-1})(1 + \boldsymbol{R}_{i,T-1})|\boldsymbol{R}_{T-2} = R_{T-2}\right\}, \end{aligned}$$

and so on with

$$Q_t(x_{t-1}, R_t) = M_t(R_t) \sum_{i=1}^{n} (1 + R_{it})x_{i,t-1}, \tag{6.29}$$

and $M_t(R_t) := \max_{1\le i\le n} \mathbb{E}\{M_{t+1}(\boldsymbol{R}_{t+1})(1+\boldsymbol{R}_{i,t+1})|\boldsymbol{R}_t = R_t\}$, $t = T-2,\ldots,1$.

In particular, suppose for the moment that the process $\boldsymbol{R}_t$ is between stages independent. Then the optimal value v^* of the multistage problem is given by

$$v^* = W_0 \prod_{t=1}^{T} \left(\max_{1\le i\le n} \mu_{it}\right), \tag{6.30}$$

where $\mu_{it} := 1 + \mathbb{E}[\boldsymbol{R}_{it}]$. On the other hand, maximization of the expectation (6.26), subject to the feasibility constraints of problem (6.22), gives the following "best" lower bound of the form (6.26):

$$v_L = W_0 \max_{1\le i\le n}\left(\prod_{t=1}^{T} \mu_{it}\right). \tag{6.31}$$

Therefore, one can easily construct examples where the ratio v^*/v_L can be arbitrary large. It should be not surprising that this ratio could be large in cases where the maximal value of the expected return moves from one asset to another with the stages of the process. Recall that, in the considered case $U(W)\equiv W$, the multistage problem can be solved in a completely myopic fashion by investing everything in the asset with the maximal expected return in the next period. Therefore, if $\mu_{it} = \mu_i$ do not depend on t, then there is no need to rebalance the portfolio and $v^* = v_L = W_0(\max_{1\le i\le n} \mu_i)^T$.

Suppose now that we relax the nonanticipativity constraints starting from the second stage of the multistage problem (see section "The Case of Finitely Many Scenarios" of chapter "Stochastic Programming Models" for a discussion of such relaxation). The obtained two-stage stochastic program has the following optimal value

$$v_U = W_0 \mathbb{E}\left[\prod_{t=1}^{T} \max_{1\le i\le n} (1+\boldsymbol{R}_{it})\right], \tag{6.32}$$

which gives an upper bound for v^*. If we assume, further, the between stages independence, then $v_U = W_0 \prod_{t=1}^{T} \mathbb{E}[\max_{1\le i\le n}(1+\boldsymbol{R}_{it})]$. We have that, for all t,

$$\mathbb{E}\left[\max_{1\le i\le n} (1+\boldsymbol{R}_{it})\right] \ge \max_{1\le i\le n} \mathbb{E}[1+\boldsymbol{R}_{it}] = \max_{1\le i\le n} \mu_{it}. \tag{6.33}$$

The difference between the left and right hand sides of (6.33) tends to be bigger if there is a larger number of assets with the maximal (or nearly) maximal expected return.

7 Stochastic generalized equations

In this section we discuss the following so-called *stochastic generalized equations*. Consider a random vector ξ, whose distribution is supported on a set $\Xi \subset \mathbb{R}^d$, a mapping $\Phi: \mathbb{R}^n \times \Xi \to \mathbb{R}^n$ and a multifunction[27] $\Gamma: \mathbb{R}^n \rightrightarrows \mathbb{R}^n$. Suppose that the expectation $\phi(x) := \mathbb{E}[\Phi(x, \xi)]$ is well defined. We refer to

$$\phi(x) \in \Gamma(x) \tag{7.1}$$

as true, or expected value, generalized equation and say that a point $x^* \in \mathbb{R}^n$ is a solution of (7.1) if $\phi(x^*) \in \Gamma(x^*)$.

The above abstract setting includes the following cases. If $\Gamma(x) := \{0\}$ for every $x \in \mathbb{R}^n$, then (7.1) becomes the ordinary equation $\phi(x) = 0$. As another example, let $\Gamma(\cdot) := N_X(\cdot)$, where X is a nonempty closed convex subset of $\mathbb{R}^n$ and $N_X(x)$ denotes the outwards normal cone to X at x. Recall that by the definition $N_X(x) = \varnothing$ if $x \notin X$. In that case x^* is a solution of (7.1) iff $x^* \in X$ and the following, so-called *variational inequality*, holds

$$(x - x^*)^T \phi(x^*) \geq 0, \quad \forall\, x \in X. \tag{7.2}$$

Since the mapping $\phi(x)$ is given in the form of the expectation, we refer to such variational inequalities as stochastic variational inequalities. Note that if $X = \mathbb{R}^n$, then $N_X(x) = \{0\}$ for any $x \in \mathbb{R}^n$, and hence in that case the above variational inequality is reduced to the equation $\phi(x) = 0$. Let us also remark that if $\Phi(x, \xi) := -\nabla_x F(x, \xi)$, for some real valued function $F(x, \xi)$, and the interchangeability formula (4.17) holds, i.e., $\phi(x) = -\nabla f(x)$, then (7.2) represents first order necessary, and if $f(x)$ is convex, sufficient conditions for x^* to be an optimal solution for the optimization problem (1.1).

Also if the feasible set X of the optimization problem (1.1) is defined by constraints in the form (4.13), with $g_j(x) := \mathbb{E}[G_j(x, \xi)]$, $j = 1, \ldots, p$, then the corresponding KKT optimality conditions (4.14) can be written in a form of variational inequality. That is, let $z := (x, \lambda) \in \mathbb{R}^{n+p}$ and

$$L(z, \xi) := F(x, \xi) + \sum_{j=1}^{p} \lambda_j G_j(x, \xi) \tag{7.3}$$

be the Lagrangian associated with the stochastic problem (1.1). Define

$$\Phi(z, \xi) := (\nabla_x L(z, \xi), G_1(x, \xi), \ldots, G_p(x, \xi)) \quad \text{and} \quad \Gamma(z) := N_K(z), \tag{7.4}$$

[27] Recall that $\Gamma(x)$ is said to be a multifunction if it maps a point $x \in \mathbb{R}^n$ into a subset of $\mathbb{R}^n$.

where $K := \mathbb{R}^n \times \mathbb{R}^q \times \mathbb{R}^{p-q}_+ \subset \mathbb{R}^{n+p}$. Then assuming that the interchangeability formula holds, we have that

$$\phi(z) := \mathbb{E}[\Phi(z, \xi)] = \left(\nabla f(x) + \sum_{j=1}^{p} \lambda_j \nabla g_j(x),\ g_1(x),\ \ldots,\ g_p(x) \right), \quad (7.5)$$

and hence variational inequality $\phi(z) \in N_K(z)$ represents the KKT optimality conditions for the true optimization problem.

We make the following assumption about the multifunction $\Gamma(x)$.

(E1) The multifunction $\Gamma(x)$ is *closed*, that is, the following holds: if $x_k \to x$, $y_k \in \Gamma(x_k)$ and $y_k \to y$, then $y \in \Gamma(x)$.

The above assumption implies that the multifunction $\Gamma(x)$ is closed valued, i.e., for any $x \in \mathbb{R}^n$ the set $\Gamma(x)$ is closed. For variational inequalities assumption (E1) always holds, i.e., the multifunction $x \mapsto N_X(x)$ is closed.

Now let $\xi^1, \ldots, \xi^N$ be a random sample of N realizations of the random vector ξ, and $\hat{\phi}_N(x) := N^{-1} \sum_{i=1}^{N} \Phi(x, \xi^i)$ be the corresponding sample average estimate of $\phi(x)$. We refer to

$$\hat{\phi}_N(x) \in \Gamma(x) \quad (7.6)$$

as the SAA generalized equation. There are standard numerical algorithms for solving nonlinear equations which can be applied to (7.6) in the case $\Gamma(x) \equiv \{0\}$, i.e., when (7.6) is reduced to the ordinary equation $\hat{\phi}_N(x) = 0$. More recently numerical procedures were developed for solving variational inequalities. We are not going to discuss such numerical algorithms but rather concentrate on statistical properties of solutions of SAA equations. We denote by S and $\hat{S}_N$ the sets of (all) solutions of the true (7.1) and SAA (7.6) generalized equations, respectively.

7.1 Consistency of solutions of the SAA generalized equations

In this section we discuss convergence properties of the SAA solutions. Recall that $\mathbb{D}(A, B)$ denotes the deviation of the set A from the set B.

Proposition 19. *Let C be a compact subset of $\mathbb{R}^n$ such that $S \subset C$. Suppose that:* (i) *assumption* (E1) *holds,* (ii) *the mapping $\phi(x)$ is continuous on C,* (iii) *w.p.1 for N large enough the set $\hat{S}_N$ is nonempty and $\hat{S}_N \subset C$,* (iv) *$\hat{\phi}_N(x)$ converges to $\phi(x)$ w.p.1 uniformly on C as $N \to \infty$. Then $\mathbb{D}(\hat{S}_N, S) \to 0$ w.p.1 as $N \to \infty$.*

Proof. The above result basically is deterministic in the sense that if we view $\hat{\phi}_N(x) = \hat{\phi}_N(x, \omega)$ as defined on a common probability space, then it should be

verified for a.e. ω. Therefore, we omit saying "w.p.1". Consider a sequence $\hat{x}_N \in \hat{S}_N$. Because of the compactness assumption (iii), by passing to a subsequence if necessary, we only need to show that if $\hat{x}_N$ converges to a point $\overline{x}$, then $\overline{x} \in S$ (compare with the proof of Theorem 4). Now since it is assumed that $\phi(\cdot)$ is continuous and $\hat{\phi}_N(x)$ converges to $\phi(x)$ uniformly, it follows that $\hat{\phi}_N(\hat{x}_N) \to \phi(\overline{x})$. Since $\hat{\phi}_N(\hat{x}_N) \in \Gamma(\hat{x}_N)$, it follows by assumption (E1) that $\phi(\overline{x}) \in \Gamma(\overline{x})$, which completes the proof. □

A few remarks about the assumptions involved in the above consistency result are now in order. By Proposition 7 we have that, in the case of iid sampling, the assumptions (ii) and (iv) of the above proposition are satisfied for any compact set C if the following assumption holds.

(E2) For every $\xi \in \Xi$ the function $\Phi(\cdot, \xi)$ is continuous on C, and $||\Phi(x, \xi)||$, $x \in C$, is dominated by an integrable function.

There are two parts to the assumption (iii) of Proposition 19, namely, that the SAA generalized equations do not have a solution which escapes to infinity, and that they possess at least one solution w.p.1 for N large enough. The first of these assumptions can be often verified by ad hoc methods. The second assumption is more subtle. We are going to discuss it next. The following concept of strong regularity is due to Robinson (1980).

Definition 20. Suppose that the mapping $\phi(x)$ is continuously differentiable. We say that a solution $x^* \in S$ is strongly regular if there exist neighborhoods $\mathcal{N}_1$ and $\mathcal{N}_2$ *of* $0 \in \mathbb{R}^n$ and x^*, respectively, such that for every $\delta \in \mathcal{N}_1$ the following (linearized) generalized equation

$$\delta + \phi(x^*) + \nabla\phi(x^*)(x - x^*) \in \Gamma(x) \tag{7.7}$$

has a unique solution in $\mathcal{N}_2$, denoted $\overline{x}(\delta)$, and $\overline{x}(\cdot)$ is Lipschitz continuous on $\mathcal{N}_1$.

Note that it follows from the above conditions that $\overline{x}(0) = x^*$. In case $\Gamma(x) \equiv \{0\}$, strong regularity simply means that the $n \times n$ Jacobian matrix $J := \nabla\phi(x^*)$ is invertible, or in other words nonsingular. Also in the case of variational inequalities, the strong regularity condition was investigated extensively, we discuss this later.

Let V be a compact neighborhood of x^*. Consider the space $C^1(V, \mathbb{R}^n)$ of continuously differentiable mappings $\psi: V \to \mathbb{R}^n$ equipped with the norm:

$$\|\psi\|_{1,V} := \sup_{x \in V}(\|\phi(x)\| + \|\nabla\phi(x)\|).$$

The following (deterministic) result is essentially due to Robinson (1982) (see also Bonnans and Shapiro, 2000, Theorem 5.13, and the following up discussion).

Suppose that $\phi(x)$ is continuously differentiable on V, i.e., $\phi \in C^1(V, \mathbb{R}^n)$. Let x^* be a strongly regular solution of the generalized equation (7.1). Then there exists $\varepsilon > 0$ such that for any $u \in C^1(V, \mathbb{R}^n)$ satisfying $||u - \phi||_{1,V} \leq \varepsilon$, the generalized equation $u(x) \in \Gamma(x)$ has a unique solution $\hat{x} = \hat{x}(u)$ in a neighborhood of x^*, such that $\hat{x}(\cdot)$ is Lipschitz continuous (with respect the norm $||\cdot||_{1,V}$), and

$$\hat{x}(u) = \overline{x}(u(x^*) - \phi(x^*)) + o(\|u - \phi\|_{1,V}). \tag{7.8}$$

By employing the above results for the mapping $u(\cdot) := \hat{\phi}_N(\cdot)$ we immediately obtain the following.

Proposition 21. *Let x^* be a strongly regular solution of the true generalized equation (7.1), and suppose that $\phi(x)$ and $\hat{\phi}_N(x)$ are continuously differentiable in a neighborhood V of x^* and $\|\hat{\phi}_N - \phi\|_{1,V} \to 0$ w.p.1 as $N \to \infty$. Then w.p.1 for N large enough the SAA generalized equation (7.6) possesses a unique solution $\hat{x}_N$ in a neighborhood of x^*, and $\hat{x}_N \to x^*$ w.p.1 as $N \to \infty$.*

The assumption that $\|\hat{\phi}_N - \phi\|_{1,V} \to 0$ w.p.1, in the above theorem, means that $\hat{\phi}_N(x)$ and $\nabla\hat{\phi}_N(x)$ converge w.p.1 to $\phi(x)$ and $\nabla\phi(x)$, respectively, uniformly on V. By Proposition 7, in the case of iid sampling this is ensured by the following assumption.

(E3) For every $\xi \in \Xi$ the mapping $\Phi(\cdot, \xi)$ is continuously differentiable on V, and $||\Phi(x, \xi)||$ and $||\nabla_x \Phi(x, \xi)||$, $x \in V$, are dominated by an integrable function.

Note that the assumption that $\Phi(\cdot, \xi)$ is continuously differentiable on a neighborhood of x^* is essential in the above analysis. By combining Propositions 19 and 21 we obtain the following result.

Theorem 22. *Let C be a compact subset of $\mathbb{R}^n$ and x^* be a unique in C solution of the true generalized equation (7.1). Suppose that:* (i) *assumption* (E1) *holds*, (ii) *for every $\xi \in \Xi$ the mapping $\Phi(\cdot, \xi)$ is continuously differentiable on C, and $||\Phi(x, \xi)||$ and $||\nabla_x \Phi(x, \xi)||$, $x \in C$, are dominated by an integrable function*, (iii) *the solution x^* is strongly regular*, (iv) *$\hat{\phi}_N(x)$ and $\nabla\hat{\phi}_N(x)$ converge w.p.1 to $\phi(x)$ and $\nabla\phi(x)$, respectively, uniformly on C. Then w.p.1 for N large enough the SAA generalized equation possesses unique in C solution $\hat{x}_N$ converging to x^* w.p.1 as $N \to \infty$.*

Note again that if the sample is iid, then the assumption (iv) in the above theorem is implied by the assumption (ii) and hence is redundant.

7.2 Asymptotics of SAA estimators

By using the first order approximation (7.8) it is also possible to derive asymptotics of $\hat{x}_N$. Suppose for the moment that $\Gamma(x) \equiv \{0\}$. Then strong regularity means that the Jacobian matrix $J := \nabla \phi(x^*)$ is nonsingular, and $\overline{x}(\delta)$ is the solution of the corresponding linear equations and hence can be written in the form

$$\overline{x}(\delta) = x^* - J^{-1}\delta. \tag{7.9}$$

By using (7.9) and (7.8) with $u(\cdot) := \hat{\phi}_N(\cdot)$, we obtain under certain regularity conditions which ensure that the remainder in (7.8) is of order $o_p(N^{-1/2})$, that

$$N^{1/2}(\hat{x}_N - x^*) = -J^{-1} Y_N + o_p(N^{-1/2}), \tag{7.10}$$

where $Y_N := N^{1/2}[\hat{\phi}_N(x^*) - \phi(x^*)]$. Moreover, in the case of iid sample, we have by the CLT that $Y_N \Rightarrow N(0, \Sigma)$, where Σ is the covariance matrix of the random vector $\Phi(x^*, \xi)$. Consequently, $\hat{x}_N$ has asymptotically normal distribution with mean vector x^* and the covariance matrix $N^{-1}J^{-1}\Sigma J^{-1}$.

Suppose now that $\Gamma(\cdot) := N_X(\cdot)$, with the set X being nonempty closed convex and *polyhedral*, and let x^* be a strongly regular solution of (7.1). Let $\overline{x}(\delta)$ be the (unique) solution, of the corresponding linearized variational inequality (7.7), in a neighborhood of x^*. Consider the cone

$$C(x^*) := \{y \in T_X(x^*) : y^T \phi(x^*) = 0\},$$

called the *critical cone*, and the Jacobian matrix $J := \nabla \phi(x^*)$. Then for all δ sufficiently close to $0 \in \mathbb{R}^n$, we have that $\overline{x}(\delta) - x^*$ coincides with the solution $\overline{d}(\delta)$ of the variational inequality

$$\delta + Jd \in N_{C(x^*)}(d). \tag{7.11}$$

Note that the mapping $\overline{d}(\cdot)$ is positively homogeneous, i.e., for any $\delta \in \mathbb{R}^n$ and $t \geq 0$, it follows that $\overline{d}(t\delta) = t\overline{d}(\delta)$. Similar to (7.10) we have, under regularity conditions ensuring that the remainder term is of order $o_p(N^{-1/2})$, that

$$N^{1/2}(\hat{x}_N - x^*) \Rightarrow \overline{d}(Y), \tag{7.12}$$

where $Y \sim N(0, \Sigma)$. It follows that $\hat{x}_N$ is asymptotically normal iff the mapping $\overline{d}(\cdot)$ is linear. This, in turn, holds if the cone $C(x^*)$ is a linear space.

In the case $\Gamma(\cdot) := N_X(\cdot)$, with the set X being nonempty closed convex and polyhedral, there is a complete characterization of the strong regularity in terms of the so-called *coherent orientation* associated with the matrix (mapping) $J := \nabla \phi(x^*)$ and the critical cone $C(x^*)$. The interested reader is referred to Robinson (1992) and Gürkan et al. (1999) for a discussion of this topic. Let us just remark that if $C(x^*)$ is a linear subspace of $\mathbb{R}^n$, then the variational inequality (7.11) can be written in the form

$$P\delta + PJd = 0, \tag{7.13}$$

where P denotes the orthogonal projection matrix onto the linear space $C(x^*)$. Then x^* is strongly regular iff the matrix (mapping) PJ restricted to the space $C(x^*)$ is invertible, or in other words nonsingular.

Suppose now that $S = \{x^*\}$ is such that $\phi(x^*)$ belongs to the interior of the set $\Gamma(x^*)$. Then, since $\hat{\phi}_N(x^*)$ converges w.p.1 to $\phi(x^*)$, it follows that the event $\{\hat{\phi}_N(x^*) \in \Gamma(x^*)\}$ happens w.p.1 for N large enough. Moreover, by the LD principle (see (8.15)) we have that this event happens with probability approaching one exponentially fast. Of course, $\hat{\phi}_N(x^*) \in \Gamma(x^*)$ means that $\hat{x}_N = x^*$ is a solution of the SAA generalized equation (7.6). Therefore, in such case one may compute an exact solution of the true problem (7.1), by solving the SAA problem, with probability approaching one exponentially fast with increase of the sample size. Note that if $\Gamma(\cdot) := N_X(\cdot)$ and $x^* \in S$, then $\phi(x^*) \in \operatorname{int} \Gamma(x^*)$ iff the critical cone $C(x^*)$ is equal to $\{0\}$. In that case the variational inequality (7.11) has solution $d^* = 0$ for any δ, i.e., $\overline{d}(\delta) \equiv 0$.

The above asymptotics can be applied, in particular, to the generalized equation (variational inequality) $\phi(z) \in N_X(z)$, where $\phi(z)$ and $N_X(z)$ are defined in (7.4) and (7.5). Recall that this variational inequality represents the KKT optimality conditions of the corresponding expected value optimization problem. Therefore, in that way the asymptotics of the optimal solutions and Lagrange multipliers of the associated SAA optimization problems can be derived. In the case of optimization problems strong regularity of a point $z^* = (x^*, \lambda^*)$, where $x^* \in S$, is equivalent to a certain (strong) form of second order optimality conditions. Also in that case the critical cone is a linear space, and hence the SAA estimator $\hat{z}_N = (\hat{x}_N, \hat{\lambda}_N)$ is asymptotically normal, iff the strict complementarity condition holds.

Let us finally remark the following. Consider variational inequality (7.2). Suppose that the set X is defined by

$$X := \{x \in \mathbb{R}^n : g_j(x) \leq 0, j = 1, \ldots, p\}, \tag{7.14}$$

where $g_j(x)$, $j=1,\ldots,p$, are convex real valued functions. Suppose further that the Slater condition holds. Then, for any $x \in X$,

$$N_X(x) = \left\{ z \in \mathbb{R}^n : z = \sum_{j \in \mathcal{I}(x)} \alpha_j \nabla g_j(x),\ \alpha_j \geq 0,\ j \in \mathcal{I}(x) \right\}, \tag{7.15}$$

where $\mathcal{I}(x) := \{j : g_j(x) = 0, j = 1, \ldots, p\}$. Therefore, in that case one can proceed with statistical testing of a candidate solution $\bar{x}$ in exactly the same way as in testing KKT optimality conditions in Section 4.2.

8 Appendix

8.1 Epiconvergence

Consider a sequence $f_k : \mathbb{R}^n \to \overline{\mathbb{R}}, k=1,\ldots,$ of extended real valued functions. It is said that the functions f_k *epiconverge* to a function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, written $f_k \xrightarrow{e} f$, if the epigraphs of the functions f_k converge, in a certain set valued sense, to the epigraph of f. It is also possible to define the epiconvergence in the following equivalent way.

Definition 23. It is said that f_k epiconverge to f if for any point $x \in \mathbb{R}^n$ the following two conditions hold: (i) for any sequence x_k converging to x one has

$$\liminf_{k\to\infty} f_k(x_k) \geq f(x), \tag{8.1}$$

(ii) there exists a sequence x_k converging to x such that

$$\limsup_{k\to\infty} f_k(x_k) \leq f(x). \tag{8.2}$$

Epiconvergence $f_k \xrightarrow{e} f$ implies that the function f is lower semicontinuous.

Epiconvergence is discussed extensively in Rockafellar and Wets (1998), Chapter 7. We need a few basic results from that theory. We denote by $\arg\min f$ the set of minimizers of f, i.e., $\bar{x} \in \arg\min f$ iff $f(\bar{x}) = \inf f(x)$, where the infimum is taken over all $x \in \mathbb{R}^n$. For $\varepsilon \geq 0$ we say that a point $\bar{x} \in \mathbb{R}^n$ is an ε-minimizer[28] of f if $f(\bar{x}) \leq \inf f(x) + \varepsilon$. Clearly, for $\varepsilon = 0$ the set of ε-minimizers of f coincides with $\arg\min f$.

[28] For the sake of convenience we allow in this section for a minimizer, or ε-minimizer, $\bar{x}$ to be such that $f(\bar{x})$ is not finite, i.e., can be equal to $+\infty$ or $-\infty$.

Proposition 24. *Suppose that* $f_k \overset{e}{\to} f$. *Then*

$$\limsup_{k\to\infty} [\inf f_k(x)] \le \inf f(x). \tag{8.3}$$

Suppose, further, that: (i) *for some* $\varepsilon_k \downarrow 0$ *there exists an* ε_k*-minimizer* x_k *of* $f_k(\cdot)$ *such that the sequence* x_k *converges to a point* $\bar{x}$. *Then* $\bar{x} \in \arg\min f$ *and*

$$\lim_{k\to\infty} [\inf f_k(x)] = \inf f(x). \tag{8.4}$$

Proof. Consider a point $x \in \mathbb{R}^n$ and let x_k be a sequence converging to x such that the inequality (8.2) holds. Then $f_k(x_k) \ge \inf f_k(x)$. Together with (8.2) this implies that

$$f(x) \ge \limsup_{k\to\infty} f_k(x_k) \ge \limsup_{k\to\infty} [\inf f_k(x)].$$

Since the above holds for any x, the inequality (8.3) follows.

Now let x_k be a sequence of ε_k-minimizers of f_k converging to a point $\bar{x}$. We have then that $f_k(x_k) \le \inf f_k(x) + \varepsilon_k$, and hence by (8.3) we obtain

$$\liminf_{k\to\infty} [\inf f_k(x)] = \liminf_{k\to\infty} [\inf f_k(x) + \varepsilon_k] \ge \liminf_{k\to\infty} f_k(x_k) \ge f(\bar{x}) \ge \inf f(x).$$

Together with (8.3) this implies (8.4) and $f(\bar{x}) = \inf f(x)$. This completes the proof. □

Assumption (i) in the above proposition can be ensured by various boundedness conditions.

The following result is taken from Rockafellar and Wets (1998), Theorem 7.17.

Theorem 25. *Let* $f_k : \mathbb{R}^n \to \overline{\mathbb{R}}$ *be a sequence of convex functions and* $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ *be a convex lower semicontinuous function such that* $\operatorname{dom} f$ *has a nonempty interior. Then the following are equivalent:* (i) $f_k \overset{e}{\to} f$, (ii) *there exists a dense subset* D *of* $\mathbb{R}^n$ *such that* $f_k(x) \to f(x)$ *for all* $x \in D$, (iii) $f_k(\cdot)$ *converges uniformly to* $f(\cdot)$ *on every compact set* C *that does not contain a boundary point of* $\operatorname{dom} f$.

8.2 Uniform integrability, and $O_p(\cdot)$ *and* $o_p(\cdot)$ *notation*

Let Y_k, $k = 1, \ldots,$ be a sequence of random variables converging in distribution to a random variable Y. In general, convergence in distribution does not imply convergence of the expected values $\mathbb{E}[Y_k]$ to $\mathbb{E}[Y]$, as $k \to \infty$,

even if all these expected values are finite. This implication holds under the additional condition that Y_k are *uniformly integrable*, that is

$$\lim_{c\to\infty} \sup_{k\in\mathbb{N}} \mathbb{E}[Y_k(c)] = 0, \tag{8.5}$$

where $Y_k(c) := |Y_k|$ if $|Y_k| \geq c$, and $Y_k(c) := 0$ otherwise. A simple sufficient condition ensuring uniform integrability, and hence the implication that $Y_k \Rightarrow Y$ implies $\mathbb{E}[Y_k] \to \mathbb{E}[Y]$, is the following: there exists $\varepsilon > 0$ such that $\sup_k \mathbb{E}[|Y_k|^{1+\varepsilon}] < \infty$.

The notation $O_p(\cdot)$ and $o_p(\cdot)$ stands for a probabilistic analogue of the usual order notation $O(\cdot)$ and $o(\cdot)$, respectively. That is, let X_k and Y_k be sequences of random variables. It is written that $Y_k = O_p(X_k)$ if for any $\varepsilon > 0$ there exists $c > 0$ such that $\mathbb{P}(|Y_k/X_k| > c) \leq \varepsilon$ for all $k \in \mathbb{N}$. It is written that $Y_k = o_p(X_k)$ if for any $\varepsilon > 0$ it holds that $\lim_{k\to\infty} \mathbb{P}(|Y_k/X_k| > \varepsilon) = 0$. Usually this is used with the sequence X_k being deterministic. In particular, the notation $Y_k = O_p(1)$ asserts that the sequence Y_k is bounded in probability, and $Y_k = o_p(1)$ means that the sequence Y_k converges in probability to zero.

8.3 Large deviations theory

Consider an iid sequence $Y_1, \ldots, Y_N$ of replications of a real valued random variable Y, and let $Z_N := N^{-1} \sum_{i=1}^N Y_i$ be the corresponding sample average. Then for any real numbers a and $t > 0$ we have that $\mathbb{P}(Z_N \geq a) = \mathbb{P}(e^{tZ_N} \geq e^{ta})$, and hence, by Chebyshev's inequality

$$\mathbb{P}(Z_N \geq a) \leq e^{-ta}\mathbb{E}[e^{tZ_N}] = e^{-ta}[M(t/N)]^N,$$

where $M(t) := \mathbb{E}[e^{tY}]$ is the *moment generating function* of Y. Suppose that Y has finite mean $\mu := \mathbb{E}[Y]$ and let $a \geq \mu$. By taking the logarithm of both sides of the above inequality, changing variables $t' = t/N$ and minimizing over $t' > 0$, we obtain

$$\frac{1}{N} \log[\mathbb{P}(Z_N \geq a)] \leq -I(a), \tag{8.6}$$

where

$$I(z) := \sup_{t\in\mathbb{R}} \{tz - \Lambda(t)\} \tag{8.7}$$

is the conjugate of the logarithmic moment generating function $\Lambda(t) := \log M(t)$. In the LD theory, $I(z)$ is called the (*large deviations*) *rate* function, and the inequality (8.6) corresponds to the upper bound of Cramér's LD Theorem.

Note that the constraint $t>0$ is removed in the above definition of the rate function $I(\cdot)$. This is because of the following. Consider the function $\psi(t):=ta-\Lambda(t)$. The function $\Lambda(t)$ is convex, and hence $\psi(t)$ is concave. Suppose that the moment generating function $M(\cdot)$ is finite valued at some $\bar{t}>0$. Then $M(t)$ is finite for all $t\in[0,\bar{t}]$, and it follows by the Dominated Convergence Theorem that $M(t)$, and hence the function $\Lambda(t)$, are right side differentiable at $t=0$. Moreover, the right side derivative of $M(t)$ at $t=0$ is μ, and hence the right side derivative of $\psi(t)$ at $t=0$ is positive if $a>\mu$. Consequently, in that case $\psi(t)>\psi(0)$ for all $t>0$ small enough, and hence $I(a)>0$ and the supremum in (8.7) is not changed if the constraint $t>0$ is removed. If $a=\mu$, then the supremum in (8.7) is attained at $t=0$ and hence $I(a)=0$. In that case the inequality (8.6) trivially holds. Now if $M(t)=+\infty$ for all $t>0$, then $I(a)=0$ for any $a\geq\mu$ and the inequality (8.6) trivially holds.

Note that for $a\leq\mu$ the upper bound (8.6) takes the form

$$\frac{1}{N}\log[\mathbb{P}(Z_N\leq a)]\leq -I(a). \tag{8.8}$$

The rate function $I(z)$ is convex and has the following properties. Suppose that the random variable Y has finite mean μ. Then $I(\mu)=0$ and $I(z)$ attains its minimum at $z=\mu$. Suppose, further, that the moment generating function $M(t)$ is finite valued for all t in a neighborhood of $t=0$. Then it follows by the Dominated Convergence Theorem that $M(t)$, and hence the function $\Lambda(t)$, are infinitely differentiable at $t=0$, and $\Lambda'(0)=\mu$ and $\Lambda''(0)=\sigma^2$, where $\sigma^2:=\mathbb{V}\mathrm{ar}[Y]$. It follows by the above discussion that in that case $I(a)>0$ for any $a\neq\mu$. We also have then that $I'(\mu)=0$ and $I''(\mu)=\sigma^{-2}$, and hence by Taylor's expansion,

$$I(a)=\frac{(a-\mu)^2}{2\sigma^2}+o(|a-\mu|^2). \tag{8.9}$$

Consequently, for a close to μ we can approximate $I(a)$ by $(a-\mu)^2/2\sigma^2$. Moreover, for any $\varepsilon>0$ there is a neighborhood $\mathcal{N}$ of μ such that

$$I(a)\geq\frac{(a-\mu)^2}{(2+\varepsilon)\sigma^2},\quad\forall\, a\in\mathcal{N}. \tag{8.10}$$

In particular, we can take $\varepsilon=1$.

The constant $I(a)$ in (8.6) gives, in a sense, the best possible exponential rate at which the probability $\mathbb{P}(Z_N\geq a)$ converges to zero. This follows from the

lower bound

$$\liminf_{N\to\infty} \frac{1}{N} \log[\mathbb{P}(Z_N \geq a)] \geq -I(a) \tag{8.11}$$

of Cramér's LD Theorem, which holds for $a \geq \mu$.

We also need to consider cases where the random variable Y can take value $+\infty$ with positive probability p. In such a case $\mathbb{E}[Y] = +\infty$ and $Z_N = +\infty$ if at least one $Y_i = +\infty$. We have then that for any $a \in \mathbb{R}$,

$$\mathbb{P}(Z_N \leq a) = (1-p)^N \mathbb{P}\{Z_N \leq a | Y_i < +\infty, i = 1, \ldots, N\}. \tag{8.12}$$

Let Y_0 be a random variable with probability distribution given by the probability distribution of Y conditional on $Y < +\infty$, and let $I_0(\cdot)$ be the large deviations rate function of Y_0. It follows then from (8.8) and (8.12) that

$$\frac{1}{N} \log[\mathbb{P}(Z_N \leq a)] \leq -\big[\log(1-p)^{-1} + \varrho(a) I_0(a)\big], \tag{8.13}$$

where $\varrho(a) := 1$ if $a < \mathbb{E}[Y_0]$, and $\varrho(a) := 0$ if $a \geq \mathbb{E}[Y_0]$.

The above, one dimensional, LD results can be extended to multivariate and even infinite dimensional settings, and also to non iid random sequences. In particular, suppose that Y is a d-dimensional random vector and let $\mu := \mathbb{E}[Y]$ be its mean vector. We can associate with Y its moment generating function $M(t)$, of $t \in \mathbb{R}^d$, and the rate function $I(z)$ defined in the same way as in (8.7) with the supremum taken over $t \in \mathbb{R}^d$ and tz denoting the standard scalar product of vectors $t, z \in \mathbb{R}^d$. Consider a (Borel) measurable set $A \subset \mathbb{R}^d$. Then, under certain regularity conditions, the following Large Deviations Principle holds:

$$\begin{aligned} -\inf_{z\in \mathrm{int}(A)} I(z) &\leq \liminf_{N\to\infty} N^{-1} \log[\mathbb{P}(Z_N \in A)] \\ &\leq \limsup_{N\to\infty} N^{-1} \log[\mathbb{P}(Z_N \in A)] \leq -\inf_{z\in \mathrm{cl}(A)} I(z), \end{aligned} \tag{8.14}$$

where $\mathrm{int}(A)$ and $\mathrm{cl}(A)$ denote the interior and topological closure, respectively, of the set A. In the above one dimensional setting, the LD principle (8.14) was derived for sets $A := [a, +\infty)$.

We have that if $\mu \in \mathrm{int}(A)$ and the moment generating function $M(t)$ is finite valued for all t in a neighborhood of $0 \in \mathbb{R}^d$, then $\inf_{z\in\mathbb{R}^d\setminus(\mathrm{int}\, A)} I(z)$ is positive. Moreover, if the sequence is iid, then

$$\limsup_{N\to\infty} N^{-1} \log[\mathbb{P}(Z_N \notin A)] < 0, \tag{8.15}$$

i.e., the probability $\mathbb{P}(Z_N \in A) = 1 - \mathbb{P}(Z_N \notin A)$ approaches one exponentially fast as N tends to infinity. For a thorough discussion of the Large Deviations theory, the interested reader is referred to Dembo and Zeitouni (1998), for example.

9 Bibliographic notes

The idea of using Monte Carlo sampling for solving stochastic optimization problems of the form (1.1) certainly is not new. There is a variety of sampling based optimization techniques, which were suggested in the literature. It will be beyond the scope of this chapter to give a comprehensive survey of these methods, we mention a few approaches related to the material of this chapter. One approach uses the Infinitesimal Perturbation Analysis (IPA) techniques to estimate the gradients of $f(\cdot)$, which consequently are employed in the Stochastic Approximation (SA) method. For a discussion of the IPA and SA methods we refer to Ho and Cao (1991), Glasserman (1991), Kushner and Clark (1978) and Nevelson and Hasminskii (1976), respectively. For an application of this approach to optimization of queueing systems see Chong and Ramadge (1993) and L'Ecuyer and Glynn (1994), for example. Closely related to this approach is the Stochastic Quasi-Gradient method (see Ermoliev, 1983).

Another class of methods uses sample average estimates of the values of the objective function, and may be its gradients (subgradients), in an "interior" fashion. Such methods are aimed at solving the true problem (1.1) by employing sampling estimates of $f(\cdot)$ and $\nabla f(\cdot)$ blended into a particular optimization algorithm. Typically, the sample is updated or a different sample is used each time function or gradient (subgradient) estimates are required at a current iteration point. In this respect we can mention, in particular, the statistical L-shaped method of Infanger (1994) and the stochastic decomposition method of Higle and Sen (1996b).

In this chapter we mainly discussed an "exterior" approach, in which a sample is generated outside of an optimization procedure and consequently the constructed SAA problem is solved by an appropriate deterministic optimization algorithm. There are several advantages in such approach. The method separates sampling procedures and optimization techniques. This makes it easy to implement and, in a sense, universal. From the optimization point of view, given a sample $\xi^1, \ldots, \xi^N$, the obtained optimization problem can be considered as a stochastic program with the associated scenarios $\xi^1, \ldots, \xi^N$, each taken with equal probability N^{-1}. Therefore, any optimization algorithm which is developed for a considered class of stochastic programs can be applied to the constructed SAA problem in a straightforward way. Also the method is ideally suited for a parallel implementation. From the theoretical point of view, as it was shown in the previous sections, there is available a quite well developed statistical inference of the SAA method. This,

in turn, gives a possibility of error estimation, validation analysis and hence stopping rules. Finally, various variance reduction techniques can be conveniently combined with the SAA method.

It is difficult to point out an exact origin of the SAA method. The idea is simple indeed and it was used by various authors under different names. Variants of this approach are known as the stochastic counterpart method (Rubinstein and Shapiro, 1990, 1993) and sample-path optimization (Plambeck et al., 1996; Robinson, 1996), for example. Also similar ideas were used in statistics for computing maximum likelihood estimators by Monte Carlo techniques based on Gibbs sampling (see, e.g., Geyer and Thompson, 1992 and references therein). Numerical experiments with the SAA approach, applied to linear and discrete (integer) stochastic programming problems, can be also found in more recent publications (Ahmed and Shapiro, 2002; Linderoth et al., 2002; Verweij et al., 2003).

Statistical theory of the SAA estimators is closely related to the statistical inference of the Maximum Likelihood (ML) method and M-estimators. Starting with a pioneering work of Wald (1949), consistency properties of the Maximum Likelihood and M-estimators were studied in numerous publications. Epi-convergence approach to studying consistency of statistical estimators was developed in Dupačová and Wets (1988). In the context of stochastic programming, consistency of SAA estimators was also investigated by tools of epi-convergence analysis in King and Wets (1991) and Robinson (1996). Uniform Laws of Large Numbers take their origin in the Glivenko–Cantelli theorem. For a further discussion of the uniform LLN we refer to van der Vaart and Wellner (1996), and for applications to stochastic programming to the recent paper by Pflug et al. (1998).

Asymptotic normality of M-estimators was proved, under quite weak differentiability assumptions, in Huber (1967). For a further discussion, and additional references, of the asymptotics of SAA optimal solutions see chapter "Stochastic Optimization and Statistical Influence" of this book. Theorems 10 and 11 are taken from Shapiro (1991).

It is possible to show that in smooth (differentiable) cases, and under some mild regularity conditions, the (unconstrained) SAA estimators $\hat{x}_N$ converge to the (unique) optimal solution x^* of the true problem at the same asymptotic rate as the corresponding stochastic approximation estimators based on the asymptotically optimal step sizes (Shapiro, 1996). As it is shown in Section 3, the situation is quite different in cases where the feasible set X is finite or the problem is polyhedral. Presentation of Sections 3.1 and 3.2 is based on Kleywegt et al. (2001), and of Section 3.3 on Shapiro and Homem-de-Mello (2000). The discussion of conditioning of stochastic programs is based on Shapiro et al. (2002). It is somewhat surprising that some problems with astronomically large number of scenarios are very well conditioned and can be solved exactly with a small sample size (see Linderoth et al., 2002 for such numerical examples). Exponential rates of convergence in stochastic programming were also studied in Kaniovski et al. (1995) and Dai et al. (2000).

The statistical bounds of Section 4.1 were suggested in Norkin et al. (1998), and developed further in Mak et al. (1999). The "common random numbers" estimator $\widehat{\mathrm{gap}}_{N,M}(\overline{x})$ of the optimality gap was introduced in Mak et al. (1999). Proposition 16 is due to Norkin et al. (1998) and Mak et al. (1999). The KKT statistical test, discussed in Section 4.2, was developed in Shapiro and Homem-de-Mello (1998), so that the material of that section is based on Shapiro and Homem-de-Mello (1998). See also Higle and Sen (1996a). Proposition 17 is taken from Shapiro and Homem-de-Mello (2000).

For a discussion of variance reduction techniques in Monte Carlo sampling we refer to Fishman (1999) and a survey paper by Avramidis and Wilson (1996), for example. In the context of stochastic programming, variance reduction techniques were discussed in Rubinstein and Shapiro (1993), Dantzig and Infanger (1991), Higle (1998) and Bailey et al. (1999), for example.

Section 6 is based on Shapiro (2002). It appears by now that statistical behavior of two-stage stochastic programming problems is quite well understood with theoretical developments supported by numerical experiments. It was demonstrated that huge two-stage linear programs, with astronomically large number of scenarios, can be solved efficiently with a proved accuracy. On the other hand, little is known about large scale multistage programs. In a sense the conclusions of Section 6 are mostly negative. At this moment we do not have a useful theory or numerical evidence which can guide us in solving large multistage problems. So further research in this direction is needed.

An extension of the SAA method to stochastic generalized equations is a natural one. Stochastic variational inequalities were discussed by Gürkan et al. (1999). Proposition 21 and Theorem 22 are similar to Theorems 1 and 2 in Gürkan et al. (1999).

References

Ahmed, S., A. Shapiro (2002). The sample average approximation method for stochastic programs with integer recourse. *E-print available at* http://www.optimization-online.org.

Avramidis, A.N., J.R. Wilson (1996). Integrated variance reduction strategies for simulation. *Operations Research* 44, 327–346.

Bailey, T.G., P. Jensen, D.P. Morton (1999). Response surface analysis of two-stage stochastic linear programming with recourse. *Naval Research Logistics* 46, 753–778.

Bonnans, J.F., A. Shapiro (2000). *Perturbation Analysis of Optimization Problems*, Springer-Verlag, New York, NY.

Chong, E.K.P., P.J. Ramadge (1993). Optimization of queues using an infinitesimal perturbation analysis-based stochastic algorithm with general update times. *SIAM Journal of Control and Optimization* 31, 698–732.

Dai, L., C.H. Chen, J.R. Birge (2000). Convergence properties of two-stage stochastic programming. *Journal of Optimization Theory and Applications* 106, 489–509.

Dantzig, G.B., G. Infanger (1992). Large-scale stochastic linear programs—importance sampling and Benders decomposition: *Computational and Applied Mathematics, I (Dublin, 1991)*, North-Holland, Amsterdam, pp. 111–120.

Dembo, A., O. Zeitouni (1998). *Large Deviations Techniques and Applications*, Springer-Verlag, New York, NY.

Dupačová, J., R.J.B. Wets (1988). Asymptotic behavior of statistical estimators and of optimal solutions of stochastic optimization problems. *The Annals of Statistics* 16, 1517–1549.

Ermoliev, Y. (1983). Stochastic quasi-gradient methods and their application to systems optimization. *Stochastics* 4, 1–37.

Fishman, G.S. (1999). *Monte Carlo, Concepts, Algorithms and Applications*, Springer-Verlag, New York.

Geyer, C.J., E.A. Thompson (1992). Constrained Monte Carlo maximum likelihood for dependent data, (with discussion). *Journal of the Royal Statistical Society, Series B* 54, 657–699.

Glasserman, P. (1991). *Gradient Estimation via Perturbation Analysis*, Kluwer Academic Publishers, Norwell, MA.

Gürkan, G., A.Y. Özge, S.M. Robinson (1999). Sample-path solution of stochastic variational inequalities. *Mathematical Programming* 21, 313–333.

Higle, J.L. (1998). Variance reduction and objective function evaluation in stochatsic linear programs. *INFORMS Journal on Computing* 10(2), 236–247.

Higle, J.L., S. Sen (1996). Duality and statistical tests of optimality for two stage stochastic programs. *Mathematical Programming (Series B)* 75(2), 257–275.

Higle, J.L., S. Sen (1996). *Stochastic Decomposition: A Statistical Method for Large Scale Stochastic Linear Programming*, Kluwer Academic Publishers, Dordrecht.

Ho, Y.C., X.R. Cao (1991). *Perturbation Analysis of Discrete Event Dynamic Systems*, Kluwer Academic Publishers, Norwell, MA.

Huber, P.J. (1967). *The Behavior of Maximum Likelihood Estimates under Nonstandard Conditions*, Vol. 1, Univ. California Press pp. 221–233.

Infanger, G. (1994). *Planning Under Uncertainty: Solving Large Scale Stochastic Linear Programs*, Boyd and Fraser Publishing Co, Danvers, MA.

Kaniovski, Y.M., A. King, R.J.-B. Wets (1995). Probabilistic bounds for the solutions of stochastic programming problems. *Annals of Operations Research* 56, 189–208.

King, A.J., R.J.-B. Wets (1991). Epi-consistency of convex stochastic programs. *Stochastics Stochastics Rep*. 34(1–2), 83–92.

Kleywegt, A.J., A. Shapiro, T. Homem-De-Mello (2001). The sample average approximation method for stochastic discrete optimization. *SIAM Journal of Optimization* 12, 479–502.

Kushner, H.J., D.S. Clark (1978). *Stochastic Approximation Methods for Constrained and Unconstrained Systems*, Springer-Verlag, Berlin, Germany.

L'Ecuyer, P., P.W. Glynn (1994). Stochastic optimization by simulation: Convergence proofs for the GI/G/1 queue in steady-state. *Management Science* 11, 1562–1578.

Linderoth, J., A. Shapiro, S. Wright (2002). The empirical behavior of sampling methods for stochastic programming. Optimization Technical Report 02-01, Computer Sciences Department, University of Wisconsin-Madison.

Mak, W.K., D.P. Morton, R.K. Wood (1999). Monte Carlo bounding techniques for determining solution quality in stochastic programs. *Operations Research Letters* 24, 47–56.

Nevelson, M.B., R.Z. Hasminskii (1976). *Stochastic Approximation and Recursive Estimation*. American Mathematical Society Translations of Mathematical Monographs, Vol. 47, Rhode Island.

Norkin, V.I., G.Ch. Pflug, A. Ruszczyński (1998). A branch and bound method for stochastic global optimization. *Mathematical Programming* 83, 425–450.

Pflug, G.Ch., A. Ruszczyński, R. Schultz (1998). On the Glivenko Cantelli problem in stochastic programming: Linear recourse. *Mathematics of Operations Research* 23, 204–220.

Plambeck, E.L., B.R. Fu, S.M. Robinson, R. Suri (1996). Sample-path optimization of convex stochastic performance functions. *Mathematical Programming, Series B* 75, 137–176.

Robinson, S.M. (1980). Strongly regular generalized equations. *Mathematics of Operations Research* 5, 43–62.

Robinson, S.M. (1982). Generalized equations and their solutions, part ii: Applications to nonlinear programming. *Mathematical Programming Study* 19, 200–221.

Robinson, S.M. (1992). Normal maps induced by linear transformations. *Mathematics of Operations Research* 17, 691–714.

Robinson, S.M. (1996). Analysis of sample-path optimization. *Mathematics of Operations Research* 21, 513–528.

Rockafellar, R.T., R.J.-B. Wets (1998). *Variational Analysis*, Springer, Berlin.

Rubinstein, R.Y., A. Shapiro (1990). Optimization of static simulation models by the score function method. *Mathematics and Computers in Simulation* 32, 373–392.

Rubinstein, R.Y., A. Shapiro (1993). *Discrete Event Systems: Sensitivity Analysis and Stochastic Optimization by the Score Function Method*, John Wiley & Sons, Chichester, England.

Shapiro, A. (1991). Asymptotic analysis of stochastic programs. *Annals of Operations Research* 30, 169–186.

Shapiro, A. (1996). Simulation-based optimization: Convergence analysis and statistical inference. *Stochastic Models* 12, 425–454.

Shapiro, A. (2002). Statistical inference of multistage stochastic programming problems. *E-print available at* http://www.optimization-online.org.

Shapiro, A., T. Homem-de-Mello (1998). A simulation-based approach to two-stage stochastic programming with recourse. *Mathematical Programming* 81, 301–325.

Shapiro, A., T. Homem-de-Mello (2000). On rate of convergence of Monte Carlo approximations of stochastic programs. *SIAM Journal on Optimization* 11, 70–86.

Shapiro, A., T. Homem de Mello, J.C. Kim (2002). Conditioning of convex piecewise linear stochastic programs. *Mathematical Programming*, 94, 1–19.

van der Vaart, A.W., A. Wellner (1996). *Weak Convergence and Empirical Processes*, Springer-Verlag, New York.

Verweij, B., S. Ahmed, A.J. Kleywegt, G. Nemhauser, A. Shapiro (2003). The sample average approximation method applied to stochastic routing problems: A computational study. *Computational Optimization and Applications* 24, 289–333.

Wald, A. (1949). Note on the consistency of the maximum likelihood estimates. *Annals Mathematical Statistics* 20, 595–601.

A. Ruszczyński and A. Shapiro, Eds., *Handbooks in OR & MS, Vol. 10*

Chapter 7

Stochastic Optimization and Statistical Inference

G.Ch. Pflug

Department of Statistics and Decision Support Systems, University of Vienna, Universitätsstrasse 5, A-1090 Vienna, Austria

Abstract

If the distribution of the random parameters of a stochastic program is unknown, the empirical distribution based on a sample may be used as a proxy. This empirical approximation is related to the "true" stochastic program in the same way as a statistical estimate is related to the true parameter value. Properties of statistical estimators, like consistency, asymptotical distributions and the construction of confidence regions are reviewed in the realm of stochastic optimization. The entropic size of a stochastic program determines the quality of the approximation. In case that random constraints are present, the notion of epiconvergence replaces in a natural way the notion of uniform convergence of functions. The asymptotic structures are described by the asymptotic stochastic program associated to the sequence of empirical programs.

Key words: Empirical program, statistical estimates, asymptotic statistics, risk functionals, entropic size, epiconvergence, asymptotic stochastic programs.

1 Uncertain and ambiguous optimization problems

In deterministic optimization, a decision x must be found, which minimizes a known cost function $f(x)$ among all possible candidates x lying in the *feasible set* $\mathcal{X} \subseteq \mathbb{R}^d$, a closed subset of the euclidean d-dimensional space

$$\operatorname*{Min}_{x \in \mathcal{X}} f(x).$$

In stochastic optimization, the cost function is not exactly known at the time when the decision is made. Only a stochastic model $F(x,\xi)$ for the costs is known, where ξ is some random vector defined on a probability space $(\Omega_0, \mathcal{A}_0, \mathbb{P})$, taking its values in $\mathbb{R}^m$. The crucial point is that although the particular value of ξ is unknown, its distribution is completely known. We refer to this situation as the *uncertainty problem*. Since the decision maker is not clairvoyant and does not know the actual value of ξ, he/she cannot minimize the cost function for each value of ξ separately, but has to minimize some real functional $\mathbb{F}$, which summarizes the random costs $F(x,\xi)$ in an appropriate manner. Think of the expectation functional $\mathbb{E}$ as the summarizing functional for the moment. The uncertainty problem reads

$$\operatorname*{Min}_{x\in\mathcal{X}}\{f(x) = \mathbb{E}[F(x,\xi)]\} \tag{1.1}$$

or—using a more general summarizing functional—

$$\operatorname*{Min}_{x\in\mathcal{X}}\{f(x) = \mathbb{F}[F(x,\xi)]\}. \tag{1.2}$$

Any functional $\mathbb{F}: \mathcal{G} \overset{\mathbb{F}}{\to} \mathbb{R}$, which maps $\mathcal{G}$, a set of distribution functions on $\mathbb{R}$, onto the real line may be used as summarizing functional. We call such functionals *risk functionals*. Examples for risk functionals are the expectation, the variance, the median, quantiles, etc. A collection of widely used risk functionals can be found in Section 4.

For simplicity, we assume for the moment that the set of constraints $\mathcal{X}$ does not depend on the random vector ξ. In Sections 7 and 8 of this chapter, we will indicate how to deal with additional "random constraints", i.e., constraints involving random variables such as

$$\mathcal{X} \cap \{x: \mathbb{E}[F_1(x,\xi)] \le b_1, \ldots, \mathbb{E}[F_J(x,\xi)] \le b_J\}.$$

Example 1. **A one-period portfolio optimization problem.** An investor has a budget B to spend in m investment categories, like bonds, stocks or other contracts. If one currency unit is invested in the ith category, then the value after the holding period is ξ_i. Here $\xi = (\xi_1, \ldots, \xi_m)$ is a vector of random price changes. If x_i denotes the amount invested in category i, then $F(x,\xi) = \sum_{i=1}^m x_i\xi_i$ is the value of the portfolio at the end of the holding period. The decision maker wants to maximize the risk-adjusted expected return. He chooses for instance the expectation minus 10% of the mean absolute deviation as risk functional, i.e.,

$$\mathbb{F}[F(x,\xi)] = \mathbb{E}[F(x,\xi)] - \frac{1}{10}\cdot\mathbb{E}|F(x,\xi) - \mathbb{E}[F(x,\xi)]|.$$

This value should be maximized (or its negative value should be minimized) under the budget and the nonnegativity constraints

$$\text{Min}\left\{-\mathbb{F}[F(x,\xi)]\colon \sum_{i=1}^{m} x_i = B,\ x_i \geq 0\right\}.$$

This is a typical stochastic optimization problem of the form (1.2).

The assumption that the distribution P, i.e., $P(A) = \mathbb{P}\{\xi \in A\}$ of the random part ξ of the cost function is exactly known is rarely fulfilled. Typically, the probability measure P is unknown (this is called the *ambiguity problem*) and only some information about it is available.

In the ambiguity situation, the stochastic optimization problem gets an additional difficulty: the unknown probability P has to be guessed through the available information and this gives an additional source of error.

Typically, the available information is contained in a sample of historic data. In other situations, moments or some other characteristics of the unknown probability measure are known (see e.g., Gaivoronski et al., 1985).

In this chapter, we consider the case that a sample $\xi_1, \ldots, \xi_N$ of independent, identically distributed random variables with distribution P is available. This sample may be used in two ways:

- *The parametric approach:* A parametric family (P_θ) of probability measures is chosen as the true model and the parameter θ is estimated by an estimate $\hat{\theta}$ on the basis of the sample. The probability measure $P_{\hat{\theta}}$ is then taken as the distribution of ξ in (1.2).
- *The nonparametric approach:* No assumption about a parametric family is made. The empirical measure $\hat{P}_N$ (see (2.4) below for its formal definition), which puts mass $1/N$ on each observation in the sample, serves as a proxy of the unknown P.

For illustration, consider the above portfolio optimization example: no financial analyst knows correct distribution of the price change vectors. Typically, either the historical data distribution is used directly (the nonparametric approach, called historic simulation by finance managers) or some theoretical models are fitted using prior knowledge and some data (the parametric approach).

In this chapter, we treat only the nonparametric approach and study the quality of this approximation. The quality of the parametric approach depends strongly on the right choice of the parametric model.

The fact that the true probability measure P is unknown and replaced by some data-based estimate apparently makes the problem harder, since one has to cope with the statistical estimation error in addition. However, as a positive effect of this approximation, one typically gets an easier problem to solve,

since the empirical distribution generated by the data sits on finitely many points. To deal with such discrete distributions is much easier than to deal with arbitrary (continuous) distributions. Just think of the fact, how much easier it is to numerically calculate a finite sum than a multidimensional integral.

The advantage in the numerical treatment of the finite-sample approximation sometimes outweighs the disadvantage of the additional approximation error. As a consequence, an artificially generated sample may be used to replace a nonambiguous model. The results for sampling approximation carry of course over to the case of artificially generated samples. Notice, however, that in the case that the sample is artificially generated, one has the freedom to choose the sample in various ways. Instead of mimicking randomness by pseudo-random numbers (Monte-Carlo method, see chapter *Monte Carlo Sampling Methods* by Shapiro in this volume), one could also use a nonrandom low-discrepancy quasi-random sequence (Quasi-Monte Carlo method, see Niederreiter, 1992) or some other, well chosen discrete approximation. The Monte-Carlo method allows to generate an approximation in a very simple way, however, it produces only random approximations of the unknown distribution P, which is a disadvantage since the quality of the approximation can only be stated in statistical terms.

In this chapter, we will only treat the approximation by sampling from a sequence of either observed or artificially generated random variables.

2 The empirical problem

Denote by G_x the distribution function of the costs, parameterized by the decision x

$$G_x(u) = \mathbb{P}\{F(x, \xi) \le u\} = \int \mathbf{1}_{\{F(x,\omega)\le u\}} \, \mathrm{d}P(\omega). \tag{2.3}$$

Here $\mathbf{1}_B$ is the indicator function of the set B

$$\mathbf{1}_B(u) = \begin{cases} 1 & u \in B \\ 0 & u \notin B \end{cases}$$

Let $\xi_1, \ldots, \xi_N$ be a random sample, independently and identically distributed with the same distribution P. Denote by $\hat{P}_N$ the *empirical distribution* generated by this sample, i.e.,

$$\hat{P}_N(A) = \frac{1}{N} \sum_{n=1}^{N} \mathbf{1}_{\{\xi_n \in A\}}. \tag{2.4}$$

Notice that $\hat{P}_N$ depends on the random sample and is therefore a *random probability measure*. For every fixed set A, $\hat{P}_N(A)$ converges almost surely to $P(A)$ due the Strong Law of Large Numbers. If P is a probability measure on a separable metric space, then $\hat{P}_N$ converges as a measure almost surely in the weak sense to P as N tends to infinity. This basic fact of statistics was proved by Varadarajan (1958).

Based on the empirical distribution $\hat{P}_N$ of the random part of the cost function, we may form the empirical cost distribution $\hat{G}_{x,N}$, parameterized by the decision x

$$\hat{G}_{x,N}(u) = \frac{1}{N}\sum_{n=1}^{N} \mathbf{1}_{\{F(x,\xi_n)\le u\}} = \int \mathbf{1}_{\{F(x,\omega)\le u\}}\, \mathrm{d}\hat{P}_N(\omega), \tag{2.5}$$

which is the *empirical counterpart* of (2.3).

The risk functional $\mathbb{F}$ maps distributions to real numbers. By a slight abuse of notation, but with no danger of confusion, we use the notation $\mathbb{F}(Z)$ for a random variable Z and $\mathbb{F}(G_Z)$ for its distribution G_Z in parallel. The main problem (1.2) in new notation reads now

$$\operatorname*{Min}_{x\in\mathcal{X}}\{f(x) = \mathbb{F}[G_x]\}. \tag{2.6}$$

Its empirical counterpart, which is based on the sample $\xi_1,\ldots,\xi_N$ is

$$\operatorname*{Min}_{x\in\mathcal{X}}\{\hat{f}_N(x) = \mathbb{F}[\hat{G}_{x,N}]\}. \tag{2.7}$$

The relation between f and $\hat{f}_N$ is like in statistical estimation theory: $\hat{f}_N(\cdot)$ is an estimate of the unknown objective function $f(\cdot)$ and $\arg\min_{x\in\mathcal{X}}\hat{f}_N(x)$ is an estimate of the unknown solution $\arg\min_{x\in\mathcal{X}} f(x)$. Recall here the notion of the arg min set

$$\arg\min_{x\in\mathcal{X}} f(x) := \left\{z\in\mathcal{X}: f(z) = \min_{x\in\mathcal{X}} f(x)\right\}.$$

The arg min set is empty if the minimum is not attained.

In statistics, the function f is called the criterion function. Estimates of the arg min type are the most important group of estimates in statistics: they include maximum-likelihood estimates, M-estimates, minimum-contrast and minimum-distance estimates. In the usual setup of statistical estimation theory, the sample $\xi_1,\ldots,\xi_N$ stems from a distribution P, which is a member of some family $\mathcal{P}$. Here $\mathcal{P}$ may be a parametric family $\mathcal{P} = (P_\theta)_{\theta\in\Theta}$ or a nonparametric family. In either case, suppose that a parametric function

Table 7.1
The relation between Statistical Estimation and Stochastic Optimization

Statistics	Stochastic optimization
The true probability measure P^* is a member of a family $\mathcal{P}$. P^* is unknown, but a sample $\xi_1, \ldots, \xi_N$ is available	
A parametric function $P \mapsto \theta(P)$, $P \in \mathcal{P}$ is given. The statistician wants to estimate $\theta^* = \theta(P^*)$, where P^* is the true probability measure	The cost function F and a functional $\mathbb{F}$, both representing some economic decision problem are given
The statistician chooses a criterion function F and a functional $\mathbb{F}$ in such a way that $\theta(P) = \arg\min_{x \in X} \mathbb{F}[F(x, \xi)]$, where ξ is distributed according to P	
The estimate is in both cases $\hat{\theta}_N \in \arg\min_x \{\hat{f}_N = \mathbb{F}[F(x, \xi)]\}$, where ξ is distributed according to the empirical measure $\hat{P}_N$ based on the sample $\xi_1, \ldots, \xi_N$	

$P \in \mathcal{P} \to \theta(P)$ is given. This unknown parametric function is estimated by an estimator $\hat{\theta}_N = \hat{\theta}_N(\xi_1, \ldots, \xi_N)$ based on the observed data.

Statistical convergence results are formulated for the whole family $\mathcal{P}$, i.e., a typical result reads: *if the data stem from* P^*, *then the estimate* $\hat{\theta}_N$ *converges to the true* $\theta^* = \theta(P^*)$, *for all* $P^* \in \mathcal{P}$ (Table 7.1).

Example 2. **Quantile estimation and the newsboy problem.** Suppose that we are interested in estimating the α-quantile θ^* from a random distribution P with distribution function G and continuous density g. Suppose that $g(x) > 0$, if $G(x) = \alpha$. Then the α-quantile $\theta^* = G^{-1}(\alpha)$ is uniquely determined. The statistician estimates θ^* by the sample quantile.

The newsboy problem, introduced in the chapter *Stochastic Programming Models* by Ruszczynski and Shapiro in this volume is a stochastic optimization version of the quantile estimation problem. Recall that the profit function, i.e., the negative cost function is

$$F(x, \xi) = (s - c)x + (r - s)[x - \xi]^+,$$

where x is the decision, ξ is the demand variable, s is the sell price, c is the buy price and r is the return price ($0 < r < c < s$). Let $G(u) = \mathbb{P}\{\xi \le u\}$ and $G^{(1)}(u) = \mathbb{E}([u - \xi]_+) = \int_{-\infty}^{u} (u - v)\, \mathrm{d}G(v) = \int_{-\infty}^{u} G(v)\, \mathrm{d}v$. The objective function is

$$f(x) = \mathbb{E}[F(x, \xi)] = (s - c)x + (r - s)G^{(1)}(x).$$

By simple calculus one sees that every optimal solution $\theta^* \in \arg\min(-f)$ satisfies the necessary condition $G(\theta^*) = (s-c)/(s-r)$, i.e., is an α-quantile of G,

where $\alpha = (s-c)/(s-r)$. Moreover, the arg min is unique. Thus, the newsboy problem of stochastic optimization turns out to be equivalent to the quantile estimation problem in statistics. All results about properties of the sample quantile estimate carry over to the newsboy stochastic optimization problem.

Generally spoken, the same type of results about convergence, asymptotics and confidence regions hold for the estimation of statistical parameters of the arg min type and for the solution of empirical stochastic programs. Notice, however, that the interpretation is slightly different: in statistics the emphasis is on the identification of a parameter or the distribution itself. In stochastic optimization the goal is to make optimal decisions, the unknown probability measure is not in the center of interest. In statistical estimation, the loss function plays the role of the cost function. It is typically up to the statistician to choose an appropriate loss function. In contrast, the cost function is given in a stochastic optimization problem.

The estimate of the objective function (the criterion function) lies in some metric space of functions, for instance the space of continuous functions endowed with the sup-norm, the space of lower semicontinuous functions endowed with epigraphical distance (see Section 7) or the space of cad-lag functions endowed with the Skorohod-distance. The arg mins need not be singletons in general, they may be sets, and the arg min estimates lie in some metric space of sets (e.g., the metric space of closed sets endowed with the Hausdorff distance or some variant of it).

We denote by $\hat{\theta}_N$ an estimate of an unknown value θ^*, based on a sample $\xi_1, \ldots, \xi_N$. Assume that $\hat{\theta}_N$ and θ^* lie in a metric space (Θ, d). This notation is valid for the estimation of criterion functions (in which case Θ is a space of functions) or arg min sets (in which case Θ is a space of sets) or singleton arg mins (in which case Θ is the $\mathbb{R}^d$).

There are some fundamental concepts of describing the quality of statistical estimates: an estimate $\hat{\theta}_N$ is good if it is close to the true value θ^*, i.e., if the random variable $d(\hat{\theta}_N, \theta^*)$ is small. Since there are several ways of expressing the fact that a random variable is "small", there are several quality notions for estimates. We describe some of them in the next section.

3 Properties of statistical estimates

Let $\hat{\theta}_N = \hat{\theta}_N(\xi_1, \ldots, \xi_N)$ be a sequence of estimates for an unknown true value θ^* based on the sample $\xi_1, \ldots, \xi_N$. Assume that the estimates and the true value lie in a metric space (Θ, d).

- **Consistency.**
  A sequence $(\hat{\theta}_N)$ of estimates is *weakly consistent*, if

$$d(\hat{\theta}_N, \theta^*) \to 0 \tag{3.8}$$

in probability, i.e., for every $\varepsilon > 0$

$$\mathbb{P}\{d(\hat{\theta}_N, \theta^*) > \varepsilon\} \to 0 \quad \text{as } N \to \infty.$$

The sequence $(\hat{\theta}_N)$ is *strongly consistent*, if

$$d(\hat{\theta}_N, \theta^*) \to 0 \quad \text{almost surely, as } N \to \infty, \tag{3.9}$$

i.e.,

$$\mathbb{P}\{d(\hat{\theta}_N, \theta^*) \to 0\} = 1.$$

Strong consistency implies weak consistency. Consistency is a minimal requirement to qualify the estimate as an acceptable approximation of the unknown true value.

- **Speed of convergence.**
Let $\beta(N)$ be an appropriate blow-up function, i.e., a function which satisfies $\beta(N) \to \infty$ as $N \to \infty$.

The sequence $(\hat{\theta}_N)$ is *weakly* $\beta(N)$*-consistent*, if $\beta(N)\, d(\hat{\theta}_N, \theta^*)$ stays bounded in probability as N tends to infinity, i.e., for every $\varepsilon > 0$, there is a K_ε such that

$$\mathbb{P}\{\beta(N)\, d(\hat{\theta}_N, \theta^*) > K_\varepsilon\} \le \varepsilon \tag{3.10}$$

for sufficiently large N (by possibly enlarging K_ε, the condition (3.10) may be assumed to hold for all N).

The sequence $(\hat{\theta}_N)$ is *strongly* $\beta(N)$*-consistent*, if $\beta(N)\, d(\hat{\theta}_N, \theta^*)$ is almost surely bounded, i.e.,

$$\mathbb{P}\left\{\sup_N \beta(N)\, d(\hat{\theta}_N, \theta^*) < \infty\right\} = 1. \tag{3.11}$$

- **Asymptotic distribution.**
If the parameter space is a linear normed space $(\Theta, \|\cdot\|)$, then one may consider the difference between estimate and true value rescaled with some blow-up function $\beta(N) \to \infty$, i.e.,

$$\beta(N)[\hat{\theta}_N - \theta^*]. \tag{3.12}$$

The sequence of estimates $(\hat{\theta}_N)$ has an asymptotic distribution, if the expression in (3.12) converges in distribution to a nondegenerate limit. In regular cases the appropriate blow-up function is $\beta(N) = N^{1/2}$ (see Section 9).

The knowledge of the asymptotic distribution gives a precise statement about the situation for large sample sizes, but may give only little information about the quality of the estimate for a realistic sample size N: suppose one knows that $\beta(N)[\hat{\theta}_N - \theta^*]$ converges in distribution to some nondegenerated random variable V. Then for every K, which is a continuity point of $\|V\|$

$$\mathbb{P}\{\beta(N)\|\hat{\theta}_N - \theta^*\| \geq K\} \to \mathbb{P}\{\|V\| \geq K\} \tag{3.13}$$

as N tends to infinity. A confidence set $\hat{\Theta}_N^*$ may be constructed as

$$\hat{\Theta}_N^* = \{\theta\colon d(\hat{\theta}_N, \theta) \leq K/\beta(N)\},$$

where $d(\theta_1, \theta_2) = \|\theta_1 - \theta_2\|$ and K is chosen such that the right hand side in (3.13) is not larger than ε_0. Then $\hat{\Theta}_N^*$ is an *asymptotic confidence set*, i.e., it satisfies

$$\limsup_N \mathbb{P}\{\theta^* \notin \hat{\Theta}_N^*\} \leq \varepsilon_0.$$

- **Universal bounds.**
For getting good small sample confidence bounds, it is not necessary that the sequence $\beta(N)[\hat{\theta}_N - \theta^*]$ possesses an asymptotic distribution. It is only necessary that $\beta(N)d(\hat{\theta}_N, \theta^*)$ is *bounded in probability*. We say that $\beta(N)d(\hat{\theta}_N, \theta^*)$ is *bounded in probability with explicit tail behavior*, if we can find an explicit function $K \mapsto \varepsilon(K)$, with the property that $\varepsilon(K) \to 0$ for $K \to \infty$ such that for all K

$$\sup_N \mathbb{P}\{\beta(N)d(\hat{\theta}_N, \theta^*) \geq K\} \leq \varepsilon(K).$$

Explicit tail functions allow the construction of universal confidence sets. Let

$$\hat{\Theta}_N^* = \{\theta\colon d(\hat{\theta}_N, \theta) \leq K/\beta(N)\}.$$

This is a universal confidence set: choosing first the desired level $1-\varepsilon_0$ and then K such that $\varepsilon(K) \leq \varepsilon_0$ one gets that

$$\sup_N \mathbb{P}\{\theta^* \notin \hat{\Theta}_N^*\} \leq \varepsilon(K) \leq \varepsilon_0. \tag{3.15}$$

Based on the size of $\hat{\Theta}_N^*$ the decision maker may decide to stop or to increase the sample size by continuing sampling and data collection.

It is one purpose of this chapter to discuss universal bounds for empirical stochastic programs. The basic notion is rigorously introduced in the following definition.

Definition 2. Let $\varepsilon(K)$ be a nonnegative function on $(0,\infty)$ with the property that $\varepsilon(K)\to 0$ as $K\to\infty$.

(i) A sequence of random variables (Z_N) is *bounded in probability with tail function* $\varepsilon(K)$, if for all $K>0$

$$\sup_N \mathbb{P}\{|Z_N|\geq K\}\leq \varepsilon(K).$$

(ii) The sequence of random variables (Z_N) is *bounded in probability with normal tails*, if it fulfills (i) and the function $\varepsilon(K)$ satisfies

$$\varepsilon(K)\leq C_1\exp(-C_2K^2) \tag{3.16}$$

for all $K>0$ and some constants $C_1>0$, $C_2>0$.

The name in (ii) comes from the following property of the standard normal distribution function Φ: the two sided tail probability $\varepsilon(K)=\Phi(-K)+1-\Phi(K)$ satisfies (3.16) with $C_1=\sqrt{2/\pi}$ and $C_2=1/2$ (Mill's ratio).

Recall that the main goal is to judge the quality of approximation of a stochastic optimization problem by the pertaining empirical problem and to find good error bounds. To this end, we discuss three different, but related questions: (Q1) the quality of the approximation of the objective function as a whole, (Q2) the quality of the approximation of the optimal value and (Q3) the quality of the approximation of the solution (or the solution set if the solution is not unique).

3.1 The three basic questions about the approximation quality of a stochastic program

Consider a stochastic optimization problem of type (2.6)

$$\operatorname*{Min}_{x\in\mathcal{X}}\{f(x)=\mathbb{F}[G_x]\} \tag{3.17}$$

and its empirical counterpart (see (2.7))

$$\operatorname*{Min}_{x\in\mathcal{X}}\{\hat{f}_N(x)=\mathbb{F}[\hat{G}_{x,N}]\}. \tag{3.18}$$

The basic questions about the approximation quality of the true problem by the empirical problem are:

(Q1) **Approximation of the objective function:**
Under which conditions is

$$\beta(N) \sup_{x \in \mathcal{X}} |\hat{f}_N(x) - f(x)| \tag{3.19}$$

bounded in probability for some blow-up function $\beta(N)$? Does one get the explicit tail behavior? Does the normal tail behavior hold?

(Q2) **Quality of the solution:**
Under which conditions is

$$\beta(N)\left[\sup\left\{f(y): y \in \arg\min_{x \in \mathcal{X}} \hat{f}_N(x)\right\} - \min_{x \in \mathcal{X}} f(x)\right] \tag{3.20}$$

bounded in probability for some blow-up function $\beta(N)$? Does one get the explicit tail behavior? Does the normal tail behavior hold?

(Q3) **Approximation of the arg min:**
Suppose for the moment that all arg mins are singletons. Then the question is: under which conditions is

$$\beta(N)\|\arg\min_{x \in \mathcal{X}} \hat{f}_N(x) - \arg\min_{x \in \mathcal{X}} f(x)\| \tag{3.21}$$

bounded in probability for some blow-up function $\beta(N)$? Does one get the explicit tail behavior? Does the normal tail behavior hold?

The assumption that all arg mins are singletons is restrictive. Often arg mins are (closed) sets. We introduce therefore the Hausdorff distance between sets: the distance between a point $x \in \mathbb{R}^d$ and a set $B \subseteq \mathbb{R}^d$ is defined as

$$\mathbb{d}(x, B) = \inf\{\|x - y\| : y \in B\}.$$

The *one-sided Hausdorff distance* between sets A and B is

$$\mathbb{d}(A \parallel B) = \sup\{\mathbb{d}(x, B): x \in A\}. \tag{3.22}$$

We set $\mathbb{d}(B \parallel \emptyset) = \infty$ and $\mathbb{d}(\emptyset, B) = 0$ for the empty set $\emptyset$. $\mathbb{d}(A \| B)$ is asymmetric as is illustrated in Fig. 1. The *Hausdorff distance* between A and B is

$$\mathbb{d}(A, B) = \max\{\mathbb{d}(A \parallel B), \mathbb{d}(B \parallel A)\}. \tag{3.23}$$

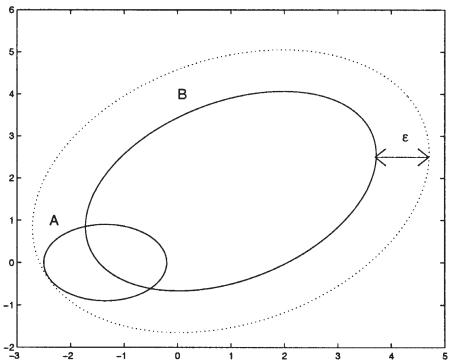

Fig. 1. An example, where $d(A \,||\, B) = \varepsilon$ but $\mathrm{d}(B \,||\, A)$ is large.

For closed sets A and B, $\mathbb{d}(A \,\|\, B) = 0$ if and only if $A \subseteq B$ and $\mathbb{d}(A, B) = 0$ if and only if $A = B$.

In the set-valued notation, the question (Q3) has to be formulated as (Q3a) or (Q3b).

(Q3a) **Asymptotic dominance of the arg min set**
Under which conditions is

$$\beta(N)\, \mathbb{d}\left(\arg\min_{x \in \mathcal{X}} \hat{f}_N(x) \,\|\, \arg\min_{x \in \mathcal{X}} f(x) \right) \tag{3.24}$$

bounded in probability for some blow-up function $\beta(N)$. Does one get the explicit tail behavior? Does the normal tail behavior hold?

Introducing the ε-fattening of a set B as

$$[B]^{\varepsilon} = \{x \colon \mathbb{d}(x, B) \le \varepsilon\},$$

the property (3.24) can be reformulated as follows: the relation

$$\arg\min_{x \in \mathcal{X}} \hat{f}_N(x) \subseteq \left[\arg\min_{x \in \mathcal{X}} f(x) \right]^{K/\beta(N)}$$

holds with arbitrarily high probability, if K and N are large enough.

(Q3b) **Asymptotic convergence of the arg min set**
Under which conditions is

$$\beta(N)\, \mathbb{d}\left(\arg\min_{x \in \mathcal{X}} \hat{f}_N(x), \arg\min_{x \in \mathcal{X}} f(x) \right) \tag{3.25}$$

bounded in probability for some blow-up function $\beta(N)$. Does one get the explicit tail behavior? Does the normal tail behavior hold?

The property (3.25) can be reformulated that both relations

$$\arg\min_{x\in\mathcal{X}} f(x) \subseteq \left[\arg\min_{x\in\mathcal{X}} \hat{f}_N(x)\right]^{K/\beta(N)}$$
$$\arg\min_{x\in\mathcal{X}} \hat{f}_N(x) \subseteq \left[\arg\min_{x\in\mathcal{X}} f(x)\right]^{K/\beta(N)}$$

hold with arbitrarily high probability, if K and N are large enough.

If there is a positive answer to (Q3a) or (Q3b), then a confidence region for the arg min may be constructed in a canonical way. Let the confidence region $\hat{C}_N$ be

$$\hat{C}_N = \left\{x: \mathbb{d}\left(x, \arg\min_{x\in\mathcal{X}} \hat{f}_N(x)\right) \leq K/\beta(N)\right\}$$

where K has to be chosen in such a way that $\varepsilon(K) \leq \varepsilon_0$, with $1-\varepsilon_0$ being the desired level of confidence. If (3.24) has been established with tail function $\varepsilon(K)$, then $\hat{C}_N$ is a *weak universal confidence set*, i.e.,

$$\sup_N \mathbb{P}\left\{\hat{C}_N \cap \arg\min_{x\in\mathcal{X}} f(x) = \emptyset\right\} \leq \varepsilon_0. \tag{3.26}$$

If even (3.25) holds, then $\hat{C}_N$ is a *strong universal confidence set*, i.e.,

$$\sup_N \mathbb{P}\left\{\arg\min_{x\in\mathcal{X}} f(x) \not\subseteq \hat{C}_N\right\} \leq \varepsilon_0.$$

The term *universal* refers to the fact that the inequality is valid for all N, in contrast to an *asymptotic* confidence set, which is valid only for large N. Notice that (3.26) can be established, if (3.24) holds and $\arg\min_{x\in\mathcal{X}} f(x)$ is a singleton. Unfortunately, no other method for proving (3.26) is known at the time being.

3.2 *Relations between the different types of approximation errors*

The basic question (Q1) is the fundamental one. Questions (Q2) and (Q3) are related to (Q1) as is shown in this section.

Proposition 3. *If the answer to* (Q1) *is positive, then the answer to* (Q2) *is also positive. More precisely, if*

$$\beta(N) \sup_{x\in\mathcal{X}} |\hat{f}_N(x) - f(x)|$$

is bounded in probability with tail function $\varepsilon(K)$, *and* $\arg\min_{x\in\mathcal{X}} \hat{f}_N(x)$ *is nonempty, then*

$$\beta(N)\left[\sup\left\{f(y)\colon y\in \arg\min_{x\in\mathcal{X}} \hat{f}_N(x)\right\} - \min_{x\in\mathcal{X}} f(x)\right]$$

is also bounded in probability with tail function $\varepsilon(K/2)$.

Proof. The proposition is established, if we show that for two functions f, g

$$\sup\left\{f(y)\colon y\in \arg\min_{x\in\mathcal{X}} g(x)\right\} - \min_{x\in\mathcal{X}} f(x) \le 2\sup_{u\in\mathcal{X}} |g(u)-f(u)| \tag{3.27}$$

and apply the result for $g(x)=\hat{f}_N(x)$. Suppose first that the minima of f and g are attained. Let $y\in\arg\min_{u\in\mathcal{X}} g(u)$ and $x\in\arg\min_{u\in\mathcal{X}} f(u)$. Then $f(y)-f(x)\le g(y)-g(x)+|g(y)-f(y)|+|g(x)-f(x)|\le 2\sup_{u\in\mathcal{X}}|g(u)-f(u)|$, since $g(y)-g(x)\le 0$. If the minima of f and g are not attained, they may be approximated with an error at most ε. The same argument as before together with the fact that ε is arbitrary leads to (3.27). □

The relation between (Q3) and (Q1) depends on the validity of some growth condition.

Definition 4. **Growth functions.** The function f possesses the growth function $\Psi(d)$ on $\mathcal{X}$, if Ψ is a strictly increasing function on $(0,\infty)$ satisfying $\Psi(d)\to 0$ as $d\to 0$ and

$$f(x)\ge \inf_{y\in\mathcal{X}} f(y) + \Psi\left(\mathbb{d}\left(x, \arg\min_{z\in\mathcal{X}} f(z)\right)\right). \tag{3.28}$$

Notice that this definition requires that $\arg\min_{z\in\mathcal{X}} f(z)$ is nonempty.

Proposition 5. *Suppose that f is a lower semicontinuous function* (*see* Section 7 *for the definition*) *having growth function* Ψ *where* $\lim_{d\to 0}(\Psi(d))/d^{\delta}=C$. *If the answer to* (Q1) *is positive, then also the answer to* (Q3) *is positive in the following sense:*

(i) *If* $\beta(N)\sup_{x\in\mathcal{X}}|\hat{f}_N(x)-f(x)|$ *is bounded in probability with tail function* $\varepsilon(K)$, *then*

$$[\beta(N)]^{1/\delta}\,\mathbb{d}\left(\arg\min_{x\in\mathcal{X}} \hat{f}_N(x) \parallel \arg\min_{x\in\mathcal{X}} f(x)\right)$$

is bounded in probability with tail function $\varepsilon(CK^{\delta/2})$.

(ii) *Let* $\overline{\Psi}(d) = \Psi(d)/d$ *and suppose that* $\overline{\Psi}(d)$ *is strictly increasing and satisfies* $\overline{\Psi}(d) \to 0$ *as* $d \to 0$. *If the difference quotients have the property that*

$$\beta(N) \sup_{x,y \in \mathcal{X}, x \neq y} \frac{|\hat{f}_N(x) - \hat{f}_N(y) - f(x) + f(y)|}{\|x - y\|}$$

is bounded in probability with tail function $\varepsilon(K)$, *then*

$$[\beta(N)]^{1/(\delta-1)\mathbb{d}} \left(\arg\min_{x \in \mathcal{X}} \hat{f}_N(x) \parallel \arg\min_{x \in \mathcal{X}} f(x) \right)$$

is bounded in probability with tail function $\varepsilon(CK^{\delta-1})$.

Proof. The proof rests on the following inequalities:

(i) If $\sup_x |f(x) - g(x)| \leq \varepsilon$ then for each minimizer x^+ of g

$$\mathbb{d}(x^+, \arg\min f) \leq \Psi^{-1}(2\varepsilon). \tag{3.29}$$

(ii) If for all $x \neq y$ we have that $|f(x) - g(x) - f(y) + g(y)| \leq \varepsilon \|x - y\|$, then for each minimizer x^+ of g

$$\mathbb{d}(x^+, \arg\min f) \leq \overline{\Psi}^{-1}(\varepsilon). \tag{3.30}$$

where $\overline{\Psi}(d) = \Psi(d)/d$. We show (3.29) first. Let $x^* \in \arg\min f$ such that $\|x^+ - x^*\| = \mathbb{d}(x^+, \arg\min f)$. Such a point x^+ exists because argmin f is closed by the lower semicontinuity property. Since Ψ is a growth function,

$$\begin{aligned} 0 &\geq g(x^+) - g(x^*) \geq f(x^+) - f(x^*) - |f(x^*) - g(x^*)| - |f(x^+) - g(x^+)| \\ &\geq \Psi(\|x^* - x^+\|) - 2\varepsilon \end{aligned}$$

whence (3.29) follows. To show (ii), we use the following inequality

$$\begin{aligned} 0 &\geq g(x^+) - g(x^*) \geq f(x^+) - f(x^*) - |f(x^*) - g(x^*) - f(x^+) + g(x^+)| \\ &\geq \Psi(\|x^* - x^+\|) - \varepsilon \|x^* - x^+\| \end{aligned}$$

whence $\Psi(\|x^* - x^+\|)/\|x^* - x^+\| = \overline{\Psi}(\|x^* - x^+\|) \le \varepsilon$ and we have established (3.29). The proposition follows from setting $g(x) = \hat{f}_N(x)$ and $\varepsilon = (k)/(\beta(N))$. □

Example 2. **The Newsboy problem (continued)**. Recall again this problem from the chapter *Stochastic Programming Models* of this volume. It was already shown earlier that the objective function has the representation

$$f(x) = \mathbb{E}[F(x, \xi)] = (s - c)x + (r - s)G^{(1)}(x),$$

where $G^{(1)}(u) = \int_{-\infty}^{u} (u - v)\,\mathrm{d}G(v) = \int_{-\infty}^{u} G(v)\,\mathrm{d}v$. Suppose that G has continuous density g and that it is known that in a symmetric neighborhood of diameter $2h$ this density is not smaller than a constant $g_0 > 0$. Then the following function Ψ is a growth function for f:

$$\Psi(d) = \begin{cases} (r - s)\frac{1}{2}\, g_0 d^2 & d \le h \\ (r - s)(g_0 dh - \frac{1}{2}\, g_0 h^2) & d > h \end{cases}$$

Thus f is locally quadratic in a neighborhood of x^*, but grows globally only at a linear rate. The curvature at the minimum is $f''(x^*) = (r - s)g(x^*) > 0$. The function $-f(x)$ is depicted in Fig. 1. In addition, this figure shows three empirical estimates $\hat{f}_N(x)$, each based on a sample size of $N = 10$. Observe that the approximating function may be quite far from the true function, but the arg mins may be rather close (Fig. 2).

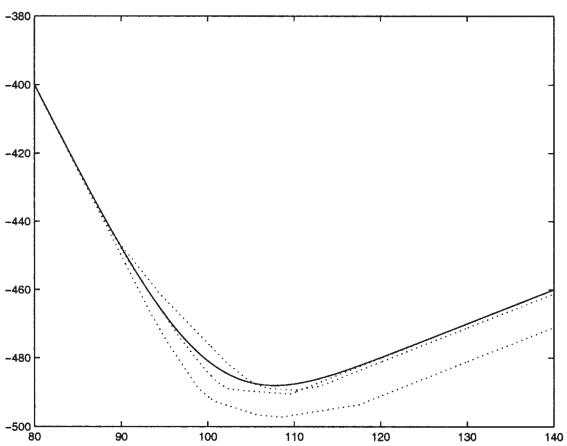

Fig. 2. Solid line: The negative objective function of the newsboy problem ($r = 9$, $c = 10$, $s = 15$, $\xi \sim N(100, 8)$). This function is locally quadratic, but globally only of linear growth. Dotted lines: Three empirical approximations of the objective function.

The growth function depends on the feasible set. If one knows that the approximating solution is already in a small neighborhood of the true solution, then usually a better growth function can be established and hence a smaller confidence interval can be found. For practical purposes, a two stage procedure may be used: in a first stage, the validity of some confidence region may be shown using a global growth function. In a second stage this confidence region may be improved using a local growth function.

3.3 The uniform boundedness property

We have seen so far that the crucial question is (Q1), i.e., to bound the approximation error of the objective function. How can this be done? The approach we will use here is to impose Lipschitz continuity properties on the risk functional $\mathbb{F}$ and to use uniform bounds for the Law of Large Numbers.

The first step is to introduce some sup-metrics $d_{\mathcal{H}}$. Let $\mathcal{H}$ be a set of measurable functions on $\mathbb{R}$ and let $\mathcal{G}_{\mathcal{H}}$ the class of those distributions G, for which all functions from $\mathcal{H}$ have finite integrals:

$$\mathcal{G}_{\mathcal{H}} = \left\{ G \colon \int h(u)\, \mathrm{d}G(u) \text{ is well defined and finite for all } h \in \mathcal{H} \right\}.$$

On $\mathcal{G}_{\mathcal{H}}$ one may define a semidistance through

$$d_{\mathcal{H}}(G_1, G_2) = \sup\left\{ \left| \int h(u)\, \mathrm{d}G_1(u) - \int h(u)\, \mathrm{d}G_2(u) \right| : h \in \mathcal{H} \right\}.$$

Recall the definition of the Lipschitz property: $\mathbb{F}$ is Lipschitz continuous w.r.t. $d_{\mathcal{H}}$ if there is a constant L (called the Lipschitz constant) such that

$$\left| \mathbb{F}(G_1) - \mathbb{F}(G_2) \right| \leq L \cdot d_{\mathcal{H}}(G_1, G_2). \tag{3.31}$$

Example. If $\mathbb{F} = \mathbb{E}$, i.e., if the risk functional is the expectation, then (3.31) holds with $L = 1$, if $\mathcal{H}$ contains the identity function $h(x) = x$. Other, less trivial examples are discussed in Section 4.

The relation between distances of probability measures and the pertaining stochastic programs is extensively studied in chapter *Stability of Stochastic Programming Problems* by Römisch in this volume. Here we are interested in uniform approximations of the true program by the empirical program. The uniform approximation may be expressed in terms of a uniform boundedness principle of appropriate families $\mathcal{F}$ of integrable functions.

Definition 6. A family $\mathcal{F}$ of integrable functions defined on a probability space $(\Omega, \mathcal{A}, P)$ has the *uniform boundedness property with blow-up function* $\beta(N)$ *and explicit tail function* $\varepsilon(K)$, if for i.i.d. random variables (ξ_n) with distribution P,

$$\beta(N) \sup_{F\in\mathcal{F}} \left| \frac{1}{N} \sum_{n=1}^{N} F(\xi_n) - \mathbb{E}[F(\xi)] \right|$$

is bounded in probability with tail function $\varepsilon(K)$, i.e.,

$$\sup_N \mathbb{P}\left\{ \beta(N) \sup_{F\in\mathcal{F}} \left| \frac{1}{N} \sum_{n=1}^{N} F(\xi_n) - \mathbb{E}[F(\xi)] \right| \geq K \right\} \leq \varepsilon(K)$$

(compare with Definition 2 (i)). If the tail function $\varepsilon(K)$ satisfies (3.16), then the tails are called normal.

The uniform boundedness property of classes of function is the key for proving uniform bounds for the error in the objective function, i.e., the difference between $f(x) = \mathbb{F}(G_x)$ and $\hat{f}_N(x) = \mathbb{F}(\hat{G}_{x,N})$, compare (2.6) and (2.7). The following theorem is obvious, but crucial.

Theorem 7. *Suppose that* $\mathbb{F}$ *is Lipschitz continuous w.r.t.* $d_{\mathcal{H}}$, *i.e., that* (3.31) *is fulfilled.*

(i) *If the family* $\mathcal{F}_{\mathcal{H}} = \{h \circ F(x,\cdot)\colon x \in \mathcal{X}, h \in \mathcal{H}\}$ *satisfies the uniform boundedness property with blow-up function* $\beta(N)$ *and explicit tail function* $\varepsilon(K)$, *then*

$$\beta(N) \sup_{x\in\mathcal{X}} |\hat{f}_N(x) - f(x)|$$

is bounded in probability (*see Definition 2*) *with tail function* $\varepsilon(K/L)$. *Thus, under the assumptions of Proposition 5*,

$$[\beta(N)]^{1/\delta}\, \mathbb{d}\left(\arg\min_{x\in\mathcal{X}} \hat{f}_N(x) \parallel \arg\min_{x\in\mathcal{X}} f(x) \right)$$

is bounded in probability with tail function $\varepsilon(C(K/L)\delta/2)$.

(ii) *If the family* $\mathcal{F}^{\nabla}_{\mathcal{H}} = \{(h \circ F(x,\cdot) - h \circ F(y,\cdot))/(\|x-y\|)\colon x,y \in \mathcal{X};\ x \neq y\}$ *satisfies the uniform boundedness property with blow-up function* $\beta(N)$ *and explicit tail function* $\varepsilon(K)$, *then*

$$\beta(N) \sup_{x,y\in\mathcal{X}, x\neq y} \frac{|\hat{f}_N(x) - \hat{f}_N(y) - f(x) + f(y)|}{\|x-y\|}$$

is bounded in probability with tail function $\varepsilon(K/L)$. *Thus, under the assumptions of Proposition 5,*

$$[\beta(N)]^{1/(\delta-1)}\, \mathbb{d}\left(\arg\min_{x\in\mathcal{X}} \hat{f}_N(x) \parallel \arg\min_{x\in\mathcal{X}} f(x)\right)$$

is bounded in probability with tail function $\varepsilon(C(K/L)^{\delta-1})$.

To summarize the results of this section, we have seen that bounds for the uniform approximation of the objective function gives automatically bounds for the quality of the solution. Bounding the approximation error of the arg min is a more difficult problem, since it depends on the curvature of the objective f (the growth condition). If the objective is flat near its arg min, a small error in function value may cause a large error in the arg min. Also, bounds for the differences of the objective give much better bounds for the solutions than bounds for function values only.

If the risk functional $\mathbb{F}$ is the expectation, uniformity in the class $\mathcal{F} = \{F(x,\cdot)\colon x \in \mathcal{X}\}$ or $\mathcal{F}^{\nabla} = \{(F(x,\cdot) - F(y,\cdot))/(\|x-y\|)\colon x, y \in \mathcal{X};\ x \neq y\}$ is required. If the risk functional is not the expectation, but at least Lipschitz continuous w.r.t. $d_{\mathcal{H}}$, then uniform versions of the Law of Large Numbers are needed for larger classes $\mathcal{F}_{\mathcal{H}} = \{h \circ F(x,\cdot)\colon x \in \mathcal{X}, h \in \mathcal{H}\}$ or $\mathcal{F}_{\mathcal{H}}^{\nabla} = \{(h \circ F(x,\cdot) - h \circ F(y,\cdot))/(\|x-y\|)\colon x, y \in \mathcal{X};\ x \neq y\}$.

In most applications, the quality of the solution matters much more than the arg min error. But even if one is interested in the arg min error, the error in the objective function has to be bounded first: all methods of dealing with the arg min error are based on bounds for the error in the objective function. Two steps are necessary for this goal: first to establish the Lipschitz continuity for the risk functional (this step can be omitted if the risk functional is the expectation) and then to prove uniformity for the appropriate function class $\mathcal{F}$. The first step will be considered in the next Section 4, the second step is done in the subsequent Section 5.

4 Risk functionals and Lipschitz properties

The expectation is not the only functional which appears in stochastic optimization problems. Especially in recent years it has been well understood that risk management is done by controlling various risk parameters like variance, quantiles etc., of the cost, wealth or income distribution.

To minimize the expected costs makes sense for situations which contain many repetitions of the same problem. Examples are the optimal design of service systems, manufacturing systems and inventory systems. If, however, the decision is only to be made once, then it is questionable whether the expectation is the appropriate risk functional for the problem. In portfolio optimization for instance, the use of variance risk adjusted expectation

$\mathbb{F}(G_x) = \mathbb{E}(G_x) - \rho \cdot \mathbb{V}\mathrm{ar}(G_x)$ as objective has become a standard (the so called Markowitz model).

Recall that any probability functional can be seen as a statistical parameter. The basic question, whether and how fast the empirical version of this parameter converges to the true one, depends on the continuity properties of this parameter with respect to appropriate distances.

Based on a set of integrable functions $\mathcal{H}$, we have already defined the semidistance $d_{\mathcal{H}}$ for distributions G_1, G_2 as

$$d_{\mathcal{H}}(G_1, G_2) = \sup\left\{\left| \int h(u)\, \mathrm{d}G_1(u) - \int h(u)\, \mathrm{d}G_2(u)\right| : h \in \mathcal{H}\right\}.$$

If $\mathcal{H}$ is *separating* on $\mathcal{G}_{\mathcal{H}}$, i.e., if $\int h(u)\,\mathrm{d}G_1(u) = \int h(u)\,\mathrm{d}G_2(u)$ for all $h \in \mathcal{H}$ implies that $G_1 = G_2$, then $d_{\mathcal{H}}$ is a distance. The larger the class $\mathcal{H}$, the finer is the topology generated by $d_{\mathcal{H}}$.

Notice also that enlarging the set $\mathcal{H}$ to its convex hull $\mathrm{conv}(\mathcal{H})$ does not change the semidistance $d_{\mathcal{H}}$.

4.1 A list of distances for probability measures

- If $\mathcal{H}$ is chosen as the class of indicator functions of half-open intervals in $\mathbb{R}^1$

  $$\mathcal{KS} = \{\mathbf{1}_{(-\infty,a]}(u) : a \in \mathbb{R}\},$$

  the *Kolmogorov–Smirnov distance* $d_{\mathcal{KS}}$ is obtained.
- If $\mathcal{H}$ is chosen as the class of Lipschitz-continuous functions with Lipschitz-constant 1

  $$\mathcal{L}\mathrm{ip} = \{h : |h(u) - h(v)| \le |u - v|\},$$

  the *Wasserstein distance* $d_{\mathcal{L}\mathrm{ip}}(G_1, G_2)$ is obtained.
- Define the Lipschitz-constant of order p of h as

  $$L_p(f) = \inf\{L : |h(u) - h(v)| \le L|u - v| \max(1, |u|^{p-1}, |v|^{p-1})\}.$$

  Choosing $\mathcal{H}$ as the class of all functions with $L_p \le 1$, we get the *Fortet–Mourier distance*.
- If $\mathcal{H}$ is chosen as the class of the indicator functions $\mathbf{1}_A$ of all measurable sets (or equivalently the class $\mathcal{M}_b^+(1)$ of all nonnegative measurable functions bounded by 1), the *variational distance* is obtained

  $$d_{\mathcal{M}_b^+(1)}(G_1, G_2) = \sup_A \int \mathbf{1}_A(u)\, d[G_1 - G_2](u).$$

- Consider the class of power functions given by

$$\mathcal{H}_k = \{\text{sign}(u)|u|^k \mathbf{1}_{\{u \geq a\}} : a \in \mathbb{R}\} \tag{4.1}$$

for integer k. The corresponding distance will be briefly denoted by

$$d_k(G_1, G_2) \quad \text{instead of} \quad d_{\mathcal{H}_k}(G_1, G_2). \tag{4.2}$$

By definition, if $X_1 \sim G_1$ and $X_2 \sim G_2$, then

$$d_k(G_1, G_2) = \sup\{|\mathbb{E}(X_1^k \mathbf{1}_{\{X_1 \geq a\}}) - \mathbb{E}(X_2^k \mathbf{1}_{\{X_2 \geq a\}})| : a \in \mathbb{R}\}.$$

In particular, $|\mathbb{E}(X_1^k) - \mathbb{E}(X_2^k)| \leq d_k(G_1, G_2)$. The distance d_0 coincides with the Kolmogorov–Smirnov metric $d_{\mathcal{KS}}$, since $\sup_u |G_1(u) - G_2(u)| = \sup_u |(1 - G_1(u)) - (1 - G_2(u))|$. We also introduce

$$\overline{d}_k = \max\{d_\ell : \ell = 0, \ldots, k\}. \tag{4.3}$$

4.2 A list of frequently used risk functionals

Recall that risk functionals are real valued mappings $\mathcal{G} \overset{\mathbb{F}}{\to} \mathbb{R}$ defined on the family of $\mathcal{G}$ of all probability measures on $\mathbb{R}$ or some subfamily.

- **Location equivariant measures (location parameters).**
  They have the property that

$$\mathbb{F}[G(\cdot - a)] = \mathbb{F}[G(\cdot)] + a$$

for all a. Examples are the expectation

$$\mathbb{E}(G) = \int u \, \mathrm{d}G(u),$$

the value-at-risk (at level α)

$$\mathbb{V}@\mathrm{R}_\alpha(G) = G^{-1}(1 - \alpha),$$

the conditional value-at-risk (at level α)

$$\mathbb{CV}@\mathrm{R}_\alpha(G) = \frac{1}{\alpha} \int_{1-\alpha}^{1} G^{-1}(v) \, \mathrm{d}v = \frac{1}{\alpha} \int_{(G^{-1}(1-\alpha), \infty)} u \, \mathrm{d}G(u).$$

Alternatively, $\mathbb{CV}@\mathrm{R}$ may be defined as

$$\mathbb{CV}@\mathrm{R}(G) = \inf\left\{a + \frac{1}{1-\alpha}\int [u-a]_+ \,\mathrm{d}G(u) \colon a \in \mathbb{R}\right\}$$

(see Rockafellar and Uryasev, 1999).

- **Location invariant measures (pure risk measures).**
They have the property that

$$\mathbb{F}[G(\cdot - a)] = \mathbb{F}[G(\cdot)]$$

for all a. Examples arethe mean absolute deviation

$$\mathbb{M}\mathrm{ad}(G) = \int |u - \mathbb{E}(G)| \,\mathrm{d}G(u),$$

the upper semi-deviation

$$\mathbb{M}\mathrm{ad}_+(G) = \int [u - \mathbb{E}(G)]_+ \,\mathrm{d}G(u),$$

the variance

$$\mathbb{V}\mathrm{ar}(G) = \int [u - \mathbb{E}(G)]^2 \,\mathrm{d}G(u),$$

the upper semi-variance

$$\mathbb{V}\mathrm{ar}_+(G) = \int ([u - \mathbb{E}(G)]_+)^2 \,\mathrm{d}G(u).$$

- **Other risk measures.**
The expected utility

$$\int U(v) \,\mathrm{d}G(v),$$

where U is a nonnegative, monotone utility function, the excess probability (over threshold t)

$$1 - G(t).$$

4.3 Continuity properties of risk functionals

Suppose that we may prove that the functional $\mathbb{F}$ is continuous w.r.t $d_{\mathcal{H}}$. Then the convergence of $\hat{G}_{x,N}$ to G_x in metric $d_{\mathcal{H}}$ implies the convergence of $\mathbb{F}(\hat{G}_{x,N})$ to $\mathbb{F}(G_x)$. If $\mathbb{F}$ is Lipschitz, then every estimate for $d_{\mathcal{H}}(\hat{G}_{x,N}, G_x)$ implies in a simple manner an estimate for $|\mathbb{F}(\hat{G}_{x,N}) - \mathbb{F}(G_x)|$. Uniformity of the convergence in x carries over.

The following Proposition 8, which is stated without proof uses the distances d_k and $\overline{d}_k$ introduced in (4.2) and (4.3).

Proposition 8. Continuity of probability functionals. *Let $X_1 \sim G_1$ and $X_2 \sim G_2$. Then*

(i) $|G_1(t) - G_2(t)| \leq d_0(G_1, G_2)$

(ii) $|\mathbb{E}(G_1) - \mathbb{E}(G_2)| \leq d_1(G_1, G_2)$

(iii) $|\mathbb{M}\text{ad}(G_1) - \mathbb{M}\text{ad}(G_2)| \leq \overline{d}_1(G_1, G_2) + 2d_1(G_1, G_2)$

(iv) $|\mathbb{V}\text{ar}(G_1) - \mathbb{V}\text{ar}(G_2)| \leq d_2(G_1, G_2) + d_1(G_1, G_2)(\mathbb{E}X_1 + d_1(G_1, G_2)$

(v) $|\mathbb{V}\text{ar}_+(G_1) - \mathbb{V}\text{ar}_+(G_2)| \leq \overline{d}_2(G_1, G_2)(2\mathbb{E}X_1 + 1) + d_1(G_1, G_2)[3 + \mathbb{E}|X_1 - \mathbb{E}X_1|]$

(vi) *Suppose that $G_1(q) = \alpha$ and that $G_1(u) - G_1(q) \geq c(u - q)$ for $|u - q| \leq \varepsilon$. Then, if $d_0(G_1, G_2) \leq \varepsilon$,*

$$|\mathbb{V}@\mathrm{R}_\alpha(G_1) - \mathbb{V}@\mathrm{R}_\alpha(G_2)| \leq \frac{1}{c}\, d_0(G_1, G_2).$$

(vii) $|\mathbb{CV}@\mathrm{R}_\alpha(G_1) - \mathbb{CV}@\mathrm{R}_\alpha(G_2)| \leq d_0(G_1, G_2) + |\mathbb{V}@\mathrm{R}_\alpha(G_1) - \mathbb{V}@\mathrm{R}_\alpha(G_2)|$

5 Arithmetic means of of i.i.d. random variables

This section reviews some results from probability theory, which are relevant for deriving statistical properties of the estimates in stochastic optimization.

5.1 Basic properties

Let (X_n) be a sequence of i.i.d. random variables. We are interested in the behavior of their arithmetic mean

$$\overline{X}_N = \frac{1}{N} \sum_{n=1}^{N} X_n$$

as N tends to infinity. There are four basic theorems, which describe the pointwise and distributional properties of $\overline{X}_N$: the Strong Law of Large Numbers (SLLN), the Central Limit Theorem (CLT), the Law of Iterated Logarithm (LIL) and the Large Deviations Theorem (LDT).

Theorem 9. **(Strong Law of Large Numbers-SLLN)** *Let* (X_n) *be a sequence of i.i.d. random variables. If* X_1 *is integrable, then*

$$\overline{X}_N = \frac{1}{N}\sum_{n=1}^{N} X_n \to \mathbb{E}X_1 \quad \textit{almost surely}.$$

Proof. See Etemadi (1981). □

Theorem 10. **(Central Limit Theorem-CLT)**. *Assume that* X_1 *has finite variance* $\sigma^2 = \mathbb{V}\mathrm{ar}(X_1)$. *Then, for all* u,

$$\mathbb{P}\left\{\frac{1}{\sigma\sqrt{N}}\sum_{n=1}^{N}[X_n - \mathbb{E}X_1] \le u\right\} \to \Phi(u) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{u} e^{-v^2/2}\,\mathrm{d}v.$$

Proof. See Shiryaev (1996, pp. 328–341) for this and more on the CLT. □

Theorem 11. **(Law of iterated logarithm-LIL)**. *Under the assumptions of the previous theorem the following limit relations hold almost surely*

$$\limsup_{N\to\infty} \frac{1}{\sigma\, 2\sqrt{N\log\log N}}\sum_{n=1}^{N}[X_n - \mathbb{E}X_1] = 1$$

$$\liminf_{N\to\infty} \frac{1}{\sigma\, 2\sqrt{N\log\log N}}\sum_{n=1}^{N}[X_n - \mathbb{E}X_1] = -1.$$

Proof. See Shiryaev (1996), Theorem 1, p. 397. □

The LDT will be presented in the next section.

5.2 Universal probability bounds

Assume as before that (X_n) is an i.i.d sequence. Let all X_n have the same distribution as X. If X has exponential moments of order $s \neq 0$ (i.e., $\mathbb{E}(e^{sX})$ exists for some $s \neq 0$), then interesting inequalities and tight bounds may be derived. The first basic step is Chernoff's inequality for $s > 0$:

$$\mathbb{P}\{X \ge t\} = \mathbb{P}\{e^{sX} \ge e^{st}\} \le e^{-st}\mathbb{E}(e^{sX}).$$

Applied to the arithmetic mean of the centered variables $\frac{1}{N}\sum_{n=1}^{N} X_n - \mathbb{E}X = \overline{X}_N - \mathbb{E}X$, the Chernoff bound reads

$$\mathbb{P}\{\overline{X}_N - \mathbb{E}X \ge t\} \le e^{-stN}[\mathbb{E}(e^{s(X-\mathbb{E}X)})]^N. \tag{5.1}$$

This bound is one-sided, but putting together this inequality and the corresponding inequality for $-X$ one gets easily the two-sided version, since for $t \geq 0$,

$$\mathbb{P}\{|\overline{X}_N - \mathbb{E}X| \geq t\} \leq \mathbb{P}\{\overline{X}_N - \mathbb{E}X \geq t\} + \mathbb{P}\{-\overline{X}_N + \mathbb{E}X \geq t\}. \tag{5.2}$$

If the random variables X_n are bounded in $[a, b]$, then the Hoeffding bound

$$\mathbb{P}\{\overline{X}_N - \mathbb{E}X \geq t\} \leq e^{-2t^2N/(b-a)^2} \tag{5.3}$$

can be obtained (see Devroye and Lugosi, 2001). Consequently, if the X_n are bounded and take values in $[-C, C]$, we get normal tail behavior for the arithmetic mean blown up with $\sqrt{N}$ in the sense of Definition 2

$$\mathbb{P}\left\{|\overline{X}_N - \mathbb{E}X| \geq K/\sqrt{N}\right\} \leq 2e^{-K^2/2C^2} \tag{5.4}$$

by setting $t = K/\sqrt{N}$.

If the random variables X_n are unbounded, Hoeffding's bound is not applicable, but a similar bound may be derived using the logarithmic moment generating function of $X - \mathbb{E}X$:

$$M_X(s) = \log \mathbb{E}(e^{s(X-\mathbb{E}X)}).$$

We assume that $M_X(s)$ is finite for s in a neighborhood $[-\varepsilon, \varepsilon]$ of 0. Outside this neighborhood M_X may be equal ∞. The region of finiteness of M_X is determined by the tail behavior of X: if X is bounded, then M_X is finite on the whole $\mathbb{R}$. The lighter are the tails, the larger is the interval of finiteness of M_X. M_X is a convex, extended real function which satisfies

$$M_X(0) = 0, \quad M'_X(0) = 0, \quad M''_X(0) = \mathbb{V}\mathrm{ar}(X).$$

Let M^*_X be the conjugate (dual) function of M_X

$$M^*_X(u) = \sup\{su - M_X(s) \colon s \in \mathbb{R}\} \tag{5.5}$$

which is also convex. Taking the infimum over all s in (5.1) we get

$$\begin{aligned}\mathbb{P}\{\overline{X}_N - \mathbb{E}X \geq t) &\leq \inf_s e^{-stN}[\mathbb{E}(e^{s(X-\mathbb{E}X)})]^N \\ &= \inf_s \exp(N[M_X(s) - st]) = \exp(-NM^*_X(t)).\end{aligned} \tag{5.6}$$

Notice that the inequality (5.6) holds for all t and all N. The following lemma links the growth of the conjugate logarithmic moment generating function to the tail behavior of the arithmetic mean.

Lemma 12.

(i) *If the logarithmic moment generating function M_X satisfies*

$$M_X(t) \le \alpha t^2, \tag{5.7}$$

for all t and for some constant α, then the arithmetic mean has normal tail behavior:

$$\mathbb{P}\left\{\overline{X}_N - \mathbb{E}X \ge K/\sqrt{N}\right\} \le e^{-K^2/4\alpha}. \tag{5.8}$$

(ii) *If $M_X(s) \le \alpha s^2$ for $|s| \le \varepsilon$, then*

$$\mathbb{P}\left\{\overline{X}_N - X \ge K/\sqrt{N}\right\} \le \begin{cases} e^{-K^2/4\alpha} & \text{if } N \ge \frac{K^2}{4\alpha^2\varepsilon^2} \\ e^{-\sqrt{N}K\varepsilon/2} & \text{otherwise} \end{cases} \tag{5.9}$$

i.e., the arithmetic mean has normal tails for sufficiently large N.

Proof. (i) If $M_X(s) \le \alpha s^2$, then $M_X^*(t) \ge t^2/4\alpha$. (5.8) follows from setting $t = K/\sqrt{N}$ in (5.6). In case (ii) $M_X^*(t) \ge \frac{t^2}{4\alpha}\mathbf{1}_{\{|t|\le 2\alpha\varepsilon\}} + (|t|\varepsilon - \alpha\varepsilon^2)\mathbf{1}_{\{|t|>2\alpha\varepsilon\}} \ge \frac{t^2}{4\alpha}\mathbf{1}_{\{|t|\le 2\alpha\varepsilon\}} + (|t|\varepsilon/2)\mathbf{1}_{\{|t|>2\alpha\varepsilon\}}$ and this must be used in (5.6). □

A slight generalization of the just proven result is Cramer's theorem (see Deuschel and Strook, 1989).

Theorem 13. **(Large deviations theorem-LDT)**. *For every closed set A*

$$\limsup_N \frac{1}{N}\log \mathbb{P}\{\overline{X}_N \in A\} \le -\inf\{M_X^*(t)\colon t \in A\}$$

Remark. A lower bound is also valid (see also Deuschel and Strook, 1989): for every open set B

$$\liminf_N \frac{1}{N}\log \mathbb{P}\{\overline{X}_N \in B\} \ge -\inf\{M_X^*(t)\colon t \in B\}.$$

5.3 Uniformity

Let $\mathcal{F}$ be a class of measurable functions on $(\Omega, \mathcal{A}, P)$ and let (ξ_n) be a sequence of i.i.d. random variables with values in Ω. If the $\mathcal{F}$ contains only one function F, then setting $X_n = F(\xi_n)$, we know that this sequence fulfills under the appropriate conditions SLLN, CLT, LIL, LDT.

For the applications we have in mind, uniform versions of these theorems are needed. Let $\mathcal{F}$ be a class of functions defined on $\mathbb{R}^m$ and (ξ_n) a sequence of i.i.d. random variables with values in $\mathbb{R}^m$.

We want to know under which conditions for $\mathcal{F}$ a uniform Law of Large Numbers with explicit tail behavior as introduced in Definition 6 holds, i.e.,

$$\sup_{F\in\mathcal{F}} \beta(N)\left|\frac{1}{N}\sum_{n=1}^{N} F(\xi_n) - \mathbb{E}F(\xi_1)\right| \tag{5.10}$$

is uniformly bounded in probability with explicit tail behavior.

For measurability assumptions, we assume throughout this chapter that $\mathcal{F}$ is a countable family of functions.

If the measurability of the supremum error can be ensured, then the family $\mathcal{F}$ may be uncountable as well. In our application, $\mathcal{F}$ will be the family of all cost functions $\mathcal{F} = \{F(x,\xi)\colon x \in \mathcal{X}\}$. If $x \mapsto F(x,\xi)$ is a.s. lower semicontinuous (see Section 7), then countably many $x \in \mathcal{X}$ suffice to determine the supremum error. Therefore, all results hold for the uncountable set $\mathcal{F}$ in this case.

We begin with just a uniform Law of Large Numbers without blowing up function and explicit tail behavior.

Definition 14. The countable class $\mathcal{F}$ of integrable functions is said to fulfill a uniform Strong Law of Large Numbers (SLLN), if

$$\sup_{F\in\mathcal{F}} \frac{1}{N}\sum_{n=1}^{N} [F(\xi_n) - \mathbb{E}[F(\xi_1)]] \to 0 \tag{5.11}$$

almost surely, as N tends to infinity.

Uniform SLLN Theorems are also known under the name of Glivenko–Cantelli Theorems. The original Glivenko–Cantelli Theorem was independently formulated by Glivenko and Cantelli for the function class $\mathcal{F} = \{\mathbf{1}_{(-\infty,u]}(v)\colon u \in \mathbb{R}\}$. Uniformity holds in a trivial manner, if the class $\mathcal{F}$ contains only finitely many functions. Typically, sufficient conditions for uniformity for infinite $\mathcal{F}$ require that this set may be approximated in a certain sense by finitely many functions. One distinguishes between shattering type conditions and covering type conditions.

Theorem 15. Vapnik–Cervonenkis shattering type theorem *Let $\mathcal{U}$ be a subfamily of measurable sets. The Vapnik–Cervonenkis (VC) index $V(\mathcal{U})$ of $\mathcal{U}$ is defined as*

$$V(\mathcal{U}) = \inf\left\{n: \max_{x_1,\dots,x_n} \#\{U \cap \{x_1,\dots,x_n\}: U \in \mathcal{U}\} < 2^n\right\}.$$

Here # denotes the cardinality. If $V(\mathcal{U}) < \infty$, then

$$\sup_{U\in\mathcal{U}} |\hat{P}_N(U) - P(U)| \to 0 \quad a.s.$$

Proof. See Shorack and Wellner (1986), p. 828. □

The VC index was formulated for families of sets, but it has also a version for families of functions: as an easy corollary one gets the following theorem: a class of functions $\mathcal{F}$ satisfies a uniform SSLN (5.11), if the collection of all level sets $\{x: F(x) < t\}$, $t \in \mathbb{R}$, $F \in \mathcal{F}$ has finite VC index.

The finiteness of the VC index guarantees that the function class satisfies the uniform SLLN, but does not allow finer estimates of the deviation of the empirical functions from the true functions. A more refined analysis is made possible by the notion of the *metric entropy*. Define first the notion of a bracket $[G, H]$: for two real, measurable functions $G \le H$, let the *bracket* be defined as

$$[G, H] = \{F: G \le F \le H \text{ a.s.}\}.$$

The bracket is empty, if $G \le H$ a.s. does not hold.

Definition 16. Metric entropy (with bracketing). Define $\|F\|_1 = \int |F(\omega)|\, \mathrm{d}P(\omega)$. Let $\mathcal{F}$ be a family of integrable functions. The *metric entropy* of $\mathcal{F}$ is defined as

$$N_1(\epsilon) = \min\Big\{k: \text{There exist integrable functions } G_1 \le H_1, \dots, G_k \le H_k, \text{ such that } \mathcal{F} \subseteq \bigcup_{i=1}^{k} [G_i, H_i] \text{ and } \|H_i - G_i\|_1 < \varepsilon \text{ for all } i\Big\}$$

Function classes with finite entropy satisfy a uniform Law of Large Numbers as the next theorem shows.

Theorem 17. *Blum-de Hardt metric entropy type theorem. Suppose that $N_1(\varepsilon)$ is finite for every $\varepsilon > 0$. Then the uniform SLLN (5.11) holds.*

Proof. Choose an arbitrary $\varepsilon > 0$. It is sufficient to show that

$$\limsup_N \sup_{F\in\mathcal{F}} \frac{1}{N}\sum_{n=1}^{N} F(\xi_n) - \mathbb{E}[F(\xi)] \le \varepsilon$$

a.s. Let $(G_i, H_i)_{i=1,\dots,N_1(\varepsilon)}$ be chosen as in Definition 16. Then

$$\begin{aligned}\sup_{F\in\mathcal{F}} \frac{1}{N}\sum_{n=1}^{N} F(\xi_n) - \mathbb{E}[F(\xi)] &= \max_{i=1,\dots,N_1(\varepsilon)} \sup_{F\in[G_i,H_i]} \frac{1}{N}\sum_{n=1}^{N} F(\xi_n) - \mathbb{E}[F(\xi)] \\ &\le \max_{i=1,\dots,N_1(\varepsilon)} \frac{1}{N}\sum_{n=1}^{N} H_i(\xi_n) - \mathbb{E}[G_i(\xi)] \\ &\le \max_{i=1,\dots,N_1(\varepsilon)} \frac{1}{N}\sum_{n=1}^{N} H_i(\xi_n) - \mathbb{E}[H_i(\xi)] + \varepsilon\end{aligned}$$

By the standard Law of Large Numbers, the lim sup of this expression for $N\to\infty$ is bounded by ε and the theorem is shown. □

A more refined analysis relates the growth of the function $N_1(\varepsilon)$ to the tail behavior of the sequence (5.10). If an assumption about the growth of the function $\varepsilon \mapsto N_1(\varepsilon)$ can be made, then good estimates of the approximation error are possible. The assumptions about $N_1(\varepsilon)$ are called *covering type* conditions.

5.4 *Covering types and universal bounds*

Definition 18. A set $A \subset \mathbb{R}^d$ is said to be of *covering type* (v, V), if for every $\varepsilon > 0$ one can find at most $\lfloor (V_\varepsilon/\varepsilon)^v \rfloor$ balls $B_1, B_2, \dots, B_{N_\varepsilon}$, each with diameter ε, which cover A, i.e., $A \subseteq \cup_{i=1}^{N_\varepsilon} B_i$ and $\lim_{\varepsilon\to 0} V_\varepsilon = V$.

Example 5. The unit cube in $\mathbb{R}^d$ is of covering type $(d, 2\sqrt{d})$.

Definition 19. Let $(\Omega, \mathcal{A}, P)$ be a probability space. A family $\mathcal{F}$ of L^2-functions is of covering type (v, V), if for every $\varepsilon > 0$ there are at most $N_2(\varepsilon) = \lfloor (V/\varepsilon)^v \rfloor$ pairs of functions $(G_1, H_1), \dots, (G_{N_2(\varepsilon)}, H_{N_2(\varepsilon)})$ with the properties

(i) $G_i(\omega) \le H_i(\omega)$ a.s. for $1 \le i \le N_2(\varepsilon)$;
(ii) $\int (H_i(\omega) - G_i(\omega))^2 P(\mathrm{d}\omega) \le \varepsilon^2$;
(iii) For each $F \in \mathcal{F}$ there is a index $i \in \{1, \dots, N_2(\varepsilon)\}$ such that

$$G_i(\omega) \le F(\omega) \le H_i(\omega).$$

Property (iii) may be expressed in the following way:

$$\mathcal{F} \subseteq \bigcup_{i=1}^{N_2(\varepsilon)} [G_i, H_i].$$

The covering type is essential for uniform confidence bands as was demonstrated by Talagrand (1994).

Theorem 20. *Let* $|F(w)| \le 1$ *for all* $F \in \mathcal{F}$. *Suppose that* $\mathcal{F}$ *is countable and of covering type* (v, V). *Then*

$$P\left\{\sup_{F\in\mathcal{F}} \left|\frac{1}{N}\sum_{n=1}^{N} F(\xi_n) - \mathbb{E}F\right| \ge \frac{K}{\sqrt{N}}\right\} \le \left(M(V)\,\frac{K}{\sqrt{v}}\right)^{v} \exp(-2K^2), \tag{5.12}$$

where $M(\cdot)$ is a universal function. To put it differently, $\mathcal{F}$ satisfies the uniform boundedness property of Definition 6 with blow-up function $\beta(N) = \sqrt{N}$ and normal tails.

Proof. See Talagrand (1994) or Van der Vaart and Wellner (1996). □

The theorem holds also for uncountable classes $\mathcal{F}$, provided that the supremum is measurable.

Remark. Notice that the factor K^v appearing in the tail function does not affect the normal tail behavior, since

$$K^v \exp(-2K^2) \le \exp(-v/2)\left(\frac{v}{2\delta}\right)^{v/2} \exp(-(2-\delta)K^2)$$

for all $0 < \delta < 2$. Thus the tail behavior is normal for all $v > 0$. If the functions F are bounded by C, then the right hand side has to be replaced by $(M(V)\,K/(C\sqrt{v}))^v \exp(-2K^2/C^2)$, still with normal tails.

If the family $\mathcal{F}$ is unbounded, then the foregoing result has to be modified. If uniform exponential moments exist, then a similar result may be proved, but the price is an extra $[\log(N)]^{-1}$ term in the blow-up function. A bounding technique is used to show this result: let for any random variable X and $c > 0$

$$X^c = \begin{cases} c & \text{if } X > c \\ -c & \text{if } X < -c \\ X & \text{otherwise} \end{cases}$$

Theorem 21. *Suppose that $\mathcal{F}$ is countable and of covering type (v, V). Suppose further that the functions $S_c = \sup_{F\in\mathcal{F}}\{F - F^c - \mathbb{E}[F] - \mathbb{E}[F^c]\}$* satisfy the following conditions

(i) $\mathbb{E}S_c \le e^{-c/2}$
(ii) $\log \mathbb{E}\exp(s[S_c - \mathbb{E}S_c]) \le \alpha s^2$ for all c. Then

$$\mathbb{P}\left\{\sup_{F\in\mathcal{F}}\left|\frac{1}{N}\sum_{n=1}^{N}F(\xi_n) - \mathbb{E}F\right| \ge \frac{K\log(N)}{\sqrt{N}}\right\}$$
$$\le \left(M(V)\,\frac{K}{\sqrt{v}}\right)^{v}\exp(-K^2/2) + \exp\left(\frac{1}{4\alpha}\right)\exp\left(\frac{-K^2\log^2 N}{32\alpha}\right), \quad (5.13)$$

where $M(\cdot)$ is a universal function. To put it differently, $\mathcal{F}$ satisfies the uniform boundedness property (see Definition 6) with blow-up function $\beta(N) = \sqrt{N}/\log(N)$ and normal tails.

Proof. We use a truncation argument and set $c_N = \log(N)$.

$$\mathbb{P}\left\{\sup_{F\in\mathcal{F}}\frac{1}{N}\sum_{n=1}^{N}F(\xi_n) - \mathbb{E}F \ge \frac{K\log(N)}{\sqrt{N}}\right\}$$
$$\le \mathbb{P}\left\{\sup_{F\in\mathcal{F}}\frac{1}{N}\sum_{n=1}^{N}F^{c_N}(\xi_n) - \mathbb{E}F^{c_N} \ge \frac{K\log(N)}{2\sqrt{N}}\right\}$$
$$+ \mathbb{P}\left\{\sup_{F\in\mathcal{F}}\frac{1}{N}\sum_{n=1}^{N}[F(\xi_n) - F^{c_N}(\xi_n) - \mathbb{E}F(\xi_n) + \mathbb{E}F^{c_N}] \ge \frac{K\log(N)}{2\sqrt{N}}\right\}$$
$$\le \mathbb{P}\left\{\sup_{F\in\mathcal{F}}\frac{1}{N}\sum_{n=1}^{N}\frac{1}{c_N}\,F^{c_N}(\xi_n) - \mathbb{E}\left[\frac{1}{c_N}\,F^{c_N}\right] \ge \frac{K}{2\sqrt{N}}\right\}$$
$$+ \mathbb{P}\left\{\frac{1}{N}\sum_{n=1}^{N}\sup_{F\in\mathcal{F}}[F(\xi_n) - F^{c_N}(\xi_n) - \mathbb{E}F + \mathbb{E}F^{c_N}] \ge \frac{K\log(N)}{2\sqrt{N}}\right\}$$
$$\le \left(M(V)\,\frac{K}{2\sqrt{v}}\right)^{v}\exp(-K^2/2)$$
$$+ \mathbb{P}\left\{\frac{1}{N}\sum_{n=1}^{N}S_{c_N}(\xi_n) - \mathbb{E}S_{c_N} \ge \frac{K\log(N)}{2\sqrt{N}} - \mathbb{E}S_{c_N}\right\}.$$

It remains to show the bound for the second summand.

By assumption, $\mathbb{E}S_{c_N} \le \exp(-\log N/2) = N^{-1/2}$. Therefore, $(K\log(N)/(2\sqrt{N}) - \mathbb{E}S_{c_N})^2 \ge ((K-2)\log(N)/(2\sqrt{N})^2$.

Using Lemma 12, the second summand can be estimated by

$$\begin{aligned}
&\mathbb{P}\left\{\frac{1}{N}\sum_{n=1}^{N} S_{c_N}(\xi_n) - \mathbb{E}S_{c_N} \geq \frac{K\log(N)}{2\sqrt{N}} - \mathbb{E}S_{c_N}\right\} \\
&\leq \mathbb{P}\left\{\frac{1}{N}\sum_{n=1}^{N} S_{c_N}(\xi_n) - \mathbb{E}S_{c_N} \geq \frac{K\log(N)-2}{2\sqrt{N}}\right\} \\
&\leq \exp\left(-\frac{(K\log N-2)^2}{16\alpha}\right) \leq \exp\left(\frac{1}{4\alpha}\right)\exp\left(-\frac{K^2\log^2 N}{32\alpha}\right),
\end{aligned}$$

which gives the claimed bound. □

The following inequality is helpful for finding bounds for truncated random variables, which have exponential moments.

Lemma 22. *Suppose that the random variable X satisfies $\mathbb{E}[\exp sX] < \infty$ for some $s > 0$. Then for $c > 0$*

$$\mathbb{E}[X-c]_+ \leq \frac{e^{-sc}}{s}\,\mathbb{E}[\exp sX].$$

Proof. By the convexity of the exponential function, $\exp(sX) \geq \exp(sc) + s\exp(sc)(X-c) \geq s\exp(sc)(X-c)$. Thus,

$$\mathbb{E}([X-c]_+) \leq \frac{e^{-sc}}{s}\,\mathbb{E}[\exp sX 1_{X\geq c}] \leq \frac{e^{-sc}}{s}\,\mathbb{E}[\exp sX]. \qquad \square$$

Example. Consider the Newsboy Example of Section 1. Suppose that $x \in [a,b]$, but ξ is unbounded. Then $F(x,\xi) = \alpha x + \beta[x-\xi]_+$ is also unbounded. Here we have set $s-c=\alpha$ and $r-s=\beta$. Using now the symbol c for the truncation and assuming that $c \geq \alpha b$, we get $F - F^c = [(\alpha+\beta)x - \beta\xi - c]_+$. The supremum of $[F - F^c - \mathbb{E}F + \mathbb{E}F^c]$ is bounded by $\beta[(\alpha+\beta-c)/\beta - \xi]_+ - \beta G^{(1)}((\alpha+\beta-c)/\beta)$. If ξ has some finite exponential moments (e.g., if ξ has normal tails) then conditions (i) and (ii) are fulfilled.

6 Entropic sizes of stochastic programs

In this section we specialize the results about covering types to stochastic optimization problems. We treat only problems with nonrandom constraints of the form $\mathrm{Min}_{x\in\mathcal{X}}\{f(x) = \mathbb{F}[F(x,\xi)]\}$.

If the risk functional is the expectation, then—in view of Section 5.4—one has to consider the entropic size of the family $\mathcal{F} = \{F(x,\xi)\colon x \in \mathcal{X}\}$.

Suppose that the covering type of $\mathcal{F}$ is (v, V) and that the functions F are bounded by 1. Then by Theorem 20,

$$P\left\{\sqrt{N}\sup_{x\in\mathcal{X}}|\hat{f}_N(x)-f(x)|\geq K\right\}\leq\left(M(V)\,\frac{K}{\sqrt{v}}\right)^{v}\exp(-2K^2).$$

This gives a positive answer to the basic question (Q1) of Section 3, showing the blowing up function is $\sqrt{N}$ and the tail behavior is normal. If the functions $F(x,\xi)$ are unbounded, then Theorem 21 has to be used instead. The answers to questions (Q2), resp. (Q3), i.e., about the quality of the solution and the approximation of the arg min set, can then be found using Propositions 3 resp. 5.

In view of Theorem 7, a better estimate for the approximation of the arg min set can be found, if the entropic size of the family $\mathcal{F}^{\nabla}=\{(F(x,\xi)-F(y,\xi))/(\|x-y\|)\colon x,y\in\mathcal{X};\ x\neq y\}$ can be calculated.

If the risk functional $\mathbb{F}$ is not the expectation, but Lipschitz w.r.t. the distance $d_{\mathcal{H}}$, then the entropic sizes of the families $\mathcal{F}_{\mathcal{H}}=\{h\circ F(x,\xi)\colon x\in\mathcal{X};\ h\in\mathcal{H}\}$ resp. $\mathcal{F}_{\mathcal{H}}^{\nabla}=\{(h\circ F(x,\xi)-h\circ F(y,\xi))/(\|x-y\|)\colon x,y\in\mathcal{X};\ x\neq y;\ h\in\mathcal{H}\}$ have to be determined.

We illustrate the method of determining the entropic size of a stochastic program for the case $\mathbb{F}=\mathbb{E}$. It consists in studying: (1) the covering type of the feasible set $\mathcal{X}$, and (2) the structure of the mappings $x\mapsto F(x,\xi)$.

Assumption A1. The constraint set $\mathcal{X}$ is a compact set in $\mathbb{R}^d$, which has covering type (d, V). This means that it can be covered by $N_\varepsilon=\lfloor(\frac{V}{\varepsilon})^d\rfloor$ balls of radius ε, i.e., for every $\varepsilon>0$, there are balls $B(x_1,\varepsilon),\ldots,B(x_{N_\varepsilon},\varepsilon)$ with $x_i\in\mathcal{X}$ such that

$$\mathcal{X}\subseteq\bigcup_{i=1}^{N_\varepsilon}B(x_i,\varepsilon).$$

For a collection of square integrable functions $\mathcal{F}=\{F(\xi)\}$, let the diameter be defined as

$$\operatorname{diam}^2(\mathcal{F})=\mathbb{E}[\sup\{F\colon F\in\mathcal{F}\}-\inf\{F\colon F\in\mathcal{F}\}]^2.$$

The diameter is well defined only if sup and inf are measurable and square integrable.

A "Lipschitz condition" of order κ links the covering type of $\mathcal{X}$ with the covering type of $\mathcal{F}$.

Lemma 23. *Suppose that the Assumption A1 is fulfilled, i.e., $\mathcal{X}$ is of covering type (d, V), and that $F(x,\xi)$ satisfies for each $x_0\in\mathcal{X}$ and $\eta>0$*

$$\operatorname{diam}(\{F(x,\xi)\colon x\in B(x_0,\eta)\})\leq L\eta^{\kappa} \tag{6.1}$$

for some exponent $\kappa > 0$ and some constant L. Then $\mathcal{F} = \{F(x,\xi)\colon x \in \mathcal{X}\}$ is of covering type $(d/\kappa, V^\kappa L)$.

Proof. Let $\eta = (\varepsilon/L)^{1/\kappa}$. Cover $\mathcal{X}$ by $N_\eta = \lfloor \frac{V}{\eta} \rfloor^d$ balls $B(x_i,\eta)$ of radius η. By (6.1), the diameter of $\{F(x,\xi)\colon x \in B(x_0,\eta)\}$ is at most $L\eta^\kappa = \varepsilon$. The number of balls needed is

$$N_\eta = \left(\frac{V}{\eta}\right)^d = \left(\frac{V^\kappa L}{\varepsilon}\right)^{d/\kappa}. \qquad \square$$

Theorem 24. *Suppose that Assumption A1 holds and that the following conditions are fulfilled*

(i) *$x \mapsto F(x,\xi)$ is Lipschitz continuous with Lipschitz constant $L(\xi)$, i.e.,*

$$|F(x,\xi) - F(y,\xi)| \le L(\xi)\,\|x - y\|.$$

(ii) $\overline{L} = \sqrt{\mathbb{E}L^2(\xi)} < \infty$. Then $\mathcal{F} = \{F(x,\xi)\colon x \in \mathcal{X}\}$ is of covering type $(d, V\overline{L})$.

Proof. This result is a consequence of Lemma 23, since $\operatorname{diam}^2(\{F(x,\xi)\colon x \in B(x_0,\eta)\} \le \eta^2\mathbb{E}(L^2(\xi)) = \eta^2\overline{L}^2$. Thus, $\kappa = 1$ in (6.1). $\square$

6.1 Example: The entropic size of a linear recourse problem

As a typical example, consider a linear recourse problem, where the functions $F(x,\xi)$ are of the form

$$F(x,\xi) = \min\{q(\xi)^T y\colon W(\xi)y = b(x,\xi),\ y \ge 0\}$$

with $q(\xi) \in \mathbb{R}^m$, $W(\xi)$ a random $[k \times m]$ matrix and $b(x,\xi)$ a $\mathbb{R}^k$-valued random function.

We make the following assumption:

Assumption A2.

(i) There exists a measurable function $\tilde{u}\colon \Omega \to \mathbb{R}^m$ and a constant C_1 such that

$$\tilde{u}(\xi) \in \{u\colon W(\xi)^T u \le q(\xi)\} \subseteq \{u\colon \|u\| \le C_1\}.$$

(ii) The function $b\colon \mathcal{X} \times \Omega \to \mathbb{R}^k$ is differentiable w.r.t. x and satisfies $\|b(x,\xi)\| \le C_0$ a.s. and $\int \sup_{x\in X} \|\nabla b(x,\xi)\|^2 P(\mathrm{d}\xi) = C_2^2 < \infty$.

Theorem 25. *Let Assumptions* A1 *and* A2 *be fulfilled. Then*

$$\mathcal{F}_X = \{F(x, \xi) : x \in \mathcal{X}\} = \{\min\{q(\xi)^T y : W(\xi)y = b(x, \xi), y \geq 0\} : x \in \mathcal{X}\}$$

is of covering type (d, C_1C_2V).

Proof. By duality, we may write F as the solution of the dual program, i.e., as the maximum of a finite number of K functions $v_1, \ldots, v_K$.

$$F(x, \xi) = \max_{k=1,\ldots,K} b(x, \xi)^T v_k(\xi). \tag{6.2}$$

Here v_k are the vertices of the dual feasible polyhedron $\{u : W(\xi)^T u \leq q(\xi)\}$ and K is their maximal number. Since $\|b(x, \xi)\| \leq C_0$ and $\|v_k(\xi)\| \leq C_1$, we get

$$|F(x, \xi)| \leq C_0C_1.$$

Finally, for all x_0,

$$\begin{aligned} \operatorname{diam}^2\{F(x, \xi) : x \in B(x_0, \eta)\} &\leq \eta^2 \int C_1^2 \sup_{x \in X} \|\nabla b(x, \xi)\|^2 P(\mathrm{d}\xi) \\ &\leq \eta^2 C_1^2 C_2^2. \end{aligned} \tag{6.3}$$

Apply now Lemma 23. □

The most important special case is that of a linear $b(x, \omega)$:

$$b(x, \xi) = k(\xi) - T(\xi)x.$$

Since $\|\nabla b(x, \xi)\| = \|T(\xi)\|$, independent of x, the constant C_2 is here simply

$$C_2^2 = \int \|T(\xi)\|^2 P(\mathrm{d}\xi).$$

7 Epiconvergence

Up to now, the distance for measuring the approximation error of the objective functions was the supremum norm. The pertaining topology is uniform convergence on compacta. This setup is standard in statistical parameter estimation. However, the approach fails for stochastic programs with probabilistic constraints, i.e., constraints involving random variables. A conceptually sound way to deal with constraints and especially with stochastic

constraints is to consider the extended problem by introducing the *extended indicator function*

$$i_{\mathcal{X}}(x) = \begin{cases} 0 & \text{if } x \in \mathcal{X} \\ \infty & \text{if } x \notin \mathcal{X} \end{cases} \tag{7.1}$$

and to add $i_{\mathcal{X}}$ to the objective function as a penalty term. The two optimization problems

$$\text{(I)} \quad \min\{f(x): x \in \mathcal{X} \subseteq \mathbb{R}^d\}$$

and

$$\text{(II)} \quad \min\{f(x) + i_{\mathcal{X}}(x): x \in \mathbb{R}^d\}$$

are clearly equivalent. The problem (II) is only seemingly unconstrained. For algorithmic purposes, there is no difference between the formulations (I) and (II). Conceptually, there is, however, a difference: problem (II) has the advantage that is entirely formulated in terms of the objective extended real function. A disadvantage of (II) is that the notion of uniform convergence is not applicable. This is the main reason for introducing a new concept of convergence, especially suited for extended real functions: the concept of epiconvergence. The standard reference is the book by Rockafellar and Wets (1998). We give here a short introduction into the theory.

Recall that an extended real valued function f defined on $\mathbb{R}^d$ is called *lower semicontinuous* (lsc), if for all x and all sequences $x_N \to x$, it follows that $\liminf_N f(x_N) \geq f(x)$. f is lower semicontinuous if and only if it is the pointwise supremum of a countable family of continuous functions.

The epigraph of a function f defined on $\mathbb{R}^d$ is

$$\operatorname{epi} f = \{(x, \alpha): \alpha \geq f(x);\ x \in \mathbb{R}^d;\ \alpha \in \mathbb{R}\}.$$

A function f is lower semicontinuous if and only if $\operatorname{epi} f$ is a closed set in $\mathbb{R}^{d+1}$.

Let $B(x, r)$ be the closed ball with center x and radius r in an euclidean space, i.e., $B(x, r) = \{y: \|y - x\| \leq r\}$. The unit ball is denoted by $\mathbb{B} = B(0, 1)$.

Let $\mathcal{C}$ be the family of all closed sets in $\mathbb{R}^{d+1}$ (including the empty set $\emptyset$). We introduce the topology of setwise convergence. Let (C_N) be a sequence of closed sets in $\mathcal{C}$. We define the (topological) *limes superior* (ls) and *limes inferior* (li) as

$$\operatorname{li} C_N = \left\{x: \exists (x_N),\ x_N \in C_N,\ x = \lim_N x_N\right\}$$

$$\operatorname{ls} C_N = \left\{x: \exists (x_{N_k}),\ x_{N_k} \in C_{N_k},\ x = \lim_k x_{N_k}\right\}.$$

Obviously, li $C_N \subseteq$ ls C_N. The sequence (C_N) of closed sets converges setwise to C (in the sense of Painlevé–Kuratowski, in symbol $C_N \xrightarrow{\text{set}} C$), if

$$\text{li } C_N = \text{ls } C_N = C. \tag{7.2}$$

The topology of $\xrightarrow{\text{set}}$ convergence can be metrizised by the following metric:

$$\delta(C_1, C_2) = \sum_{k=1}^{\infty} 2^{-k} \frac{\mathbb{d}(C_1 \cap k\mathbb{B}, C_2 \cap k\mathbb{B})}{1 + \mathbb{d}(C_1 \cap k\mathbb{B}, C_2 \cap k\mathbb{B})}$$

where $\mathbb{B}$ is the unit ball in $\mathbb{R}^{d+1}$ and $\mathbb{d}$ is the Hausdorff metric (see (3.23)). The localization to the sets $k\mathbb{B}$ is necessary, because $\mathbb{d}(C_1, C_N)$ may be infinite for all N, even if $C_N \xrightarrow{\text{set}} C$. The family $\mathcal{C}$ of all closed sets $\mathbb{R}^{d+1}$ is a separable, metric space. A countable dense family of sets is given by all finite sets with points having rational coordinates.

A sequence of lsc functions (f_N) *epiconverges* to a lsc function f ($f_N \xrightarrow{\text{epi}} f$) if the following properties hold for all $x \in \mathbb{R}^d$

(i) for all sequences (x_N) with $x_N \to x$, $\liminf_N f_N(x_N) \geq f(x)$
(ii) there is a sequence $y_N \to x$ such that $\limsup_N f_N(y_N) \leq f(x)$.

Setwise convergence and epiconvergence of functions are closely related: f_N epiconverges to f if and only if $\text{epi} f_N \xrightarrow{\text{set}} \text{epi} f$. On the other hand, the sequence of closed sets A_N converges setwise to A if and only if the extended indicator functions i_{A_N} epiconverge to i_A.

The family of all l.s.c. extended real functions is a separable metric space under the topology of epiconvergence. A countable dense set of functions is given by functions of the form $f(x_i) = r_i$ for a finite set of rational points x_i and rational values r_i and $f(x) = \infty$ elsewhere.

Epiconvergence does not imply and is not implied by pointwise convergence.

Example 26. Consider the following sequence of functions

$$f_N^{(1)} = \begin{cases} -Nx & \text{if } 0 \leq x \leq 1/N \\ Nx - 2 & \text{if } 1/N \leq x \leq 2/N \\ 0 & \text{otherwise} \end{cases} \qquad f_N^{(2)} = \begin{cases} -Nx - 1 & \text{if } -1/N \leq x \leq 0 \\ -1 + Nx & \text{if } 0 \leq x \leq 1/N \\ 0 & \text{otherwise} \end{cases}$$

Both sequences epiconverge to $f(x) = -\mathbf{1}_{\{x=0\}}$. The first sequence converges pointwise to another limit, namely 0. The second sequence converges also pointwise, but not uniformly to $f(x)$. Uniform convergence of a sequence of lsc functions to a limit function implies, however, that the limit is lsc and that epiconvergence holds.

Most important is the relation between epiconvergence and the convergence of minima and argminima.

Lemma 27. *Let* $f_N \overset{\text{epi}}{\to} f$. *Then*

(i) $\text{ls}_N \arg\min f_N \subset \arg\min f$.

(ii) *If for all* $\varepsilon > 0$ *there is an index* N_ε *and a compact set* K_ε *such that* $\inf\{f_N(x): x \in K_\varepsilon\} \le \inf\{f_N(x): x \in \mathbb{R}^d\} + \varepsilon$ *for all* $N \ge N_\varepsilon$, *then*

$$\inf_x f_N(x) \to \inf_x f(x).$$

(iii) *Suppose that there is a compact set* K *such that* $\arg\min f_N \subset K$, *for sufficiently large* N. *Suppose further that* $\arg\min f = \{x^*\}$ *is a singleton. Then*

$$\text{ls}_N \arg\min f_N = \{x^*\},$$

i.e., for every selection $x_N \in \arg\min f_N$, $x_N \to x^*$ *as* $N \to \infty$.

Proof. (i) Let $x_{N_k} \in \arg\min f_{N_k}$ and suppose that $x_{N_k} \to x$. We have to show that $x \in \arg\min f$. By epiconvergence $\liminf_k f_{N_k}(x_{N_k}) \ge f(x)$ and by (7.3) $\limsup_k f_{N_k}(x_{N_k}) \le \inf_u f(u)$. Combining the two inequalities, we get $f(x) \le \inf_u f(u)$ and therefore $x \in \arg\min f$. (ii) Let x such that $f(x) < \inf_u f(u) + \varepsilon$. There is a sequence $(y_N) \to x$ with $\limsup_N f_N(y_N) \le f(x)$ and therefore $\limsup_N \inf_u f_N(u) \le \limsup_N f_N(y_N) \le f(x) \le \inf_u f(u) + \varepsilon$. Since ε was arbitrary,

$$\limsup_N \inf_u f_N(u) \le \inf_u f(u). \tag{7.3}$$

To show the opposite inequality we choose an ε and a sequence of points $(x_N) \in K_\varepsilon$ with the property that $f_N(x_N) \le \inf_u f_N(u) + 2\varepsilon$. We can choose a subsequence (x_{N_k}) which converges to some $x \in K_\varepsilon$. By epiconvergence, $\liminf_N \inf_u f_N(u) \ge \liminf_N f_N(x_N) - 2\varepsilon \ge f(x) - 2\varepsilon \ge \inf_u f(u) - 2\varepsilon$. Since ε is arbitrary, (ii) follows. (iii) Every selection $x_N \in \arg\min f_N$ has cluster points, because of compactness. But every cluster point must be in $\arg\min f$, i.e., must be equal to x^*. □

The inclusion in (i) may be strict as is shown by the following example:

$$f_N(x) = \begin{cases} x/N & \text{if } 0 \le x \le 1 \\ 1 & \text{otherwise} \end{cases}.$$

f_N epiconverges to

$$f(x) = \begin{cases} 0 & \text{if } 0 \le x \le 1 \\ 1 & \text{otherwise} \end{cases},$$

but $\arg\min f_N = \{0\}$ does not converge to $\arg\min f = [0, 1]$.

However, if we introduce the notion of ε-arg min

$$\varepsilon\text{-arg min } f := \left\{x: f(x) \leq \inf_u f(u) + \varepsilon\right\},$$

the following lemma can be shown.

Lemma 28. *Suppose that* $f_N \xrightarrow{\text{epi}} f$ *and that* $\inf_x f_N(x) \to \inf_x f(x)$. *Then there is a sequence* $\varepsilon_N \to 0$ *such that* ε_N-arg min $f_N \to f$. *Conversely, if* $\varepsilon_N \to 0$ *and* ε_N-arg min $f_N \to f$ *then* $\inf_x f_N(x) \to \inf_x f(x)$.

Proof. See Rockafellar and Wets (1998), Theorem 7.31. □

Besides this qualitative statement, a quantitative statement of the relation between the epidistance of functions and the distance of the arg mins is highly desirable. The following Theorem 29 links the epidistance of functions with the distance of ε-arg mins. Such a theorem can be used to formulate an analogue to Proposition 5 (which was formulated for sup-norms) within the epiconvergence setup.

For the formulation of the result, we introduce two new distances for functions:

$$\hat{\mathbb{d}}(f, g) = \mathbb{d}(\text{epi } f, \text{epi } g)$$

where $\mathbb{d}$ is the Hausdorff distance introduced in (3.23) and a variant

$$\hat{\mathbb{d}}^+(f, g) = \inf\{\eta: f(x) \geq \inf g_{y \in B(x,\eta)} - \eta;\ g(x) \geq \inf f_{y \in B(x,\eta)} - \eta \text{ for all } x\}.$$

We have that

$$\hat{\mathbb{d}}^+(f, g) \leq \hat{\mathbb{d}}(f, g) \leq \sqrt{2}\hat{\mathbb{d}}(f, g)$$

(see Rockafellar and Wets, 1998, Proposition 7.61, p. 284).

Remark. The original definition of $\hat{\mathbb{d}}$ and $\hat{\mathbb{d}}^+$ in Rockafellar and Wets (1998) includes the restriction to a ball $\rho\mathbb{B}$. In some applications, this restriction may be necessary. The concept and idea, however, may be demonstrated without this restriction.

If two functions f and g are close to each other w.r.t. distance $\hat{\mathbb{d}}^+$, both their min's and their ε-arg min's are close to each other according to the following theorem.

Theorem 29. *Suppose that f and g are lsc convex functions on a compact set χ in $\mathbb{R}^d$. Suppose further that $\mathbb{d}(0, 3\varepsilon\text{-arg}\min f) \le \rho$, $\mathbb{d}(0, 3\varepsilon\text{-arg}\min g) \le \rho$ for some constant ρ. Then $\hat{\mathbb{d}}^+(f,g) = \eta$ and $\eta < \varepsilon$ implies that*

(i) $|\min f - \min g| \le \eta$
(ii) $\hat{\mathbb{d}}(\varepsilon\text{-arg}\min f, \varepsilon\text{-arg}\min g) \le \eta + \frac{2\eta}{\varepsilon+2\eta}\, 2\rho$.

Proof. The proof is a modification of the proof of Theorem 7.69, p. 291 in Rockafellar and Wets (1998). Let $x^+ \in \arg\min f$. By the definition of $\hat{\mathbb{d}}^+(f,g) < \eta$, $\min g \le \min_{y \in B(x^+,\eta)} g(y) \le f(x^+) + \eta = \min f + \eta$. Exchanging f and g, wet get (i). Next we show that

$$\varepsilon\text{-arg min } f \subseteq (\varepsilon + 2\eta)\text{-arg min } g.$$

Suppose that $x \in \varepsilon\text{-arg}\min f$. Let $y \in B(x,\eta)$ such that $g(y) \le f(x) + \eta$. Then $g(y) \le \min f + \varepsilon + \eta \le \min g + \varepsilon + 2\eta$. Finally, we show that

$$(\varepsilon + 2\eta)\text{-arg min } g \subseteq (\varepsilon\text{-arg min } g) + \left[\frac{2\eta}{\varepsilon + 2\eta}\, 2\rho\right]\mathbb{B}. \tag{7.4}$$

Let $\alpha_0 = \min g$, $\alpha_1 = \min g + \varepsilon$, $\alpha_2 = \min g + \varepsilon + 2\eta$. Let $x \in (\varepsilon + 2\eta)\text{-arg}\min g$, i.e., $g(x) \le \alpha_2$. and let $x^* \in \arg\min g$. Let $y = (1-\tau)x + \tau x^*$, where $\tau = (\alpha_2 - \alpha_1)/(\alpha_2 - \alpha_0) = 2\eta/(\varepsilon + 2\eta)$. Then, by convexity, $g(y) \le (1-\tau)\alpha_2 + \tau\alpha_0 = \alpha_1$. Thus, $y \in \varepsilon\text{-arg}\min g$. Since $\|x - y\| = \tau\|x - x^*\| \le \tau 2\mathbb{d}(0,\ 3\varepsilon\text{-arg}\min g) \le \tau 2\rho$, (7.4) is established. Putting the pieces together, we have shown that

$$\varepsilon\text{-arg min } f \subseteq \varepsilon\text{-arg min } g + \left(\eta + \frac{2\eta}{\varepsilon + 2\eta}\, 2\rho\right)\mathbb{B}$$

which is equivalent to $\mathbb{d}(\varepsilon\text{-arg}\min f \,\|\, \varepsilon\text{-arg}\min g) \le \eta + \frac{2\eta}{\varepsilon+2\eta}\, 2\rho$. By exchanging the roles of f and g we get the assertion (ii). □

For the application of the concept of epiconvergence in statistics and stochastic optimization it has to be extended to stochastic processes. A stochastic process $(x,\omega) \mapsto Z(x,\omega)$ is called a *random lower semicontinuous* function if it is jointly measurable and for all $\omega \in \Omega$, $x \mapsto Z(x,\omega)$ is lower semicontinuous.

If Z is a random lsc process on $\mathbb{R}^d$, then epi Z induces a probability measure on the Borel sets of the metric space $(\mathcal{C}^{d+1}, \delta)$. For a sequence (Z_N) of random lsc processes, the notion of a.s. convergence in the epi sense is well defined: Z_N epiconverges to Z almost surely, if

$$\mathbb{P}\{Z_N \overset{\text{epi}}{\to} Z\} = 1.$$

Of course, also the notion of convergence in distribution is defined: (Z_N) epiconverges in distribution to Z (in symbol: $Z_N \overset{\text{epi-}\mathcal{D}}{\to} Z$), if the probability distributions induced by epi Z_N on the separable metric space $(\mathcal{C}^{d+1}, \delta)$ weakly converge to that induced by epi Z.

An equivalent formulation of the epiconvergence in distribution is given by the following theorem.

Theorem 30. *Let $Z_N(x, \cdot)$ be a l.s.c. process. Then Z_N epiconverges in distribution ($Z_N \overset{\text{epi-}\mathcal{D}}{\to} Z$) if and only if for all k, all collections of closed rectangles $R_1, \ldots, R_k$ and all reals $\alpha_1, \ldots, \alpha_k$*

$$
\begin{aligned}
&P\left(\inf_{x\in R_1} Z(x,\cdot) > \alpha_1, \ldots, \inf_{x\in R_k} Z(x,\cdot) > \alpha_k\right)\\
&\le \liminf_N P\left(\inf_{x\in R_1} Z_N(x,\cdot) > \alpha_1, \ldots, \inf_{x\in R_k} Z_N(x,\cdot) > \alpha_k\right)\\
&\le \limsup_N P\left(\inf_{x\in R_1^o} Z_N(x,\cdot) \ge \alpha_1, \ldots, \inf_{x\in R_k^o} Z_N(x,\cdot) \ge \alpha_k\right)\\
&\le P\left(\inf_{x\in R_1^o} Z(x,\cdot) \ge \alpha_1, \ldots, \inf_{x\in R_k^o} Z(x,\cdot) \ge \alpha_k\right).
\end{aligned}
$$

Proof. See Salinetti and Wets (1986) and Pflug (1992). □

8 Epipointwise convergence for stochastic programs

In this section we consider a stochastic program with random constraints

$$\operatorname*{Min}_{x\in\mathcal{X}\cap y} \{ f(x) = \mathbb{E}_{\mathbb{P}}[F(x,\xi)]\}$$

where

$$\mathcal{Y} = \{x\colon \mathbb{E}[F_1(x,\xi)] \le b_1, \ldots, \mathbb{E}[F_J(x,\xi)] \le b_J\}.$$

The empirical counterpart of f is $\hat{f}_N(x) = \frac{1}{N}\sum_{n=1}^{\infty} F(x,\xi_n)$ and the empirical counterpart of the constraint set $\mathcal{Y}$ is

$$\hat{\mathcal{Y}}_N = \left\{x\colon \frac{1}{N}\sum_{n=1}^{N} F_1(x,\xi_n) \le b_1 + \frac{\log N}{\sqrt{N}}, \ldots, \frac{1}{N}\sum_{n=1}^{N} F_J(x,\xi_n) \le b_J + \frac{\log N}{\sqrt{N}}\right\}.$$

The extra term $\frac{\log N}{\sqrt{N}}$ is added for consistency reasons. Consider the extended function

$$f^e(x) = f(x) + i_{y(x)}$$

and its empirical counterpart

$$\hat{f}_N^e(x) = \hat{f}_N(x) + i_{\hat{y}_N}(x)$$

where i, the extended indicator function, is given by (7.1).

$\hat{f}_N^e(x)$ converges a.s. pointwise to $f^e(x)$, but not uniformly. Epiconvergence holds often in addition to that. The combination of pointwise convergence and epiconvergence leads to the following definition.

Definition 31. The extended real functions f_N converges epipointwise to f (in symbol: $f_N \overset{\text{epipw}}{\to} f$), if

(i) $x_N \to x$ implies that $\liminf_N f_N(x_N) \geq f(x)$
(ii) $f_N(x) \to f(x)$ for all x.

Epipointwise convergence is a stronger concept than just epiconvergence: take the functions $f_N^{(1)}$ and $f_N^{(2)}$ of Example 26. While $f_N^{(2)}$ converge to $f(x) = -1_{\{x=0\}}$ in the epipointwise sense, the functions $f_N^{(1)}$ do not converge in this sense. On the other hand, epipointwise convergence is weaker than uniform convergence. There is an equivalent definition for epipointwise convergence.

Lemma 32. *For lsc functions,* $f_N \overset{\text{epipw}}{\to} f$ *if and only if*

$$\lim \inf_{x \in R} f_N(x) = \inf_{x \in R} f(x)$$

for all compact rectangles R.

Proof. Suppose that $f_N \overset{\text{epipw}}{\to} f$. Let $x \in R$ such hat $f(x^*) = \inf_{x \in R} f(x)$. Then, since $f_N(x^*) \to f(x^*)$, $\limsup \inf_{x \in R} f_N(x) \leq \inf_{x \in R} f(x)$. Together with $\limsup \inf_{x \in R} f_N(x) \geq \inf_{x \in R} f(x)$ this implies the assertion. Conversely, let $\liminf_{x \in R} f_N(x) = \inf_{x \in R} f(x)$ for all compact R. By setting $R = \{x\}$ pointwise convergence follows. Let $x_N \to x$. Suppose that $\liminf f_N(x_N) = f(x) - \varepsilon$, where $\varepsilon > 0$. We may find a compact rectangle R such that x lies in its interior and such that $\inf_{y \in R} f(y) \geq f(x) - \varepsilon/2$. By assumption, $f(x) - \varepsilon = \liminf f_N(x_N) \geq \lim_{y \in R} \inf f_N(y) \inf_{y \in R} f(y) \geq f(x) - \varepsilon/2$. This contradiction proves the lemma. □

It is in general not true that $f_N \overset{\text{epi}}{\to} f$ and $g_N \overset{\text{epi}}{\to} g$ implies that $f_N + g_N \overset{\text{epi}}{\to} f + g$ (as a counter example, take the functions $f_N = f_N^{(2)}$ of Example 26 and $g_N = -f_N$. Then because of the asymmetry of the epiconvergence, $f_N + g_N \equiv 0$, but $f + g = -1_{\{x=0\}}$). However, epipointwise convergence is compatible with sums.

Lemma 33. *If* $f_N \overset{\text{epipw}}{\to} f$ *and* $g_N \overset{\text{epipw}}{\to} g$, *then* $f_N + g_N \overset{\text{epipw}}{\to} f + g$.

Proof. Since $\liminf_N f_N(x_N) + g_N(x_N) \geq \liminf_N f_N(x_N) + \liminf_N g_N(x_N)$, the assertion follows from the definition of epipointwise convergence (Definition 31). □

As an important consequence of this lemma one sees that for the epipointwise convergence of $\hat{f}_N^e$ one may consider the two parts $\hat{f}_N(x)$ and $i_{\hat{\mathcal{Y}}_N}(x)$ separately.

Assume that the functions $F(x, \xi)$ and $F_j(x, \xi)$, $j = 1, \ldots, J$ have for $x \in \mathcal{X}$ the uniform boundedness property (see Definition 6). Suppose that the constraint set $\mathcal{Y}$ fulfills the following condition:

Condition (B). There is a function $\varepsilon(\eta) > 0$, for $\eta > 0$ such that the following implications hold: if $x \in \mathcal{Y}$, then $\inf_{y \in B(x,\eta)} f_j(x) < b_j - \varepsilon(\eta)$ for at least one $j \in [1, \ldots, J]$. If $B(x, \eta) \cap \mathcal{Y} = \emptyset$, then $f_j(x) > b_j + \varepsilon(\eta)$ for at least one $j \in [1, \ldots, J]$.

Theorem 34. *Suppose that* $\hat{f}_N$ *converges a.s. uniformly to* f *and that* $\hat{f}_{N,j}$ *converge a.s. uniformly to* f_j *for all* j. *Suppose further that* $\sup_x \mathbb{V}\text{ar}(F_j(x, \xi)) = \sigma < \infty$. *Then, under Condition* (B), $\hat{f}_N + i_{\hat{\mathcal{Y}}_N}$ *converges a.s. epipointwise to* $f + i_{\mathcal{Y}}$.

Proof. In view of Lemma 33, we have only to show that $i_{\hat{\mathcal{Y}}_N}$ converges a.s. epipointwise to $i_{\mathcal{Y}}$. We show first the pointwise convergence. If $x \notin \mathcal{Y}$, then there is a j such that $f_j(x) > b_j$ for some j. Therefore, since $\hat{f}_{N,j}(x) \to f_j(x)$ a.s., for every ω there is an $N_0(\omega)$ such that $\hat{f}_{N,j} > b_j + (\log N/\sqrt{N})$ for $N \geq N_0$. Consequently, $i_{\hat{\mathcal{Y}}_N} \geq i_{\{\hat{f}_{N,j}(x) \leq b_j + (\log N)/\sqrt{N}\}} \to \infty = i_{\{f_j(x) \leq b_j\}}$. If $x \in \mathcal{Y}$, then by the LIL (Theorem 11) $\hat{f}_{N,j}(x) \leq b_j + 2\sigma(\log\log N)/\sqrt{N}$ a.s., where σ is the standard deviation of $F_j(x, \xi)$. Therefore, also in this case $i_{\{\hat{f}_{N,j}(x) \leq b_j + (\log N)/\sqrt{N}\}} \to 0 = i_{\{f_j(x) \leq b_j\}}$.

For the epiconvergence property, we have to show that for all x and all $\eta > 0$,

$$\mathbb{P}\left\{\sup_x \inf_{y \in B(x,\eta)} i_{\hat{\mathcal{Y}}_N}(y) - i_{\mathcal{Y}} \geq \eta\right\} \to 0$$

and

$$\mathbb{P}\left\{\inf_x \inf_{y \in B(x,\eta)} i_{\mathcal{Y}}(y) - i_{\hat{\mathcal{Y}}_N}(x) \geq \eta\right\} \to 0.$$

Since we assume that $\infty - \infty = 0$, this is equivalent to

$$\mathbb{P}\{\exists x \in \mathcal{Y}\text{: for all } y \in B(x,\eta)\text{: } y \notin \hat{\mathcal{Y}}_N\} \to 0$$
$$\mathbb{P}\{\exists x \in \hat{\mathcal{Y}}_N\text{: for all } y \in B(x,\eta)\text{: } y \notin \mathcal{Y}\} \to 0$$

By Assumption (B), if $x \in \mathcal{Y}$, then there is a j and some $y \in B(x,\eta)$ such that $f_j(y) < b_j - \varepsilon(\eta)$. But this implies that $\mathbb{P}\{f_{N,j}(y) > b_j + (\log N)/\sqrt{N}\} \to 0$. On the other hand, if for all $y \in B(x,\eta)$, $y \notin \mathcal{Y}$, then there is a j such that $f(x) > b_j + \varepsilon(\eta)$, which implies that $\mathbb{P}\{\hat{f}_{j,N}(x) \le b_j + (\log N)/\sqrt{N}\} \to 0$. □

9 Asymptotic stochastic programs

Consider a stochastic program in the extended form

$$\operatorname*{Min}_{x\in\mathbb{R}^d}\{ f^e(x) = \mathbb{E}[F(x,\xi)] + i_{\mathcal{X}}\} \tag{9.1}$$

where $i_{\mathcal{X}}$ the extended indicator function of the constraint set $\mathcal{X}$, is defined as in (7.1). We assume that the constraint set $\mathcal{X}$ is closed, convex and nonrandom. Program (9.1) has the empirical counterpart

$$\operatorname*{Min}_{x\in\mathbb{R}^d}\left\{\hat{f}^e_N(x) = \frac{1}{N}\sum_{n=1}^{N} F(x,\xi_n) + i_{\mathcal{X}}\right\}. \tag{9.2}$$

Let v be the minimal value of (9.1) and $\hat{v}_N$ the (random) minimal value of (9.2). Any measurable selection of the empirical arg min set $\hat{x}_N \in \hat{\mathcal{X}}^*_N$, where

$$\hat{\mathcal{X}}^*_N = \operatorname*{arg\,min}_{x\in\mathbb{R}^d}\left\{\frac{1}{N}\sum_{n=1}^{N} F(x,\xi_n) + i_{\mathcal{X}}\right\}$$

is called an *empirical solution*.

We study first the asymptotic behavior of the empirical minima $\hat{v}_n$ and later the asymptotic behavior of the empirical arg minima $\hat{x}_N$.

The basic observation is: if $\hat{f}^e_N(\cdot)$ converges (a.s. or in probability) to $f^e(\cdot)$ in some function space and the min-operator $f \mapsto M(f) = \min_x f(x)$ is **continuous** in this function space, then $\hat{v}_N$ converges (a.s. or in probability) to v.

A more refined analysis considers the errors in more detail: suppose that

$$Y_N(x) = \beta(N)(\hat{f}^e_N(x) - f^e(x)) \tag{9.3}$$

converges in distribution to a nondegenerate limit process Y in a function space. If the min-operator is **differentiable** in this function space, then $\beta(N)(\hat{v}_N - v)$ converges in distribution to a limit law. A theorem of this kind is stated below.

Theorem 35. *Suppose that* $\beta(N)(\hat{f}_N^e(x) - f^e(x))$ *converges in the Banach space* $C(\mathcal{X})$ *of continuous functions to a limiting distribution* Y. *If* $\mathcal{X}$ *is compact, then*

$$\beta(N)(\hat{v}_N - v) \quad \text{converges in law to} \quad \min_{x \in \mathcal{X}^*} Y(x)$$

where $\mathcal{X}^*$ *is the set of solutions of* (9.1).

Proof. The "delta" method of proof was already mentioned (see Shapiro, 1993): if M is an (at least directionally (Hadamard-) differentiable operator in $C(\mathcal{X})$ and Y_N given by (9.3) converges in law to Y, then $\beta(N)[M(\hat{f}_N^e(x)) - M(f^e(x))]$ converges in law to $M'(M(f^e(x)), Y)$. Here $M'(f, g) = \lim_{h \downarrow 0} \frac{1}{h}[M(f + hg) - M(f)]$. It is easy to see that the directional derivative of the min-operator M in $C(\mathcal{X})$ is

$$M'(f, g) = \min\{g(x) \colon x \in \arg\min f\}.$$

A more direct proof (which is also applicable, if the function space is not a Banach space) shows that the asymptotic distribution of $\beta(N)[\min_{x \in \mathcal{X}} \hat{f}_N^e(x) - \min_{x \in \mathcal{X}} f^e(x)]$ is the same as the asymptotic distribution of $\beta(N)[\min_{x \in \mathcal{X}^*} \hat{f}_N^e(x) - \min_{x \in \mathcal{X}^*} f^e(x)] = \min_{x \in \mathcal{X}^*} \beta(N)[\hat{f}_N^e(x) - f^e(x)]$. The latter expression has asymptotic distribution $\min_{x \in \mathcal{X}^*} Y(x)$. □

For studying the limiting distribution of the solutions we suppose that (9.1) has a unique solution

$$x^* = \arg\min_{x \in \mathbb{R}^d} \{\mathbb{E}[F(x, \xi)] + i_{\mathcal{X}}\}, \tag{9.4}$$

and that $\hat{x}_N$ is a measurable selection from $\hat{\mathcal{X}}_N^*$.

The goal is to find the limiting distribution of

$$\Gamma_N^{-1}(\hat{x}_N - x^*),$$

where Γ_N is an appropriately chosen sequence of regular matrices.

The basic equation is

$$\Gamma_N^{-1}(\arg\min_x \hat{f}_N^e(x) - x^*) = \arg\min_t \hat{f}_N^e(x^* + \Gamma_N t). \tag{9.5}$$

Here t is called the *local coordinate* around the solution x^*. The left hand side of (9.5) appears as the arg min of a stochastic program in local coordinates. Evidently, the right hand side may be written as

$$\begin{aligned}\arg\min_t \hat{f}_N^e(x^* + \Gamma_N t) &= \arg\min_t \rho_N[\hat{f}_N^e(x^* + \Gamma_N t) - f(x^*)] \\ &= \arg\min Z_N(t) \text{ (say)}\end{aligned}$$

where $\rho_N > 0$ is arbitrary. Now, suppose that ρ_N and Γ_N are chosen such that the extended real stochastic process $Z_N(\cdot) = \rho_N[\hat{f}_N^e(x^* + \Gamma_N t) - f(x^*)]$ converges to a limiting process $Z(\cdot)$. Then, using a **M-theorem** (Minimizer theorem) of the form

$$\lim_N \arg\min_t Z_N(t) = \arg\min_t \lim_N Z_N(t) \tag{9.7}$$

we may identify the limiting distribution of $\Gamma_N^{-1}(\hat{x}_N - x^*)$ as the distribution of $\arg\min_t Z(t)$. A useful M-theorem will be given below (see Theorem 37).

Under rather general assumptions $Z_N(\cdot)$ converges in distribution to a limiting process

$$Z(\cdot) = D(\cdot) + S(\cdot) + i_{\mathcal{X}^+}(\cdot), \tag{9.8}$$

where $D(t)$ is a regularly varying deterministic function, $S(t)$ is a zero mean stochastic process on $\mathbb{R}^d$, which by the classical result of P. Lévy, must be infinitely divisible, and $\mathcal{X}^+$ is some closed convex set.

To be more precise about the type of convergence, we assume that

$$\rho_N \sum_{n=1}^{N} [F(x^* + \Gamma_N t, \xi_n) - F(x^*, \xi_n)] \quad \text{converges to} \quad D(t) + S(t)$$

in distribution in the sup-norm sense, whereas

$$i_{\mathcal{X}}(x^* + \Gamma_N t) \quad \text{converges to} \quad i_{\mathcal{X}^+}(t)$$

in distribution in the epi-sense (which is equivalent to setwise convergence of the constraint sets).

The sum then converges also in the epi-sense. We consider the process Z_N as a random element in the separable metric space of lsc functions on $\mathbb{R}^d$ (see Section 7).

Definition 36. Suppose that $Z_N(\cdot)$ defined by (9.6) epiconverges in distribution to $Z(\cdot)$ given by (9.8). Then the stochastic program

$$\operatorname*{Min}_t Z(t) = \operatorname*{Min}_{t\in\mathcal{X}^+}\{D(t) + S(t)\} \tag{9.9}$$

is called the *asymptotic stochastic program* associated to (9.1).

Notice that $S(t)$ is a stochastic process and hence the arg min of (9.9) is a random variable or a random set.

Now we make the cited M-theorem (9.7) precise.

Theorem (M-Theorem) 37. *Suppose that the l.s.c. processes $Z_N(\cdot)$ epiconverge in distribution to a l.s.c. limiting process $Z(\cdot)$. Suppose further that these processes satisfy the assumptions of Lemma 27 (iii), i.e., there is a (random) compact set $K(\omega)$ such that* $\arg\min Z_N(\omega) \subseteq K(\omega)$ *a.s. and that* $\arg\min Z$ *is a.s. a singleton. Then any measurable selection from* $\arg\min_t Z_N(t)$ *converges in distribution to* $\arg\min_t Z(t)$.

Proof. Using the Shorohod–Wichura–Dudley theorem (see Shorack and Wellner, 1986, p. 47), one may construct, on an appropriate probability space, versions $\tilde{Z}_N$ and $\tilde{Z}$ of the processes Z_N and Z, which converge almost surely in the episense. These versions satisfy for almost all ω, the assumptions of Lemma 27(iii). Thus, all measurable selections from $\arg\min \tilde{Z}_N$ converge to $\arg\min \tilde{Z}$ almost surely. A fortiori, convergence in distribution of any selection from $\arg\min Z_N$ to the distribution of $\arg\min Z$ holds. □

The technique to prove the asymptotics of a stochastic program consists therefore in two steps

- identification of the asymptotic stochastic program,
- calculation of the arg min distribution of the asymptotic program.

There are some classes of stochastic processes $Z(\cdot)$ which may occur as limits. In principle, by letting ξ be a degenerate random variable, one sees that the class of limiting stochastic programs is at least as large as the class of deterministic programs. However, by a general result due to P. Lévy, a nondegenerate limiting process appearing in (9.8) must be a stable process. The simplest stable process is the normal shift process, i.e., $S(t) = t^T Y$, where $Y \sim N(0, \Sigma)$. Here t^T denotes the transposition. Examples for other stable processes appearing as limits of stochastic programs, like Wiener type processes and Poisson type processes were given by Pflug (1995).

We consider here only the smooth case, which leads to normal shift limit process. The other mentioned limit processes may appear only for cost functions having jumps.

There is a distinction to be made between the two cases

- The optimum x^* lies in the interior of the feasible set $\mathcal{X}$, no constraints appear in the limiting stochastic program. This is case 1.
- The optimum x^* lies on the boundary of $\mathcal{X}$. This is case 2.

9.1 Case 1: No constraints in the limit

The basic assumption concerns the local behavior of $x \mapsto F(x, \xi)$ near the optimum x^*. We introduce the following set of assumptions:

Condition C1. $F(x, \xi)$ is a smooth function in the sense that it allows an expansion

$$F(x^* + t, \xi) = F(x^*, \xi) + t^T[k(\xi)] + t^T Q(\xi)t + r(t, \xi)$$

such that $|r(t, \xi)| = \overline{r}(\xi)o(\|t\|^2)$ for a function $\overline{r}$ satisfying $\mathbb{E}(\overline{r}(\xi)) < \infty$. All entries of the vector $k(\xi)$ are square integrable and all entries of the matrix $Q(\xi)$ are integrable.

Theorem 38. *Let Assumption C1 be satisfied and suppose that x^* lies in the interior of $\mathcal{X}$. Then with $\Gamma_N = \frac{1}{\sqrt{N}} I$*

$$Z_N(t) \to Z(t) = t^T Y + t^T A t$$

in distribution, uniformly on compact sets, where $A = \mathbb{E}(Q(\xi))$ and $Y \sim N(0, \Sigma)$ with $\Sigma = \mathbb{E}(k(\xi) \cdot k(\xi)^T)$.

Proof. Expand $Z_N(t)$ as

$$\begin{aligned} Z_N(t) &= \sum_{n=1}^{N} F\left(x^* + t/\sqrt{N}, \xi_n\right) - F(x^*, \xi_n) \\ &= t/\sqrt{N} \sum_{n=1}^{N} k(\xi_n) + 1/N \sum_{n=1}^{N} t^T Q(\xi_i)t + \sum_{n=1}^{N} r\left(x^* + t/\sqrt{N}, \xi_n\right) \\ &= t^T Y_N + t^T A_N t + r_N(t) \text{ (say).} \end{aligned}$$

Here $Y_N = N^{-1/2} \sum_{n=1}^{N} k(\xi_n)$ converges by the CLT in distribution to a normal distribution $Y \sim N(0, \Sigma)$ with $\Sigma = \mathbb{E}(k \cdot k^T)$ and hence $t^T Y_N \to t^T Y$. (Notice that $\mathbb{E}(k(\xi)) = 0$ because of the optimality condition). By the Strong

Law of Large Numbers (9) $1/N \sum_{n=1}^{N} t^T Q(\xi_n) t \to t^T \mathbb{E}(Q(\xi)) t = t^T A t$ almost surely. By Assumption C1, for all T,

$$\sup_{\|t\| \le T} |r_N(t)| = \sup_{\|t\| \le T} \left| \sum_{n=1}^{N} r\left(x^* + t/\sqrt{N}, \xi_n\right) \right|$$

$$= o(1)\|T\|^2 \, 1/N \sum_{n=1}^{N} \bar{r}(\xi_n) \to 0 \quad \text{a.s.} \qquad \square$$

The arg min distribution of the process Z appearing in Theorem 38 is normal, since

$$\arg\min\{t^T Y + t^T A t : t \in \mathbb{R}^d\} = \frac{1}{2} A^{-1} Y \sim N\left(0, \frac{1}{4}[A^{-1}]^T \Sigma A^{-1}\right).$$

9.2 *Case 2: The limiting arg min lies on the boundary of the constraint set*

Suppose that the solution x^* lies on the boundary of the convex set $\mathcal{X} \subseteq \mathbb{R}^d$. Let $\bar{k}$ be the gradient of the objective function at point x^*. Necessarily, $-\bar{k}$ points outwards $\mathcal{X}$. By a possible translation and rotation of the parameter space $\mathcal{X}$, we may standardize the problem such that x^* is the origin and $\bar{k}$ is a positive multiple of the first unit vector e_1.

Suppose therefore w.l.o.g. that $x^* = 0$ and $\bar{k}/\|k\| = e_1$. The asymptotics depend on the *local curvature* of the convex set $\mathcal{X}$ near its boundary point 0 as will be defined now.

Define the function $\lambda_{\mathcal{X}}$ on the sphere $S := \{s : \|s\| = 1\}$ of $\mathbb{R}^d$ with values in $[0, \infty]$ as

$$\lambda_{\mathcal{X}}(v) = \sup\{\gamma : \gamma \cdot v \in \mathcal{X}\}. \tag{9.10}$$

$\lambda_{\mathcal{X}}(v)$ characterizes $\mathcal{X}$ completely. The tangential cone $\mathcal{X}_0$ of $\mathcal{X}$ at 0 is the closure of $\{v : \lambda_{\mathcal{X}}(v) > 0\}$. Recall that a point x is extremal in a convex set $\mathcal{Z}$, if x cannot be represented as $x = \frac{1}{2} z_1 + \frac{1}{2} z_2$, where $z_i \in \mathcal{Z}$ and $z_1 \neq z_2$.

Assumption C2. Either 0 is extremal in the tangential cone $\mathcal{X}_0$, in which case we set $\alpha = 1$ (see Figs. 3 and 4), or there is an $\alpha > 1$ such that for all $s \perp e_1$

$$\lambda_{\mathcal{X}}(s + \delta e_1) = c(s)\delta^{1/(\alpha-1)}(1 + o(1)) \tag{9.11}$$

where $o(1) \to 0$ uniformly in s as $\delta \to 0$ and $s \mapsto c(s)$ is upper semicontinuous. The constant α is called the degree of curvature of $\mathcal{X}$ at 0.

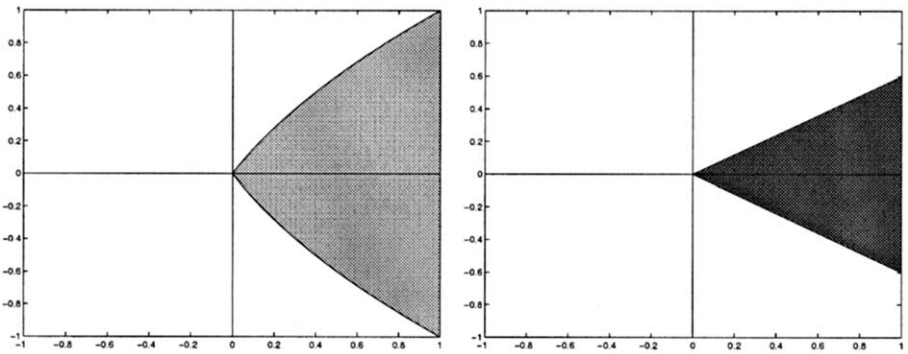

Fig. 3. A convex set X (left) and its asymptotic set X* (right) expanded at the origin using the blow-up matrices Γ_N with $\gamma_N = N^{-1}$, $\mu_N = N^{-1}$. Here the local curvature is $\alpha = 1$. The asymptotic set is a cone.

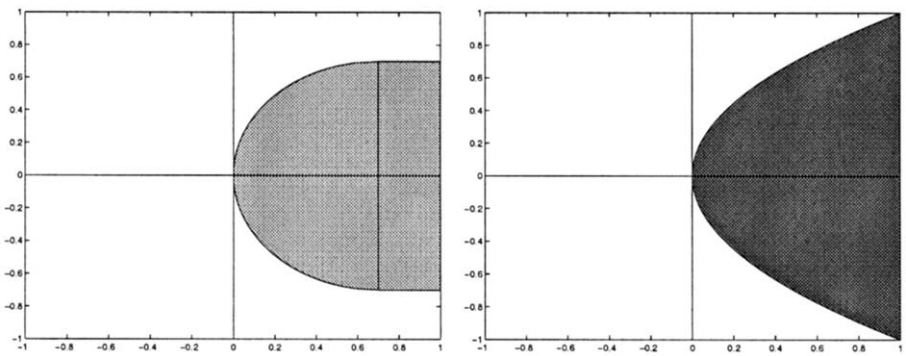

Fig. 4. A convex set X (left) and its asymptotic set X* (right) expanded at the origin using the blow-up matrices Γ_N with $\gamma_N = N^{-1}$, $\mu_N = N^{-1/2}$. Here the local curvature is $\alpha = 2$. The asymptotic set is a parabola.

The blow-up matrices will be chosen as

$$\Gamma_N = \begin{pmatrix} \gamma_N & 0 & \cdots & 0 \\ 0 & \mu_N & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ddots & \mu_N \end{pmatrix} \tag{9.12}$$

Proposition 39. *Let Assumption C2 be satisfied with* $\alpha > 1$. *If* $\mu_N \to 0$ *and* $\gamma_N \mu_N^{-\alpha} \to 1$ *then with* Γ_N *given by* (9.12) $\mathcal{X}_N = \Gamma_N^{-1}\mathcal{X} \overset{\text{set}}{\to} \mathcal{X}^*$, *where*

$$\mathcal{X}^* = \{\xi s + \eta e_1 : s \perp e_1;\ \xi \geq c(s) \cdot \eta^\alpha;\ \xi, \eta \geq 0\}. \tag{9.13}$$

If, however, $\gamma_N \mu_N^{-\alpha} \to \infty$ *then* $\mathcal{X}_N = \Gamma_N^{-1} \mathcal{X} \overset{\text{set}}{\to} \mathcal{X}^0$, *where* $\mathcal{X}^0$ *is the tangential cone*

$$\mathcal{X}^0 = \{\xi s + \eta e_1 : s \perp e_1;\ \xi, \eta \geq 0\} \tag{9.14}$$

which is a half-space.

Proof. See Pflug (1995). □

The situation is illustrated in Figs. 3 and 4.

Theorem 40. *Let Assumptions* C1 *and* C2 *be satisfied with* $\alpha > 1$ *being the degree of curvature of* $\mathcal{X}$ *at* x^*. *Set* $\gamma = \min(\alpha, 2)$. *Let*

$$\Gamma_N = \begin{pmatrix} N^{-\gamma/2(\gamma-1)} & 0 & \cdots & 0 \\ 0 & N^{-1/2(\gamma-1)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ddots & N^{-1/2(\gamma-1)} \end{pmatrix},$$

and $\rho_N = N^{(2-\gamma)/(2\gamma-2)}$. *Then* Z_N *epiconverges in distribution to* Z, *where* $Z(\cdot)$ *is of the following form: if* $\alpha < 2$, *then*

$$Z(t) = t^T \overline{k} + t^T Y + i_{\mathcal{X}^*}(t);$$

if $\alpha = 2$, *then*

$$Z(t) = t^T \overline{k} + t^T Y + t^T P^T A P t + i_{\mathcal{X}^*}(t);$$

if $\alpha > 2$, *then*

$$Z(t) = t^T \overline{k} + t^T Y + t^T P^T A P t + i_{\mathcal{X}^0}(t).$$

Here $P = I - \|\overline{k}\|^{-2} \overline{k}\, \overline{k}^T$ *is the projection operator onto the orthogonal complement of* $\overline{k}$, $\Sigma = P^T \mathbb{E}(kk^T) P$ *and* $Y \sim N(0, \Sigma)$. $\mathcal{X}^*$ *is given by* (9.13) *and* $\mathcal{X}^0$ *is the tangential cone.*

Proof. See Pflug (1995). □

Example 41. Consider the following program

$$\text{Min}\{x_1 + \mathbb{E}[x_1 \xi_1 + x_2 \xi_2] : k|x_2|^\alpha - x_1 \leq 0\} \tag{9.15}$$

where $1<\alpha<2$, $\mathbb{E}(\xi_1)=\mathbb{E}(\xi_2)=0$ and

$$\mathrm{Cov}(\xi_1,\xi_2)=\Sigma=\begin{pmatrix}\sigma_1^2 & \sigma_{12}\\ \sigma_{12} & \sigma_2^2\end{pmatrix}.$$

It is evident that $x^*=(0,0)$ is the solution of (9.15). Let $\kappa=1/(2\alpha-1)$ and

$$\Gamma_N=\begin{pmatrix}N^{-\alpha\kappa} & 0\\ 0 & N^{-\kappa}\end{pmatrix}.$$

Let $(\xi_{1,n},\xi_{2,n})$ be an i.i.d. sequence distributed like (ξ_1,ξ_2). The empirical program pertaining to (9.15) is

$$\begin{aligned}&\mathrm{Min}\left\{N^{\alpha\kappa}\left[N^{-\alpha\kappa}t_1+N^{-\kappa}N^{-1}t_2\sum_{n=1}^{N}\xi_{2,i}+N^{-\alpha\kappa}N^{-1}t_1\sum_{n=1}^{N}\xi_{1,n}\right]:k|t_2|^\alpha-t_1\le 0\right\}\\ &=\mathrm{Min}\left\{t_1+N^{-1/2}t_2\sum_{n=1}^{N}\xi_{2,i}+:k|t_2|^\alpha-t_1\le 0\right\}\end{aligned}$$

which converges to the asymptotic program

$$\mathrm{Min}\{t_1+t_2Y:k|t_2|^\alpha-t_1\le 0\} \tag{9.16}$$

where $Y\sim N(0,\sigma_2^2)$. Let (T_1,T_2) be distributed according to the arg min-distribution of (9.16). Then $T_1=k^{1/(1-\alpha)}|\frac{Y}{\alpha}|^{\alpha/(\alpha-1)}$ and $T_2=\mathrm{sign}(Y)|\frac{Y}{k\alpha}|^{1/(\alpha-1)}$. For the case $\alpha=2$, T_1 has a χ^2 distribution and T_2 has a normal distribution.

Remark. If the curvature $\alpha=1$, then an asymptotic distribution exists if $\mathbb{E}(k(\xi))=0$. In this case the asymptotic stochastic program is $\mathrm{Min}\{Z(t)=t^TY+i_{\mathcal{X}^0}(t)\}$. The appropriate localizing sequence is $\Gamma_N=N^{-1}I$.

Instead of using the just presented "local coordinates" one may use in smooth cases the "generalized delta method". The idea of this method is to relate the variations of the objective function to the variations of the arg mins: suppose that $\overline{x}(v)=\arg\min_{x\in\mathcal{X}}f(x)+x\cdot v$. The mapping $v\mapsto\overline{x}(v)$ may be complicated. Only in very simple cases it is linear or locally linear. As an example, take $\mathcal{X}=\mathbb{R}^d$, $f(x)=\frac{1}{2}x^TAx$ for a positive definite A. Then $\overline{x}(v)=-A^{-1}v$, hence linear. If the criterion function f is, however, not locally quadratic or some constraints are active in the optimum, then this function is

not locally linear. The idea of the generalized delta method is to use the following argument

$$\begin{aligned}\arg\min \hat{f}_N(x) &= \arg\min_x \{f(x) + [\hat{f}_N(x) - f(x)] - [\hat{f}_N(x^*) - f(x^*)]\} \\ &= \arg\min_x \{f(x) + (x - x^*)[\nabla \hat{f}_N(x^*) - \nabla f(x^*)] + \text{remainder}\} \\ &= \arg\min_x \{f(x) + x[\nabla \hat{f}_N(x^*) - \nabla f(x^*)] + \text{remainder}\} \\ &\sim \overline{x}(\nabla \hat{f}_N(x^*) - \nabla f(x^*)).\end{aligned}$$

(see Shapiro, 1993). The random variable $[\nabla \hat{f}_N(x^*) - \nabla f(x^*)]$ may be normal in regular cases. The random variable $\overline{x}([\nabla \hat{f}_N(x^*) - \nabla f(x^*)])$ is normal only in the special case, when $\overline{x}$ is locally linear. One way to look at $\overline{x}(\nabla \hat{f}_N(x^*) - \nabla f(x^*))$ is to find its directional derivatives $\overline{x}'(0, d)$. This method is elaborated in the book by Bonnans and Shapiro (1999).

If the constraints are also random, then the asymptotic results may be based on a generalized equation approach and a Z-theorem (**Z**ero Theorem): suppose for instance, that the stochastic program is

$$\text{Min}\{f(x) = \mathbb{E}[F(x, \xi)] : g(x) = \mathbb{E}[G(x, \xi)] \leq 0,\ x \in \mathcal{X}\}. \tag{9.17}$$

Assuming convexity of f, g and the set $\mathcal{X}$, the generalized equations, which characterize the solution, are

$$\begin{aligned}&0 \in \nabla f(x) + y \nabla g(x) + N_{\mathcal{X}}(x) \\ &0 \in -g(x) + N(y)\end{aligned}$$

Here $N_{\mathcal{X}}(x) = \{u : u^T(v - x) \leq 0,\ v \in \mathcal{X}\}$ and $N(y) = \{u \leq 0 ; u^T y = 0\}$.

The abstract Z-theorem reads: if the set-valued process $\hat{H}_N(\cdot)$ converges to $H(\cdot)$, then the solutions $\hat{x}_N$ of the generalized equation $0 \in \hat{H}_N(x)$ converge to the solution x^* of $0 \in H(x)$.

Such an approach was used by King and Rockafellar (1993) and Shapiro (1993). For Z-theorems see Anisimov and Pflug (2000) and the chapter *Stability of Stochastic Programs* by Römisch in this volume and references therein.

10 Bibliographic remarks

Glivenko and independently Cantelli proved what is now called the Glivenko–Cantelli Theorem in 1933. It states the uniformity of the empirical measure convergence for intervals. Blum (1955) and DeHardt (1971) gave the

first bracketing conditions for uniformity. This idea was further developed by Ossiander (1987) and Van der Vaart (1994) and many others. In 1971 Vapnik and Cervonenkis published their seminal paper on the shattering dimension of a class of sets and the relation to uniformity. The first uniform entropy result is due to Pollard (1982). Ledoux and Talagrand introduced in their 1991 book about probability in Banach spaces isoperimetric methods. These methods are extensively used in Talagrand's fundamental paper about sharper bounds. The 1996 book by van der Vaart and Wellner summarizes various results about Glivenko–Cantelli and Donsker classes of functions. It also contains a simplified proof of Talagrand's theorem.

The foregoing results were obtained for i.i.d. random variables. For robustness reasons, one is also interested in "slightly" dependent variables. Whereas Laws of Large Numbers are well known for dependent variables (Andrews, 1988; Andrews and Pollard, 1994), tight bounds are still missing. The most promising direction of generalization of the i.i.d. case goes through mixing conditions (Doukhan et al., 1994, 1995), see also the book by De la Peña and Giné (1999). Talagrand's measure concentration techniques were generalized by Marton (1996).

The notion of epiconvergence was introduced by Wijsman (1964). This reference and much more can be found in the book Rockafellar and Wets (1998). Salinetti and Wets (1986) introduced the concept of convergence in probability for sequences of closed-set valued random variables. Vogel (1994) considered "one-sided" versions of it by defining upper and lower semi-convergence in probability.

Asymptotic distribution for solutions of stochastic programs were given by King and Rockafellar, Dupacova and Wets, Shapiro, Römisch, Pflug and many others.

There is a close relation between the asymptotic theory of stochastic optimization programs and the asymptotic theory of statistical estimation under constraints. Among the large number of papers in the latter area, we cite here: Geyer (1994), Dong and Wets (2000) and Shapiro (2000).

References

Andrews, D.W.K. (1988). Laws of large numbers for dependent non-identically distributed random variables. *Econometric Theory* 4, 458–467.

Andrews, D.W.K., D. Pollard (1994). An introduction to functional central limit theorems for dependent stochastic processes. *International Statistical Review* 62, 119–132.

Anisimov, A., G.Ch. Pflug (2000). Z-Theorems: limits of stochastic equations. *Bernoulli* 6(5), 124–135.

Attouch, H., R.J.B. Wets (1990). Laws of large numbers for random lsc functions. *Seminaire Anal. Convexe* 20, Exp. 13, Université de Montpellier.

Billingsley, P. (1968). *Convergence of Probability Measures*, Wiley, New York.

Blum, J.R. (1955). On the convergence of empirical distributions. *Ann. Math. Statist.* 26, 527–529.

Bonnans, F., A. Shapiro (1999). *Perturbation Analysis of Optimization Problems*, Springer-Verlag, New York.

Dehardt, J. (1971). Generalizations of the Glivenko–Cantelli Theorem. *Ann. Math. Statist.* 42, 2050–2055.

De la Peña, V.H., E. Giné (1999). *Decoupling. From Dependence to Independence*, Springer, New York.

Deuschel, J.D., D.W. Strook (1989). *Large Deviations*, Academic Press, Boston.

Devroye, L., G. Lugosi, (2001). *Combinatorial Methods in Density Estimation*, Springer Series in Statistics, New York.

Dong, M., R.J.B. Wets (2000). Estimatind density functions: A constrained maximum likelihood approach. *Journal of Nonparametric Statistics* 12(4), 549–595.

Doukhan, P., P. Massart, E. Rio (1994). The functional central limit theorem for strongly mixing processes. *Annales Institut Henri Poincaré Probabilités et Statistiques* 30(1), 63–82.

Doukhan, P., P. Massart, E. Rio (1995). Invariance principles for absolutely regular empirical processes. *Annales Institut Henri Poincaré Probabilités et Statistiques* 31(2), 393–427.

Dudley, R.M. (1982). A course on empirical processes. *Lecture Notes in Mathematics,* 1097, Springer, New York, pp. 1–142.

Dudley, R.M. (1999). *Uniform Central Limit Theorems*, Cambridge University Press, Cambridge.

Dupacova, J., R.J.B. Wets (1988). Asymptotic behavior of statistical estimates and optimal solutions of stochastic optimization problems. *The Annals of Statistics* 16, 1517–1549.

Etemadi, N. (1981). An elementary proof of the strong law of large numbers. *Z. Wahrscheinlichkeitstheorie verw. Geb.* 55(1), 119–122.

Gaivoronski, A., Yu. Ermoliev, C. Nedeva (1985). Stochastic optimization problems with incomplete information on distribution functions. *SIAM Journal on Control and Optimization* 23(5), 697–716.

Geyer, Ch. (1994). On the asymptotics of constrained M-estimation. *The Annals of Statistics* 22, 1993–2010.

Giné, E., J. Zinn (1984). Some limit theorems for empirical processes. *The Annals of Probability* 12, 929–989.

Giné, E., J. Zinn (1986). Lectures on the central limit theorem for empirical processes. *Lecture Notes in Mathematics*, Vol. 1221, Springer, Berlin, pp. 50–113.

Hackenbroch, W., A. Thalmaier (1994). *Stochastische Analysis*, B.G. Teubner, Stuttgart.

Hall, P., C.C. Heyde (1980). *Martingale Limit Theory and its Application*, Academic Press, New York.

Kaniovski, Yu.M., A. King, R.J.-B. Wets (1993). Probabilistic bounds (via large deviations) for the solutions of stochastic programming problems. Research Report Rc 18752, IBM Research Center, Yorktown Heights.

King, A.J., R.T. Rockafellar (1993). Asymptotic theory for solutions in statistical estimation and stochastic programming. *Math. Oper. Research* 18, 148–162.

Ledoux, M., M. Talagrand (1991). *Probability in Banach Spaces*, Springer, New York.

Louhichi, S. (2000). Weak convergence for empirical processes of associated sequences. *Annales Institut Henri Poincaré, Probabilités et Statistiques* 36(5), 547–569.

Marton, K. (1996). Bounding $\overline{d}$-distance by informational divergence: A method to prove measure concentration. *The Annals of Probability* 24(2), 857–866.

Niederreiter, H.(1992). *Random Number Generation and Quasi Monte Carlo Methods*. CBMS-NSF Regional Conference Series in Applied Mathematics 63, SIAM Philadelphia.

Ossiander, M. (1987). A central limit theorem under metric entropy with L_2-bracketing. *The Annals of Probability* 15, 897–919.

Pflug, G.Ch. (1992). Asymptotic dominance and confidence for solutions of stochastic programs. *Czechoslovak Journal of OR* 1(1), 21–30.

Pflug, G.Ch. (1995). Asymptotic stochastic programs. *Mathematics of OR* 20(4), 769–789.

Pflug, G.Ch. (1999). Stochastic programs and statistical data. *Annals of OR* 85, 59–78.

Pflug, G.Ch., A. Ruszczynski, R. Schultz (1998). On the Glivenko Cantelli problem in stochastic programming: Linear recourse. *Mathematics of OR* 23(1), 204–220.

Pflug, G.Ch., A. Ruszczynski, R. Schultz (1998). On the Glivenko–Cantelli problem in stochastic programming: Mixed integer linear recourse. *Math. Methods of OR* 47, 39–49.

Pollard, D. (1982). A central limit theorem for empirical processes. *Journal of the Australian Mathematical Society (Series A)* 33, 235–248.

Pollard, D. (1984). *Convergence of Stochastic Processes*, Springer, New York.
Rio, E. (1993). Covariance inequalities for strongly mixing processes. *Annales Institut Henri Poincaré, Probabilités et Statistiques* 29, 587–597.
Rockafellar, T.R., R.J.B. Wets (1998). Variational Analysis. *Grundlehren der Mathematischen Wissenschaften,* Vol. 317, Springer, Berlin.
Rockafellar T.R., S.P. Uryasev (1999). Optimization of conditional value-at-risk. *The Journal of Risk* 2(3), 2000, 21–41.
Römisch, W. (1986). On the convergence of measurable selections and applications to approximations in stochastic optimization. *Zeitschrift für Analysis und Anwendungen* 5, 277–288.
Salinetti, G., R.J.B. Wets (1986). On the convergence in distribution of measurable multifunctions (random sets), normal integrands, stochastic processes and stochastic infima. *Mathematics of OR* 11, 385–419.
Shapiro, A. (1989). Asymptotic properties of statistical estimators in stochastic programming. *The Annals of Statistics* 17(2), 841–858.
Shapiro, A. (1993). Asymptotic behavior of optimal solutions in stochastic programming. *Mathematics of Operations Research* 18, 829–845.
Shapiro, A. (1994). Quantitative stability in stochastic programming. *Mathematical Programming* 67, 99–108.
Shapiro, A. (2000). Statistical inference of stochastic optimization problems, in: S.P. Urasev (ed.), *Probabilistic Constrained Optimization*, Kluwer Academic Publishers, Norwell, USA, pp. 288–313.
Shapiro, A. (2000). On the asymptotics of constrained local M-estimators. *The Annals of Statistics* 28(3), 948–960.
Shiryaev, A.N. (1996). Probability, 2nd ed., Graduate Texts in Mathematics 95, Springer-Verlag, New York.
Shorack, G.R., J.A. Wellner (1986). *Empirical Processes with Applications to Statistics*, Wiley, New York.
Talagrand, M. (1987). The Glivenko–Cantelli problem. *The Annals of Probability* 15, 837–870.
Talagrand, M. (1994). Sharper bounds for Gaussian and empirical processes. *The Annals of Probability* 22, 28–76.
Talagrand, M. (1995). Concentration of measure and isoperimetric inequalities in product spaces. *Publ. IHES* 81, 73–205.
Talagrand, M. (1996). The Glivenko–Cantelli problem, ten years later. *Journal of Theoretical Probability* 9, 371–384.
Van der Vaart, A. (1994). Bracketing smooth functions. *Stochastic Processes and their Applications* 52, 93–105.
Van der Vaart, A.W., J.W. Wellner (1996). *Weak Convergence and Empirical Processes*, Springer, New York.
Van der Vaart, A.W. (1998). *Asymptotic Statistics*, Cambridge University Press, Cambridge.
Vapnik, V.N., A.Y. Chervonenkis (1971). On the uniform convergence of relative frequences of events to their probabilities. *Theory of Probability and its Applications* 16, 264–280.
Vapnik, V.N., A.Y. Chervonenkis (1981). Necessary and sufficient conditions for the uniform convergence of means to their expectations. *Theory of Probability and its Applications* 24, 532–553.
Varadarajan, V.S. (1958). Weak convergence of measures in separable metric spaces. On the convergence of sample probability distributions. *Sankhya* 19, 15–22 and 23–26.
Vogel, S. (1994). A stochastic approach to stability in stochastic programming. *Journal of Computational and Applied Mathematics* 56, 65–96.
Wijsman, R.A. (1964). Convergence of sequences of convex sets, cones and functions. *Bulletin of the AMS* 70, 186–188.
Yukich, J.E. (1986). Rates of convergence for classes of functions: the non-iid case. *Journal of Multivariate Analysis* 20, 175–189.

A. Ruszczyński and A. Shapiro, Eds., *Handbooks in OR & MS, Vol. 10*

Chapter 8

Stability of Stochastic Programming Problems

Werner Römisch

Institute of Mathematics, Humboldt-University Berlin, D-10099 Berlin, Germany,
E-mail: romisch@mathematik.hu-berlin.de

Abstract

The behaviour of stochastic programming problems is studied in case of the underlying probability distribution being perturbed and approximated, respectively. Most of the theoretical results provide continuity properties of optimal values and solution sets relative to changes of the original probability distribution, varying in some space of probability measures equipped with some convergence and metric, respectively. We start by discussing relevant notions of convergence and distances for probability measures. Then we associate a distance with a stochastic program in a natural way and derive (quantitative) continuity properties of values and solutions by appealing to general perturbation results for optimization problems. Later we show how these results relate to stability with respect to weak convergence and how certain ideal probability metrics may be associated with more specific stochastic programs. In particular, we establish stability results for two-stage and chance constrained models. Finally, we present some consequences for the asymptotics of empirical approximations and for the construction of scenario-based approximations of stochastic programs.

Key words: Stochastic programming, stability, weak convergence, probability metric, Fortet–Mourier metric, discrepancy, risk measure, two-stage, mixed-integer, chance constrained, empirical approximation, scenario reduction.

1 Introduction

Stochastic programming is concerned with models for optimization problems under stochastic uncertainty that require a decision on the basis of given probabilistic information on random data. Typically, deterministic equivalents of such models represent finite-dimensional nonlinear programs whose

objectives and/or constraints are given by multivariate integrals with respect to the underlying probability measure. At the modelling stage these probability measures reflect the available knowledge on the randomness at hand. This fact and the numerical challenges when evaluating the high-dimensional integrals have drawn great attention to the stability analysis of stochastic programs with respect to changes in the underlying probability measure. In this chapter we present a unified framework for such a stability analysis by regarding stochastic programs as optimization problems depending on the probability measure varying in some space of measures endowed with some distance. We give stability results both for general models and for more specific stochastic programs like two-stage and chance constrained models and include most of the proofs. Moreover, we discuss some conclusions about specific approximation procedures for stochastic programs.

To specify the stochastic programming models for our analysis, we recall that many deterministic equivalents of such models are of the form

$$\min\left\{\int_{\Xi} F_0(x,\xi)\,\mathrm{d}P(\xi)\colon x\in X,\ \int_{\Xi} F_j(x,\xi)\,\mathrm{d}P(\xi)\le 0, j=1,\dots,d\right\}, \tag{1.1}$$

where the set $X\subseteq\mathbb{R}^m$ is closed, Ξ is a closed subset of $\mathbb{R}^s$, the functions F_j from $\mathbb{R}^m\times\Xi$ to the extended reals $\overline{\mathbb{R}}$ are random lower semicontinuous functions for $j=0,\dots,d$, and P is a Borel probability measure on Ξ.

The set X is used to describe all constraints not depending on P, and the set Ξ contains the supports of the relevant measures and provides some flexibility for formulating the models and the corresponding assumptions. We recall that F_j is a random lower semicontinuous function if its epigraphical mapping $\xi\mapsto \operatorname{epi} F_j(\cdot,\xi):=\{(x,r)\in\mathbb{R}^m\times\mathbb{R}\colon F_j(x,\xi)\le r\}$ is closed-valued and measurable, which implies, in particular, that $F_j(\cdot,\xi)$ is lower semicontinuous for each $\xi\in\Xi$ and $F_j(x,\cdot)$ is measurable for each $x\in\mathbb{R}^m$.

Although, our stability analysis mainly concerns model (1.1) and its specifications, we also provide an approach to the stability of more general models that contain risk functionals and are of the form

$$\min\left\{\mathbb{F}_0\left(P[F_0(x,\cdot)]^{-1}\right)\colon x\in X,\ \mathbb{F}_j(P[F_j(x,\cdot)]^{-1})\le 0, j=1,\dots,d\right\}, \tag{1.2}$$

where the risk functionals $\mathbb{F}_j$, $j=0,\dots,d$,, map from suitable subsets of the set $\mathcal{P}(\mathbb{R})$ of all probability measures on $\mathbb{R}$ to $\mathbb{R}$. In general, the functionals $\mathbb{F}_j$ depend on a measure in $\mathcal{P}(\mathbb{R})$ in a more involved way than the expectation functional $\mathbb{F}_e(G):=\int_{\mathbb{R}} r\,\mathrm{d}G(r)$, for which we have $\mathbb{F}_e(P[F_0(x,\cdot)]^{-1})=\int_{\mathbb{R}} r\,\mathrm{d}P[F_0(x,\cdot)]^{-1}(r)=\int_{\Xi} F_0(x,\xi)\,\mathrm{d}P(\xi)$. Another example is the variance

functional $\mathbb{F}_v(G) := \int_{\mathbb{R}} r^2 \, dG(r) - (\int_{\mathbb{R}} r \, dG(r))^2$. We also refer to the value-at-risk functional in Example 1 and to the examples in Section 2.4.

We illustrate the abstract models by the classical newsboy example (see e.g., Dupačová (1994), Example 1 in Ruszczyński and Shapiro (2003)).

Example 1. (newsboy problem) A newsboy must place a daily order for a number x of copies of a newspaper. He has to pay r dollars for each copy and sells a copy at c dollars, where $0 < r < c$. The daily demand ξ is random with (discrete) probability distribution $P \in \mathcal{P}(\mathbb{N})$ and the remaining copies $y(\xi) = \max\{0, x - \xi\}$ have to be removed. The newsboy might wish that the decision x maximizes his expected profit or, equivalently, minimizes his expected costs, i.e.,

$$\begin{aligned}\int_{\mathbb{R}} F_0(x, \xi) \, dP(\xi) &:= \int_{\mathbb{R}} [(r - c)x + c\max\{0, x - \xi\}] \, dP(\xi) \\ &= (r - c)x + c\sum_{k \in \mathbb{N}} \pi_k \max\{0, x - k\} \\ &= rx - cx\sum_{\substack{k \in \mathbb{N} \\ k \geq x}} \pi_k - c\sum_{\substack{k \in \mathbb{N} \\ k < x}} \pi_k k\end{aligned}$$

where π_k is the probability of demand $k \in \mathbb{N}$. The unique integer solution is the maximal $k \in \mathbb{N}$ such that $\sum_{i=k}^{\infty} \pi_i \geq \frac{r}{c}$. Another possibility is that the newsboy wishes to maximize his profit and, at the same time, to minimize his risk costs cs where s bounds the number $y(\xi)$ of copies that remain with probability p. The minimal s corresponds to his value-at-risk at level p. The resulting stochastic program reads

$$\min_{x \in \mathbb{R}_+} \{(r - c)x + c \inf \{s \in \mathbb{R}_+ : P(y(\xi) \leq s) \geq p\}\}.$$

The latter program is equivalent to the chance constrained model

$$\min_{(x,s) \in \mathbb{R}^2_+} \left\{ (r - c)x + cs : \sum_{\substack{k \in \mathbb{N} \\ x - s \leq k}} \pi_k \geq p \right\} \tag{1.3}$$

whose unique integral solution is $(k,0)$ with the maximal $k \in \mathbb{N}$ such that $\sum_{i=k}^{\infty} \pi_i \geq p$. Hence, the minimum risk solution is more pessimistic than the minimal expected cost solution if $\frac{r}{c} < p < 1$, i.e., if the newsboy wants to be sure with high probability that no copies of the newspaper remain.

However, the inherent difficulty of all these approaches is that the newsboy does not know the probability distribution P of the demand and has to use some approximation instead. Hence, he is interested in the stability of his decision which means that it does not vary too much for small perturbations of the data. For instance, his decision might be based on n independent identically distributed observations ξ_i, $i = 1, \ldots, n$, of the demand, i.e., on approximating P by the empirical measure P_n (cf. Section 4.1) and, in case of minimal expected costs, on solving the approximate problem

$$\min_{x \in \mathbb{R}_+} \left\{ (r-c)x + \frac{c}{n} \sum_{i=1}^{n} \max\{0, x - \xi_i\} \right\}. \tag{1.4}$$

Of course, this approach is only justified if some optimal solution x_n of the approximate problem (1.4) is close to some original solution for sufficiently large n. Both variants of the newsboy problem represent specific two-stage and chance constrained stochastic programs, respectively. Their discussion will be continued in the Examples 15, 19 and 54.

Throughout we will denote the set of all Borel probability measures on Ξ by $\mathcal{P}(\Xi)$, the feasible set of (1.1) by $\mathcal{X}(P)$, the optimal value by $\vartheta(P)$ and the (ε-approximate) solution set of (1.1) by $X_\varepsilon^*(P)$ and $X^*(P)$, respectively, i.e.,

$$\mathcal{X}(P) := \left\{ x \in X : \int_\Xi F_j(x, \xi)\, \mathrm{d}P(\xi) \leq 0, j = 1, \ldots, d \right\}, \tag{1.5}$$

$$\vartheta(P) := \inf \left\{ \int_\Xi F_0(x, \xi)\, \mathrm{d}P(\xi) : x \in \mathcal{X}(P) \right\}, \tag{1.6}$$

$$X_\varepsilon^*(P) := \left\{ x \in \mathcal{X}(P) : \int_\Xi F_0(x, \xi)\, \mathrm{d}P(\xi) \leq \vartheta(P) + \varepsilon \right\} \quad (\varepsilon \geq 0), \tag{1.7}$$

$$X^*(P) := X_0^*(P) = \left\{ x \in \mathcal{X}(P) : \int_\Xi F_0(x, \xi)\, \mathrm{d}P(\xi) = \vartheta(P) \right\}. \tag{1.8}$$

In this chapter, *stability* mostly refers to continuity properties of the optimal value function $\vartheta(\cdot)$ and the (ε-approximate) solution-set mapping $X_\varepsilon^*(\cdot)$ at P, where both $\vartheta(\cdot)$ and $X_\varepsilon^*(\cdot)$ are regarded as mappings given on a set of probability measures endowed with a suitable distance. The distance has to be selected such that it allows to estimate differences of objective and constraint function values, and, that it is optimum adapted to the model at hand.

Fortunately, there exists a diversity of convergence notions and metrics in probability theory and statistics that address different goals and are based on various constructions (see, e.g., Rachev (1991) and van der Vaart (1998)). We will use so-called distances with ζ-structure that are given as uniform distances of expectations of functions taken from a class $\mathcal{F}$ of measurable functions from Ξ to $\mathbb{R}$, i.e.,

$$d_{\mathcal{F}}(P,Q) = \sup_{F\in\mathcal{F}} \left| \int_{\Xi} F(\xi)\, \mathrm{d}P(\xi) - \int_{\Xi} F(\xi)\, \mathrm{d}Q(\xi) \right|. \tag{1.9}$$

In a first step we choose the class $\mathcal{F}$ as the set $\{F_j(x,\cdot)\colon x \in X \cap \mathrm{cl}\,\mathcal{U}, j = 0,\ldots,d\}$, where $\mathcal{U}$ is a properly chosen open subset of $\mathbb{R}^m$, and derive some (qualitative and quantitative) stability results in the Sections 2.2 and 2.3. Such a distance forms a kind of *minimal information* (*m.i.*) *metric* for the stability of (1.1). Some of the corresponding results (e.g., the Theorems 5 and 9) work under quite weak assumptions on the underlying data of (1.1). In particular, if possible differentiability or even continuity assumptions on the functions $x \mapsto \int_{\Xi} F_j(x,\xi)\, \mathrm{d}P(\xi)$ are avoided for the sake of generality. The approach is inspired by general perturbation results for optimization problems in Klatte (1987, 1994); Attouch and Wets (1993) and in the monographs by Bank et al. (1982) and Rockafellar and Wets (1998) and Bonnans and Shapiro (2000).

Since the m.i. metrics are often rather involved and difficult to handle, we look, on the one hand, for implications of the general qualitative result on stability with respect to the topology of weak convergence. On the other hand, we look for another metric having ζ-structure by enlarging the class $\mathcal{F}$ and, hence, bounding the m.i. metric from above. Our strategy for controlling this enlargement procedure consists in adding functions to the enlarged class that share the essential analytical properties with some of the functions $F_j(x,\cdot)$. As a result of this process we obtain *ideal* metrics that are optimum adjusted to the model (1.1) or to a whole class of models and that enjoy pleasant properties (e.g., a duality and convergence theory). In Section 3, we show for three types of stochastic programs how such ideal metrics come to light in a natural way by revealing the analytical properties of the relevant functions $F_j(x,\cdot)$. At the same time, we obtain quantitative stability results for all models.

For two-stage models containing integer variables and for chance constrained models, the relevant functions are discontinuous and their ideal classes contain products of (locally) Lipschitzian functions and of characteristic functions of sets describing regions of continuity (see Sections 3.2 and 3.3).

When using stability results for designing or analyzing approximation schemes or estimation procedures, further properties of the function classes $\mathcal{F}$ and of the metrics may become important. For example, we derive covering numbers of certain function classes and discuss their implications on probabilistic bounds for empirical optimal values and solution sets.

The chapter is organized as follows. First Section 2 contains some prerequisites on convergences and metric distances of probability measures.

This is followed by our main qualitative stability result (Theorem 5) and its conclusions on the stability with respect to weak convergence of probability measures. We continue with the quantitative stability results for solution sets of (1.1) (Theorems 9 and 12) and a Lipschitz continuity result (Theorem 13) for ε-approximate solution sets of convex models. We add a discussion of how to associate ideal metrics with more specific stochastic programs. Section 2 is finished by discussing the challenges and by presenting first results of a perturbation analysis for stochastic programs containing risk functionals (1.2). In Section 3 we consider linear two-stage, mixed-integer two-stage and linear chance constrained stochastic programs and present various perturbation results for such models. The potential of our general perturbation analysis is explained in Section 4 for two types of approximations of the underlying probability measure P. First, we consider empirical measures as nonparametric estimators of P and derive asymptotic statistical properties of values and solutions by using empirical process theory. Secondly, we discuss the optimal construction of finitely discrete measures based on probability metrics and sketch some results and heuristic algorithms for the optimal reduction of discrete measures. We conclude the chapter with some bibliographical notes on the relevant literature.

2 General stability results

2.1 *Convergences and metrics of probability measures*

Let us consider the set $\mathcal{P}(\Xi)$ of all Borel probability measures with support contained in a closed subset Ξ of $\mathbb{R}^s$. We will endow the set $\mathcal{P}(\Xi)$ or some of its subsets with different convergences and distances, which are adapted to the underlying stochastic program or to a whole class of stochastic programs. The classical convergence concept in probability theory is the *weak convergence* of measures in $\mathcal{P}(\Xi)$ (see e.g., Billingsley (1968) and Dudley (1989)). A sequence (P_n) in $\mathcal{P}(\Xi)$ is said to converge weakly to $P \in \mathcal{P}(\Xi)$, shortly $P_n \xrightarrow{w} P$, if

$$\lim_{n\to\infty} \int_{\Xi} g(\xi)\, \mathrm{d}P_n(\xi) = \int_{\Xi} g(\xi)\, \mathrm{d}P(\xi) \tag{2.10}$$

holds for each g in the space $C_b(\Xi)$ of bounded continuous functions from Ξ to $\mathbb{R}$. It is well known that the topology τ_w of weak convergence is metrizable (e.g., by the bounded Lipschitz metric (2.11)) and that $P_n \xrightarrow{w} P$ holds iff the sequence of probability distribution functions of P_n converges pointwise to the distribution function F_P of P at all continuity points of F_P. Another important property of weak convergence is the continuous mapping theorem: If $P_n \xrightarrow{w} P$ and $g\colon \Xi \to \mathbb{R}$ is measurable, bounded and P-continuous, i.e., $P(\{\xi \in \Xi\colon g \text{ is not continuous at } \xi\}) = 0$, we have (2.10).

Most of the distances on (subsets of) $\mathcal{P}(\Xi)$ that will be considered are of the form $d_{\mathcal{F}}$ in (1.9), where $\mathcal{F}$ is a class of measurable functions from Ξ to $\mathbb{R}$, and are defined on the set $\mathcal{P}_{\mathcal{F}} := \{Q \in \mathcal{P}(\Xi)\colon \sup_{F\in\mathcal{F}} |\int_{\Xi} F(\xi)\,\mathrm{d}Q(\xi)| < \infty|\}$, where $d_{\mathcal{F}}$ is finite. A uniform distance of the form (1.9) is called a distance having *ζ-structure* (see Zolotarev (1983) and Rachev (1991)). Clearly, $d_{\mathcal{F}}$ does not change if the set $\mathcal{F}$ is replaced by its convex hull. It is nonnegative, symmetric and satisfies the triangle inequality, i.e., a pseudometric on $\mathcal{P}_{\mathcal{F}}$. $d_{\mathcal{F}}$ is a metric if the class $\mathcal{F}$ is rich enough to preserve that $d_{\mathcal{F}}(P, Q) = 0$ implies $P = Q$. Next we list some important examples of distances having ζ-structure, where the classes $\mathcal{F}$ range from (locally) Lipschitz continuous functions to piecewise constant functions with a prescribed structure of discontinuity sets.

Example 2. (metrics with ζ-structure)

(a) For $p = 0$ and $p \geq 1$ we introduce classes $\mathcal{F}_p(\Xi)$ of locally Lipschitz continuous functions that increase with p

$$\mathcal{F}_p(\Xi) := \left\{F\colon \Xi \mapsto \mathbb{R}\colon \left|F(\xi) - F(\tilde{\xi})\right| \leq c_p(\xi,\tilde{\xi})\|\xi - \tilde{\xi}\|, \forall \xi,\tilde{\xi} \in \Xi\right\},$$

$$\mathcal{F}_0(\Xi) := \mathcal{F}_1(\Xi) \cap \left\{F \in C_b(\Xi)\colon \sup_{\xi\in\Xi}\left|F(\xi)\right| \leq 1\right\}.$$

Here, $\|\cdot\|$ denotes some norm on $\mathbb{R}^s$ and $c_p(\xi, \tilde{\xi}) := \max\{1, \|\xi\|, \|\tilde{\xi}\|\}^{p-1}$ for all $\xi, \tilde{\xi} \in \Xi$ and $p \geq 1$ describes the growth of the local Lipschitz constants. The corresponding distance with ζ-structure for $p = 0$ is the *bounded Lipschitz metric* (Section 11.3 of Dudley (1989))

$$\beta(P, Q) := \sup_{F\in\mathcal{F}_0(\Xi)}\left|\int_{\Xi} F(\xi)\,\mathrm{d}P(\xi) - \int_{\Xi} F(\xi)\,\mathrm{d}Q(\xi)\right| \tag{2.11}$$

and metrizes the weak convergence on $\mathcal{P}(\Xi)$. For $p = 1$ we arrive at the *Kantorovich metric*

$$\zeta_1(P, Q) := \sup_{F\in\mathcal{F}_1(\Xi)}\left|\int_{\Xi} F(\xi)\,\mathrm{d}P(\xi) - \int_{\Xi} F(\xi)\,\mathrm{d}Q(\xi)\right| \tag{2.12}$$

and for $p \geq 1$ at the *p-th order Fortet–Mourier metrics* (see Fortet and Mourier (1953) and Rachev (1991))

$$\zeta_p(P, Q) := \sup_{F\in\mathcal{F}_p(\Xi)}\left|\int_{\Xi} F(\xi)\,\mathrm{d}P(\xi) - \int_{\Xi} F(\xi)\,\mathrm{d}Q(\xi)\right| \tag{2.13}$$

on the set $\mathcal{P}_p(\Xi) := \{Q \in P(\Xi) \int_\Xi \|\xi\|^p \, dQ(\xi) < \infty\}$ of probability measures having finite p-th order absolute moments. It is known that a sequence (P_n) converges to P in $(\mathcal{P}_p(\Xi),\zeta_p)$ iff it converges weakly and

$$\lim_{n\to\infty} \int_\Xi \|\xi\|^p \, dP_n(\xi) = \int_\Xi \|\xi\|^p \, dP(\xi)$$

holds. Furthermore, the estimate

$$\left| \int_\Xi \|\xi\|^p \, dP(\xi) - \int_\Xi \|\xi\|^p \, dQ(\xi) \right| \le p\zeta_p(P, Q)$$

is valid for each $p \ge 1$ and all $P,Q \in \mathcal{P}_p(\Xi)$ (Section 6 in Rachev (1991)). Hence, closeness with respect to ζ_p implies the closeness of q-th order absolute moments for $q \in [1, p]$.

(b) Let $\mathcal{B}$ denote a set of Borel subsets of Ξ and consider the class $\mathcal{F}_\mathcal{B} := \{\chi_B : B \in \mathcal{B}\}$ of their characteristic functions χ_B taking the value 1 if the argument belongs to B and 0 otherwise. The distance with ζ-structure generated by $\mathcal{F}_B$ is defined on $\mathcal{P}(\Xi)$. It takes the form

$$\alpha_\mathcal{B}(P, Q) := d_{\mathcal{F}_\mathcal{B}}(P,Q) = \sup_{B\in\mathcal{B}} |P(B) - Q(B)|$$

and is called $\mathcal{B}$-*discrepancy*. The following instances play a special role in the context of stability in stochastic programming:

(b1) Let Ξ be convex and $\mathcal{B}_\mathrm{c}(\Xi)$ the set of all closed convex subsets of Ξ.
(b2) Let Ξ be polyhedral and $\mathcal{B}_{\mathrm{ph}_k}(\Xi)$ the set of all polyhedra being subsets of Ξ and having at most k faces.
(b3) Let $\Xi = \mathbb{R}^s$ and $\mathcal{B}_\mathrm{h}(\Xi)$ be the set of all closed half-spaces in $\mathbb{R}^s$.
(b4) Let $\Xi = \mathbb{R}^s$ and $\mathcal{B}_\mathrm{K}(\Xi) := \{(-\infty,\xi] : \xi \in \mathbb{R}^s\}$ be the set of all cells.

The corresponding distances are the *isotrope discrepancy* α_c, the *polyhedral discrepancy* α_{ph_k}, the *half-space discrepancy* α_h and the *Kolmogorov metric*. The latter metric coincides with the uniform distance of distribution functions on $\mathbb{R}^s$ and is denoted by d_K, i.e.,

$$d_\mathrm{K}(P, Q) = \alpha_{\mathcal{B}_\mathrm{K}}(P, Q) = \sup_{\xi\in\mathbb{R}^s} |P((-\infty, \xi]) - Q((-\infty, \xi])|.$$

A sequence (P_n) converges to P in $\mathcal{P}(\Xi)$ with respect to $\alpha_\mathcal{B}$, where $\mathcal{B}$ is a class of closed convex subsets of Ξ, iff (P_n) converges weakly to P

and $P(\mathrm{bd}\, B) = 0$ holds for each $B \in \mathcal{B}$ (with $\mathrm{bd}\, B$ denoting the boundary of the set B).

The examples reveal some relations between the weak convergence of probability measures and their convergence with respect to a uniform metric $d_{\mathcal{F}}$ for some classes $\mathcal{F}$. Such relations have already been explored more systematically in the literature. A class $\mathcal{F}$ of measurable functions from Ξ to $\mathbb{R}$ is called a *P-uniformity class* if

$$\lim_{n\to\infty} d_{\mathcal{F}}(P_n, P) = 0 \tag{2.14}$$

holds for each sequence (P_n) that converges weakly to P. Necessary conditions for $\mathcal{F}$ to be a P-uniformity class are that $\mathcal{F}$ is uniformly bounded and that every function in $\mathcal{F}$ is P-continuous. Sufficient conditions are given in Billingsley and Topsøe (1967), Topsøe (1967, 1977) and Lucchetti et al. (1994). For example, $\mathcal{F}$ is a P-uniformity class if it is uniformly bounded and it holds that $P(\{\xi \in \Xi\colon \mathcal{F} \text{ is not equicontinuous at } \xi\}) = 0$ (Topsøe (1967)). Unless $\mathcal{F}$ is uniformly bounded, condition (2.14) cannot be valid for any sequence (P_n) that converges weakly to P. In that case, a uniform integrability condition with respect to the set $\{P_n\colon n \in \mathbb{N}\}$ has to be additionally imposed on $\mathcal{F}$. The set $\mathcal{F}$ is called *uniformly integrable* with respect to $\{P_n\colon n \in \mathbb{N}\}$ if

$$\lim_{R\to\infty} \sup_{n\in\mathbb{N}} \sup_{F\in\mathcal{F}} \int_{F(\xi)>R} |F(\xi)|\, \mathrm{d}P_n(\xi) = 0. \tag{2.15}$$

Note that condition (2.15) is satisfied if the moment condition

$$\sup_{n\in\mathbb{N}} \sup_{F\in\mathcal{F}} \int_{\Xi} |F(\xi)|^{1+\varepsilon}\, \mathrm{d}P_n(\xi) < \infty \tag{2.16}$$

holds for some $\varepsilon > 0$ (Section 5 in Billingsley (1968)). Then the condition (2.14) is valid for any sequence (P_n) that converges weakly to P in $\mathcal{P}_{\mathcal{F}}$ and has the property that $\mathcal{F}$ is uniformly integrable with respect to $\{P_n\colon n \in \mathbb{N}\}$ if the set $\mathcal{F}^R := \{[F]_R(\cdot) := \max\{-R, \min\{F(\cdot), R\}\}\colon F \in \mathcal{F}\}$ of truncated functions of $\mathcal{F}$ is a P-uniformity class for large $R > 0$. Since the class $\mathcal{F}^R$ is uniformly bounded, it is a P-uniformity class if $P(\{\xi \in \Xi\colon \mathcal{F}^R \text{ is not equicontinuous at } \xi\}) = 0$. Sufficient conditions for classes of characteristic functions of convex sets to be P-uniformity classes are mentioned in Example 2(b).

2.2 Qualitative stability

Together with the original stochastic programming problem (1.1) we consider a perturbation $Q \in \mathcal{P}(\Xi)$ of the probability distribution P and the perturbed model

$$\min\left\{\int_{\Xi} F_0(x,\xi)\,\mathrm{d}Q(\xi)\colon x \in X, \int_{\Xi} F_j(x,\xi)\,\mathrm{d}Q(\xi) \le 0, j = 1,\ldots,d\right\} \tag{2.17}$$

under the general assumptions imposed in Section 1. To fix our setting, let $\|\cdot\|$ denote the Euclidean norm and $\langle\cdot,\cdot\rangle$ the corresponding inner product. By $\mathbb{B}$ we denote the Euclidean unit ball and by $d(x,D)$ the distance of $x \in \mathbb{R}^m$ to the set $D \subset \mathbb{R}^m$. For any nonempty and open subset $\mathcal{U}$ of $\mathbb{R}^m$ we consider the following sets of functions, elements and probability measures

$$\mathcal{F}_{\mathcal{U}} := \{F_j(x,\cdot)\colon x \in X \cap \operatorname{cl}\mathcal{U}, j = 0,\ldots,d\},$$

$$\mathcal{X}_{\mathcal{U}}(Q) := \left\{x \in X \cap \operatorname{cl}\mathcal{U}\colon \int_{\Xi} F_j(x,\xi)\,\mathrm{d}Q(\xi) \le 0, j = 1,\ldots,d\right\} \quad (Q \in \mathcal{P}_{\mathcal{F}_{\mathcal{U}}}(\Xi)),$$

$$\mathcal{P}_{\mathcal{F}_{\mathcal{U}}}(\Xi) := \{Q \in \mathcal{P}(\Xi)\colon -\infty < \int_{\Xi} \inf_{x \in X \cap r\mathbb{B}} F_j(x,\xi)\,\mathrm{d}Q(\xi) \quad \text{for each } r > 0 \text{ and}$$

$$\sup_{x \in X \cap \operatorname{cl}\mathcal{U}} \int_{\Xi} F_j(x,\xi)\,\mathrm{d}Q(\xi) < \infty \quad \text{for each } j = 0,\ldots,d\},$$

and the pseudometric on $\mathcal{P}_{\mathcal{F}_{\mathcal{U}}} := \mathcal{P}_{\mathcal{F}_{\mathcal{U}}}(\Xi)$

$$d_{\mathcal{F}_{\mathcal{U}}}(P,Q) := \sup_{F \in \mathcal{F}_{\mathcal{U}}} \left|\int_{\Xi} F(\xi)(P-Q)(\mathrm{d}\xi)\right| = \sup_{\substack{j=0,\ldots,d \\ x \in X \cap \operatorname{cl}\mathcal{U}}} \left|\int_{\Xi} F_j(x,\xi)(P-Q)(\mathrm{d}\xi)\right|.$$

Thus, $d_{\mathcal{F}_{\mathcal{U}}}$ is a distance of probability measures having ζ-structure. It is nonnegative, symmetric and satisfies the triangle inequality (see also Section 2.1). Our general assumptions and the Fatou Lemma imply that the objective function and the constraint set of (2.17) are lower semicontinuous on X and closed in $\mathbb{R}^{\mathrm{m}}$, respectively, for each $Q \in \mathcal{P}_{\mathcal{F}_{\mathcal{U}}}(\Xi)$. Our first results provide further basic properties of the model (2.17).

Proposition 3. *Let $\mathcal{U}$ be a nonempty open subset of $\mathbb{R}^m$. Then the mapping $(x, Q) \mapsto \int_\Xi F_j(x, \xi)\, \mathrm{d}Q(\xi)$ from $(X \cap \mathrm{cl}\,\mathcal{U}) \times (\mathcal{P}_{\mathcal{F}_\mathcal{U}}, d_{\mathcal{F}_\mathcal{U}})$ to $\overline{\mathbb{R}}$ is sequentially lower semicontinuous for each $j = 0, \ldots, d$.*

Proof. Let $j = 0, \ldots, d$, $x \in X \cap \mathrm{cl}\,\mathcal{U}$, $Q \in \mathcal{P}_{\mathcal{F}_\mathcal{U}}$, (x_n) be a sequence in $X \cap \mathrm{cl}\,\mathcal{U}$ such that $x_n \to x$, and (Q_n) be a sequence converging to Q in $(\mathcal{P}_{\mathcal{F}_\mathcal{U}}, d_{\mathcal{F}_\mathcal{U}})$. Then the lower semicontinuity of $F_j(\cdot, \xi)$ for each $\xi \in \Xi$ and the Fatou Lemma imply the estimate

$$\begin{aligned}\int_\Xi F_j(x,\xi)\, \mathrm{d}Q(\xi) &\le \liminf_{n\to\infty} \int_\Xi F_j(x_n,\xi)\, \mathrm{d}Q(\xi)\\ &\le \liminf_{n\to\infty} \left\{ d_{\mathcal{F}_\mathcal{U}}(Q, Q_n) + \int_\Xi F_j(x_n, \xi) Q_n(\mathrm{d}\xi) \right\}\\ &= \liminf_{n\to\infty} \int_\Xi F_j(x_n, \xi) Q_n(\mathrm{d}\xi). \end{aligned}$$

□

Proposition 4. *Let $\mathcal{U}$ be a nonempty open subset of $\mathbb{R}^m$. Then the graph of the set-valued mapping $Q \mapsto \mathcal{X}_\mathcal{U}(Q)$ from $(\mathcal{P}_{\mathcal{F}_\mathcal{U}}, d_{\mathcal{F}_\mathcal{U}})$ into $\mathbb{R}^m$ is sequentially closed.*

Proof. Let (Q_n) be a sequence converging to Q in $(\mathcal{P}_{\mathcal{F}_\mathcal{U}}, d_{\mathcal{F}_\mathcal{U}})$ and (x_n) be a sequence converging to x in $\mathbb{R}^{\mathrm{m}}$ and such that $x_n \in \mathcal{X}_\mathcal{U}(Q_n)$ for each $n \in \mathbb{N}$. Clearly, we have $x \in X \cap \mathrm{cl}\,\mathcal{U}$. For $j \in \{1, \ldots, d\}$ we obtain from Proposition 3 that the estimate

$$\int_\Xi F_j(x, \xi)\, \mathrm{d}Q(\xi) \le \liminf_{n\to\infty} \int_\Xi F_j(x_n, \xi) Q_n\, (\mathrm{d}\xi) \le 0.$$

and, thus, $x \in \mathcal{X}_\mathcal{U}(Q)$ holds. □

To obtain perturbation results for (1.1), a stability property of the constraint set $\mathcal{X}(P)$ when perturbing the *probabilistic constraints* is needed. Consistently with the general definition of metric regularity for multifunctions (see, e.g., Rockafellar and Wets (1998)), we consider the set-valued mapping $y \mapsto \mathcal{X}_y(P)$ from $\mathbb{R}^d$ to $\mathbb{R}^m$, where

$$\mathcal{X}_y(P) = \left\{ x \in X \colon \int_\Xi F_j(x, \xi)\, \mathrm{d}P(\xi) \le y_j, j = 1, \ldots, d \right\},$$

and say that its inverse $x \mapsto \mathcal{X}_x^{-1}(P) = \{y \in \mathbb{R}^d \colon x \in \mathcal{X}_y(P)\}$ from $\mathbb{R}^m$ to $\mathbb{R}^d$ is *metrically regular* at some pair $(\overline{x}, 0) \in \mathbb{R}^m \times \mathbb{R}^d$ with $\overline{x} \in \mathcal{X}(P) = \mathcal{X}_0(P)$ if

there are constants $a \geq 0$ and $\varepsilon > 0$ such that it holds for all $x \in X$ and $y \in \mathbb{R}^d$ with $\|x - \overline{x}\| \leq \varepsilon$ and $\max_{j=1,\ldots,d} |y_j| \leq \varepsilon$ that

$$d(x, \mathcal{X}_y(P)) \leq a \max_{j=1,\ldots,d} \max \left\{ 0, \int_\Xi F_j(x, \xi)\, \mathrm{d}P(\xi) - y_j \right\}.$$

To state our results we will need localized versions of optimal values and solution sets. We follow the concept proposed in Robinson (1987) and Klatte (1987), and set for any nonempty open set $\mathcal{U} \subseteq \mathbb{R}^m$ and any $Q \in \mathcal{P}_{\mathcal{F}_\mathcal{U}}$

$$\vartheta_\mathcal{U}(Q) = \inf \left\{ \int_\Xi F_0(x, \xi)\, \mathrm{d}Q(\xi) \colon x \in \mathcal{X}_\mathcal{U}(Q) \right\},$$

$$X^*_\mathcal{U}(Q) = \left\{ x \in \mathcal{X}_\mathcal{U}(Q) \colon \int_\Xi F_0(x, \xi)\, \mathrm{d}Q(\xi) = \vartheta_\mathcal{U}(Q) \right\}.$$

A nonempty set $\mathcal{S} \subseteq \mathbb{R}^m$ is called a *complete local minimizing* (*CLM*) *set* of (2.17) relative to $\mathcal{U}$ if $\mathcal{U} \subseteq \mathbb{R}^m$ is open and $\mathcal{S} = X^*_\mathcal{U}(Q) \subset \mathcal{U}$. Clearly, CLM sets are sets of local minimizers, and the set $X^*(Q)$ of global minimizers is a CLM set with $X^*(Q) = X^*_\mathcal{U}(Q)$ if $X^*(Q) \subset \mathcal{U}$.

Now, we are ready to state the main qualitative stability result.

Theorem 5. *Let $P \in \mathcal{P}_{\mathcal{F}_\mathcal{U}}$ and assume that*

(i) *$X^*(P)$ is nonempty and $\mathcal{U} \subseteq \mathbb{R}^m$ is an open bounded neighbourhood of $X^*(P)$,*
(ii) *if $d \geq 1$, the function $x \mapsto \int_\Xi F_0(x, \xi)\, \mathrm{d}P(\xi)$ is Lipschitz continuous on $X \cap \mathrm{cl}\,\mathcal{U}$,*
(iii) *the mapping $x \mapsto \mathcal{X}_x^{-1}(P)$ is metrically regular at each pair $(\overline{x}, 0)$ with $\overline{x} \in X^*(P)$.*

Then the multifunction $X^*_\mathcal{U}$ from $(\mathcal{P}_{\mathcal{F}_\mathcal{U}}, d_{\mathcal{F}_\mathcal{U}})$ to $\mathbb{R}^m$ is upper semicontinuous at P, i.e., for any open set $\mathcal{O} \supseteq X^*_\mathcal{U}(P)$ it holds that $X^*_U(Q) \subseteq \mathcal{O}$ if $d_{\mathcal{F}_\mathcal{U}}(P,Q)$ is sufficiently small. Furthermore, there are positive constants L and δ such that

$$|\vartheta(P) - \vartheta_\mathcal{U}(Q)| \leq L d_{\mathcal{F}_\mathcal{U}}(P,Q) \tag{2.18}$$

holds and $X^*_\mathcal{U}(Q)$ is a CLM set of (2.17) relative to $\mathcal{U}$ whenever $Q \in \mathcal{P}_{\mathcal{F}_\mathcal{U}}$ and $d_{\mathcal{F}_\mathcal{U}}(P,Q) < \delta$. In case $d = 0$, the estimate (2.18) is valid with $L = 1$ and for any $Q \in \mathcal{P}_{\mathcal{F}_\mathcal{U}}$.

Proof. We consider the (localized) parametric optimization problem

$$\min\left\{f(x,Q)=\int_{\Xi}F_0(x,\xi)\,\mathrm{d}Q(\xi)\colon x\in\mathcal{X}_U(Q)\right\},$$

where the probability measure Q is regarded as a parameter varying in the pseudometric space $(\mathcal{P}_{\mathcal{F}_U},d_{\mathcal{F}_U})$. Proposition 4 says that the graph of the multifunction $\mathcal{X}_{\mathcal{U}}$ from $\mathcal{P}_{\mathcal{F}_U}$ to $\mathbb{R}^m$ is sequentially closed. Hence, $\mathcal{X}_{\mathcal{U}}$ is upper semicontinuous on $\mathcal{P}_{\mathcal{F}_U}$, since $\mathrm{cl}\,\mathcal{U}$ is compact. Furthermore, we know by Proposition 3 that the function f from $(X\cap\mathrm{cl}\,\mathcal{U})\times\mathcal{P}_{\mathcal{F}_U}$ to $\overline{\mathbb{R}}$ is sequentially lower semicontinuous and finite. Let us first consider the case of $d=0$. Since $f(\cdot,Q)$ is lower semicontinuous, $X^*_{\mathcal{U}}(Q)$ is nonempty for each $Q\in\mathcal{P}_{\mathcal{F}_U}$. Let $x_*\in X^*(P)$, $Q\in\mathcal{P}_{\mathcal{F}_U}$ and $\tilde{x}\in X^*_{\mathcal{U}}(Q)$. Then the estimate

$$\begin{aligned}|\vartheta(P)-\vartheta_{\mathcal{U}}(Q)|&\le\max\left\{\int_{\Xi}F_0(x_*,\xi)(Q-P)(\mathrm{d}\xi),\int_{\Xi}F_0(\tilde{x},\xi)(P-Q)(\mathrm{d}\xi)\right\}\\&\le d_{\mathcal{F}_U}(P,Q)\end{aligned}$$

holds. This implies that the multifunction $X^*_{\mathcal{U}}$ from $(\mathcal{P}_{\mathcal{F}_U},d_{\mathcal{F}_U})$ to $\mathbb{R}^m$ is closed at P and, thus, upper semicontinuous at P.

In case $d\ge 1$, condition (ii) implies that the function f is even continuous on $(X\cap\mathrm{cl}\,\mathcal{U})\times\mathcal{P}_{\mathcal{F}_U}$. Then we use Berge's classical stability analysis (see Berge (1963) for topological parameter spaces and Theorem 4.2.1 in Bank et al. (1982) for metric parameter spaces) and conclude that $X^*_{\mathcal{U}}$ is upper semicontinuous at P if $\mathcal{X}_{\mathcal{U}}$ satisfies the following (lower semicontinuity) property at some pair $(\overline{x},P)$ with $\overline{x}\in X^*(P)$:

$$\mathcal{X}_{\mathcal{U}}(P)\cap B(\overline{x},\overline{\varepsilon})\subseteq\mathcal{X}_U(Q)+a\,d_{\mathcal{F}_U}(P,Q)\mathbb{B}\ \text{ whenever }\ d_{\mathcal{F}_U}(P,Q)<\overline{\varepsilon},\qquad(2.19)$$

where $a\ge 0$ is the corresponding constant in condition (iii), and $\overline{\varepsilon}>0$ is sufficiently small. To establish property (2.19), let $\overline{x}\in X^*(P)$, and $a=a(\overline{x})\ge 0$, $\varepsilon=\varepsilon(\overline{x})>0$ be the metric regularity constants from (iii). First we observe that the estimate $\int_{\Xi}F_j(x,\xi)(Q-P)(\mathrm{d}\xi)\le d_{\mathcal{F}_U}(P,Q)$ holds for any $x\in X\cap\mathrm{cl}\,\mathcal{U}$, $j\in\{1,\dots,d\}$ and $Q\in\mathcal{P}_{\mathcal{F}_U}$. Next we choose $\overline{\varepsilon}=\overline{\varepsilon}(\overline{x})$ such that $0<\overline{\varepsilon}<\varepsilon$ and $\overline{x}+(a+1)\overline{\varepsilon}\,\mathbb{B}\subseteq\mathcal{U}$. Hence, we have $x+a\overline{\varepsilon}\,\mathbb{B}\subseteq\mathcal{U}$ for any $x\in\overline{x}+\overline{\varepsilon}\,\mathbb{B}$. Let $Q\in\mathcal{P}_{\mathcal{F}_U}$ be such that $d_{\mathcal{F}_U}(P,Q)<\overline{\varepsilon}$. Putting $y_j=-d_{\mathcal{F}_U}(P,Q)$, $j=1,\dots,d$, the above estimate implies that $\mathcal{X}_y(P)\cap\mathrm{cl}\,\mathcal{U}\subseteq\mathcal{X}_{\mathcal{U}}(Q)$. Due to the choice of $\overline{\varepsilon}$ we have $d(x,\mathcal{X}_y(P)\cap\mathrm{cl}\,\mathcal{U})=d(x,\mathcal{X}_y(P))$ for any

$x \in \mathcal{X}_{\mathcal{U}}(P) \cap (\overline{x} + \overline{\varepsilon}\,\mathbb{B})$, and, hence, the metric regularity condition (iii) yields the estimate

$$\begin{aligned} d(x, \mathcal{X}_{\mathcal{U}}(Q)) &\le d(x, \mathcal{X}_y(P) \cap \operatorname{cl} \mathcal{U}) = d(x, \mathcal{X}_y(P)) \\ &\le a \max_{j=1,\dots,d} \max \left\{ 0, \int_{\Xi} F_j(x,\xi)\, \mathrm{d}P(\xi) + d_{\mathcal{F}_{\mathcal{U}}}(P,Q) \right\} \\ &\le a\, d_{\mathcal{F}_{\mathcal{U}}}(P,Q), \end{aligned}$$

which is equivalent to the property (2.19). Hence, $X^*_{\mathcal{U}}$ is upper semicontinuous at P and there exists a constant $\hat{\delta} > 0$ such that $X^*_{\mathcal{U}}(Q) \subset \mathcal{U}$ for any $Q \in \mathcal{P}_{\mathcal{F}_{\mathcal{U}}}$ with $d_{\mathcal{F}_{\mathcal{U}}}(P,Q) < \hat{\delta}$. Thus $X^*_{\mathcal{U}}(Q)$ is a CLM set of (2.17) relative to $\mathcal{U}$ for each such Q.

Moreover, for any $x \in \mathcal{X}_{\mathcal{U}}(Q) \cap (\overline{x} + \overline{\varepsilon}\mathbb{B})$ (iii) implies the estimate

$$\begin{aligned} d(x,\mathcal{X}_{\mathcal{U}}(P)) &= d(x,\mathcal{X}_0(P) \cap \operatorname{cl} \mathcal{U}) = d(x, \mathcal{X}_0(P)) \\ &\le a \max_{j=1,\dots,d} \max \left\{ 0, \int_{\Xi} F_j(x,\xi)\, \mathrm{d}P(\xi) \right\} \\ &\le a \max_{j=1,\dots,d} \max \left\{ 0, \int_{\Xi} F_j(x,\xi)\, \mathrm{d}P(\xi) - \int_{\Xi} F_j(x,\xi)\, \mathrm{d}Q(\xi) \right\} \\ &\le a\, d_{\mathcal{F}_{\mathcal{U}}}(P,Q)\,, \end{aligned}$$

which is equivalent to the inclusion

$$\mathcal{X}_{\mathcal{U}}(Q) \cap (\overline{x} + \overline{\varepsilon}\ \mathbb{B}) \subseteq \mathcal{X}_{\mathcal{U}}(P) + a\, d_{\mathcal{F}_{\mathcal{U}}}(P,Q)\mathbb{B}.$$

Since $X^*(P)$ is compact, we employ a finite covering argument and arrive at two analogues of both inclusions, where a neighbourhood $\mathcal{N}$ of $X^*(P)$ appears instead of the balls $\overline{x} + \overline{\varepsilon}\,\mathbb{B}$ in their left-hand sides, and a uniform constant $\hat{a}$ appears instead of a in their right-hand sides. Moreover, there exists a uniform constant $\hat{\varepsilon} > 0$ such that the (new) inclusions are valid whenever $d_{\mathcal{F}_{\mathcal{U}}}(P,Q) < \hat{\varepsilon}$. Now, we choose $\delta > 0$ such that $\delta \le \min\{\hat{\delta},\hat{\varepsilon}\}$ and $X^*_{\mathcal{U}}(Q) \subset \mathcal{N}$ whenever $d_{\mathcal{F}_{\mathcal{U}}}(P,Q) < \delta$.

Let $Q \in \mathcal{P}_{\mathcal{F}_\mathcal{U}}$ be such that $d_{\mathcal{F}_\mathcal{U}}(P,Q) < \delta$ and $\tilde{x} \in \mathcal{X}^*_\mathcal{U}(Q) \subseteq \mathcal{X}_\mathcal{U}(Q) \cap \mathcal{N}$. Then there exists an element $\overline{x} \in \mathcal{X}_\mathcal{U}(P)$ satisfying $\|\tilde{x} - \overline{x}\| \leq \hat{a}\, d_{\mathcal{F}_\mathcal{U}}(P,Q)$. We obtain

$$\begin{aligned}\vartheta(P) \leq f(\overline{x},P) &\leq f(\tilde{x},Q) + |f(\overline{x},P) - f(\tilde{x},Q)| \\ &\leq \vartheta_\mathcal{U}(Q) + |f(\overline{x},P) - f(\tilde{x},P)| + |f(\tilde{x},P) - f(\tilde{x},Q)| \\ &\leq \vartheta_\mathcal{U}(Q) + L_f\|\overline{x} - \tilde{x}\| + d_{\mathcal{F}_\mathcal{U}}(P,Q) \\ &\leq \vartheta_\mathcal{U}(Q) + (L_f\hat{a} + 1)d_{\mathcal{F}_\mathcal{U}}(P,Q)\,,\end{aligned}$$

where $L_f \geq 0$ denotes a Lipschitz constant of $f(\cdot, P)$ on $X \cap \mathrm{cl}\,\mathcal{U}$. For the converse estimate, let $\overline{x} \in X^*(P)$ and $Q \in \mathcal{P}_{\mathcal{F}_\mathcal{U}}$ be such that $d_{\mathcal{F}_\mathcal{U}}(P,Q) < \delta$. Then there exists $\tilde{x} \in \mathcal{X}_\mathcal{U}(Q)$ such that $\|\tilde{x} - \overline{x}\| \leq \hat{a}\, d_{\mathcal{F}_\mathcal{U}}(P,Q)$. We conclude

$$\vartheta_\mathcal{U}(Q) \leq f(\tilde{x}, Q) \leq \vartheta(P) + |f(\tilde{x}, Q) - f(\overline{x},P)|$$

and arrive analogously at the desired continuity property of ϑ_U by putting $L = L_f\hat{a} + 1$. □

The above proof partly parallels arguments in Klatte (1987). The most restrictive requirement in the above result is the metric regularity condition (iii). Example 40 in Section 3.3 provides some insight into the necessity of condition (iii) in the context of chance constrained models. Criteria for the metric regularity of multifunctions are given e.g., in Section 9G of Rockafellar and Wets (1998) and in Mordukhovich (1994b). Here, we do not intend to provide a specific sufficient condition for (iii), but recall that the constraint functions $\int_\Xi F_j(\cdot,\xi)\,\mathrm{d}P(\xi)$ $(j = 1,\ldots,d)$ are often nondifferentiable or even discontinuous in stochastic programming. In Section 3.3 we show how metric regularity is verified in case of chance constrained programs.

Although, Theorem 5 also asserts a quantitative continuity property for optimal values, its essence consists in a continuity result for optimal values and solution sets. As a first conclusion we derive consequences for the stability of (1.1) with respect to the weak convergence of probability measures (cf. Section 2.1). To state our main stability result for (1.1) with respect to the topology of weak convergence, we need the classes $\mathcal{F}^R_\mathcal{U}$ of truncated functions of $\mathcal{F}_\mathcal{U}$ for $R > 0$ and the uniform integrability property of $\mathcal{F}_\mathcal{U}$ (see Section 2.1).

Theorem 6. *Let the assumptions of Theorem 5 for (1.1) be satisfied. Furthermore, let $\mathcal{F}^R_\mathcal{U}$ be a P-uniformity class for large $R > 0$ and (P_n) be a sequence in $\mathcal{P}_{\mathcal{F}_\mathcal{U}}$ that is weakly convergent to P.*

Then the sequence $(\vartheta_\mathcal{U}(P_n))$ converges to $\vartheta(P)$, the sets $X^*_\mathcal{U}(P_n)$ are CLM sets relative to $\mathcal{U}$ for sufficiently large $n \in \mathbb{N}$ and

$$\lim_{n\to\infty} \sup_{x \in X^*_\mathcal{U}(P_n)} d(x, X^*(P))) = 0$$

holds if $\mathcal{F}_\mathcal{U}$ is uniformly integrable with respect to $\{P_n : n \in \mathbb{N}\}$.

Proof. Let (P_n) be a sequence in $\mathcal{P}_{\mathcal{F}_\mathcal{U}}$ that converges weakly to P and has the property that $\mathcal{F}_\mathcal{U}$ is uniformly integrable with respect to $\{P_n : n \in \mathbb{N}\}$. Then the assumption implies (see Section 2.1)

$$\lim_{n\to\infty} d_{\mathcal{F}_\mathcal{U}}(P_n, P) = 0$$

and, hence, the result is an immediate consequence of Theorem 5. □

Compared to Theorem 5, the stability of (1.1) with respect to weakly convergent perturbations of P requires additional conditions on $\mathcal{F}_\mathcal{U}$. The previous theorem provides the sufficient conditions that its truncated class $\mathcal{F}_\mathcal{U}^R$ has the P-uniformity property for large $R > 0$ and that $\mathcal{F}_\mathcal{U}$ is uniformly integrable with respect to the set of perturbations. The first condition is satisfied if $\mathcal{F}_\mathcal{U}^R$ is P-almost surely equicontinuous on Ξ (cf. Section 2.1). It implies, in particular, the P-continuity of $F_j(x,\cdot)$ for each $j = 0, \ldots, d$ and $x \in X \cap \operatorname{cl}\mathcal{U}$. The uniform integrability condition

$$\lim_{R\to\infty} \sup_{n\in\mathbb{N}} \max_{j=0,\ldots,d} \sup_{x\in X\cap \operatorname{cl}\mathcal{U}} \int_{|F_j(x,\xi)|>R} |F_j(x,\xi)| \mathrm{d}P_n(\xi) = 0 \tag{2.20}$$

is satisfied if the moment condition

$$\sup_{n\in\mathbb{N}} \max_{j=0,\ldots,d} \sup_{x\in X\cap \operatorname{cl}\mathcal{U}} \int_{\Xi} |F_j(x,\xi)|^{1+\varepsilon} \mathrm{d}P_n(\xi) < \infty \tag{2.21}$$

holds for some $\varepsilon > 0$. Assume, for example, that the functions F_j satisfy an estimate of the form

$$|F_j(x,\xi)| \le C\|\xi\|^k, \quad \forall (x,\xi) \in (X \cap \operatorname{cl}\mathcal{U}) \times \Xi,$$

for some positive constants C, k and all $j = 0, \ldots, d$ (see e.g., Sections 3.1 and 3.2). In this case, the uniform integrability condition (2.20) is satisfied if

$$\lim_{R\to\infty} \sup_{n\in\mathbb{N}} \int_{\|\xi\|>R} \|\xi\|^k \mathrm{d}P_n(\xi) = 0.$$

The corresponding sufficient moment condition reads

$$\sup_{n\in\mathbb{N}} \int_{\Xi} \|\xi\|^{k+\varepsilon} \mathrm{d}P_n(\xi) < \infty$$

for some $\varepsilon > 0$. The latter condition is often imposed in stability studies with respect to weak convergence.

The P-continuity property of each function $F_j(x,\cdot)$ and condition (2.20) are not needed in Theorem 5. However, the following examples show that both conditions are indispensable for stability with respect to weak convergence.

Example 7. Let $m = s = 1$, $d = 0$, $\Xi = \mathbb{R}$, $X = \mathbb{R}_-$, $F_0(x,\xi) = -\chi_{(-\infty,x]}(\xi)$ for $(x,\xi) \in \mathbb{R} \times \Xi$ and $P = \delta_0$, where δ_ξ denotes the measure that places unit mass at ξ. Then $\vartheta(P) = 1$ and $X^*(P) = \{0\}$. The sequence $(\delta_{\frac{1}{n}})$ converges weakly to P in $\mathcal{P}(\Xi)$, but it holds that $\vartheta(P_n) = 0$ for each $n \in \mathbb{N}$. This is due to the fact that, for some neighbourhood $\mathcal{U}$ of 0, the set $\{\chi_{(-\infty,x]}(\cdot)\colon x \in X \cap \mathrm{cl}\,\mathcal{U}\}$ is not a P-uniformity class since $P(\mathrm{bd}\,(-\infty,0]) = P(\{0\}) = 1$.

Example 8. Let $m = s = 1$, $d = 0$, $\Xi = \mathbb{R}_+$, $X = [-1,1]$, $F_0(x,\xi) = \max\{\xi - x,0\}$ for $(x,\xi) \in \mathbb{R} \times \Xi$ and $P = \delta_0$. Then $\vartheta(P) = 0$ and $X^*(P) = [0,1]$. Consider the sequence $P_n = (1 - \frac{1}{n})\delta_0 + \frac{1}{n}\delta_n$, $n \in \mathbb{N}$, which converges weakly to P. It holds that $\vartheta(P_n) = 1 - \frac{1}{n}$ and $X^*(P_n) = \{1\}$ for each $n \in \mathbb{N}$ and, thus, $(\vartheta(P_n))$ does not converge to $\vartheta(P)$. Here, the reason is that the class $\{\max\{\cdot - x, 0\}\colon x \in [-1,1]\}$ is not uniformly integrable with respect to $\{P_n\colon n \in \mathbb{N}\}$.

Indeed, the weak convergence of measures is a very weak condition on sequences and, hence, requires strong conditions on (1.1) to be stable. Many approximations of P (e.g., in Section 4.1), however, have much stronger properties than weak convergence and, hence, work under weaker assumptions than Theorem 6. To give an example, we recall that the P-continuity property of each function $F_j(x,\cdot)$ is an indispensable assumption in case of stability with respect to weak convergence, but this property is not needed when working with $d_{\mathcal{F}_\mathcal{U}}$ and with specifically adjusted ideal metrics (and the corresponding convergences of measures) in case of (mixed-integer) two-stage and chance constrained models (see Sections 3.1, 3.2 and 3.3). Consequently, we prefer to work with these distances, having in mind their relations to the topology of weak convergence.

2.3 *Quantitative stability*

The main result in the previous section claims that the multifunction $X^*_\mathcal{U}(\cdot)$ is nonempty near P and upper semicontinuous at P. In order to quantify the upper semicontinuity property, a growth condition on the objective function in a neighbourhood of the solution set to the original problem (1.1) is needed. Instead of imposing a specific growth condition (as e.g. quadratic growth), we consider the *growth function* ψ_P defined on $\mathbb{R}_+$ by

$$\psi_P(\tau) := \min\left\{\int_\Xi F_0(x,\xi)\,\mathrm{d}P(\xi) - \vartheta(P)\colon d(x, X^*(P)) \geq \tau,\ x \in \mathcal{X}_\mathcal{U}(P)\right\} \tag{2.22}$$

of problem (1.1) on $\mathrm{cl}\mathcal{U}$, i.e., near its solution set $X^*(P)$, and the associated function

$$\Psi_P(\eta) := \eta + \psi_P^{-1}(2\eta) \quad (\eta \in \mathbb{R}_+), \tag{2.23}$$

where we set $\psi_P^{-1}(t) := \sup\{\tau \in \mathbb{R}_+ : \psi_P(\tau) \le t\}$. Both functions, ψ_P and Ψ_P, depend on the data of (1.1) and, in particular, on P. They are lower semicontinuous on $\mathbb{R}_+$; ψ_P is nondecreasing, Ψ_P increasing and both vanish at 0 (cf. Theorem 7.64 in Rockafellar and Wets (1998)). The second main stability result establishes a quantitative upper semicontinuity property of (localized) solution sets and identifies the function Ψ_P as modulus of semicontinuity. In the convex case, it also provides continuity moduli of countable dense families of selections to solution sets.

Theorem 9. *Let the assumptions of* Theorem 5 *be satisfied and* $P \in \mathcal{P}_{\mathcal{F}_\mathcal{U}}$. *Then there exists a constant* $\hat{L} \ge 1$ *such that*

$$\emptyset \ne X_\mathcal{U}^*(Q) \subseteq X^*(P) + \Psi_P(\hat{L} d_{\mathcal{F}_\mathcal{U}}(P, Q))\mathbb{B} \tag{2.24}$$

holds for any $Q \in \mathcal{P}_{\mathcal{F}_\mathcal{U}}$ *with* $d_{\mathcal{F}_\mathcal{U}}(P, Q) < \delta$. *Here,* δ *is the constant in* Theorem 5 *and* Ψ_P *is given by* (2.23). *In case* $d = 0$, *the estimate* (2.24) *is valid with* $\hat{L} = 1$ *and for any* $Q \in \mathcal{P}_{\mathcal{F}_\mathcal{U}}$.

Proof. Let $L > 0$, $\delta > 0$ be the constants in Theorem 5, $Q \in \mathcal{P}_{\mathcal{F}_\mathcal{U}}$ with $d_{\mathcal{F}_\mathcal{U}}(P, Q) < \delta$ and $\tilde{x} \in X_\mathcal{U}^*(Q)$. As argued in the proof of Theorem 5, there exists an element $\overline{x} \in \mathcal{X}_\mathcal{U}(P)$ such that $\|\tilde{x} - \overline{x}\| \le \hat{a}\overline{\delta}$, where $\overline{\delta} := d_{\mathcal{F}_\mathcal{U}}(P, Q)$. Let $L_P \ge 0$ denote a Lipschitz constant of the function $x \mapsto \int_\Xi F_0(x,\xi)\,\mathrm{d}P(\xi)$ on $X \cap \mathrm{cl}\mathcal{U}$. Then the definition of ψ_P and Theorem 5 imply that

$$\begin{aligned}
\overline{\delta}(1 + L_P\hat{a} + L) &\ge \overline{\delta}(1 + L_P\hat{a}) + \vartheta_\mathcal{U}(Q) - \vartheta(P) \\
&= \overline{\delta}(1 + L_P\hat{a}) + \int_\Xi F_0(\tilde{x}, \xi)\,\mathrm{d}Q(\xi) - \vartheta(P) \\
&\ge \overline{\delta} L_P\hat{a} + \int_\Xi F_0(\tilde{x}, \xi)\,\mathrm{d}P(\xi) - \vartheta(P) \\
&\ge \int_\Xi F_0(\overline{x}, \xi)\,\mathrm{d}P(\xi) - \vartheta(P) \ge \psi_P(d(\overline{x}, X^*(P))) \\
&\ge \inf_{y \in \tilde{x} + \hat{a}\overline{\delta}\mathbb{B}} \psi_P(d(y, X^*(P))) = \psi_P(d(\tilde{x}, X^*(P) + \hat{a}\overline{\delta}\mathbb{B})).
\end{aligned}$$

Hence, we obtain

$$d(\tilde{x}, X^*(P)) \le \hat{a}\overline{\delta} + d(\tilde{x}, X^*(P) + \hat{a}\overline{\delta}\mathbb{B})$$
$$\le \hat{a}\overline{\delta} + \psi_P^{-1}(\overline{\delta}(1 + L_P\hat{a} + L)) \le \hat{L}\overline{\delta} + \psi_P^{-1}(2\hat{L}\overline{\delta}) = \Psi_P(\hat{L}\overline{\delta}),$$

where $\hat{L} := \max\{\hat{a}, \frac{1}{2}(1 + L_P\hat{a} + L)\} \ge 1$. In case $d = 0$, we may choose $\hat{x} = \tilde{x}$, $\hat{a} = 1$, $L = 1$, $L_P = 0$ and an arbitrary δ. This completes the proof. □

Parts of the proof are similar to arguments of Theorem 7.64 in Rockafellar and Wets (1998). Next, we briefly comment on some aspects of the general stability theorems, namely, specific growth conditions and localization issues.

Remark 10. Problem (1.1) is said to have *k-th order growth* at the solution set for some $k \ge 1$ if $\psi_P(\tau) \ge \gamma\tau^k$ for each small $\tau \in \mathbb{R}_+$ and some $\gamma > 0$, i.e., if

$$\int_\Xi F_0(x, \xi)\, \mathrm{d}P(\xi) \ge \vartheta(P) + \gamma d(x, X^*(P))^k$$

holds for each feasible x close to $X^*(P)$. Then $\Psi_P(\eta) \le \eta + (2\eta/\gamma)^{\frac{1}{k}} \le C\eta^{\frac{1}{k}}$ for some constant $C > 0$ and sufficiently small $\eta \in \mathbb{R}_+$. In this case, Theorem 9 provides the Hölder continuity of $X^*_{\mathcal{U}}$ at P with rate $\frac{1}{k}$. Important special cases are the linear and quadratic growth for $k = 1$ and $k = 2$, respectively.

Remark 11. In the Theorems 5 and 9 the localized optimal values $\vartheta_{\mathcal{U}}(Q)$ and solution sets $X^*_{\mathcal{U}}(Q)$ of the (perturbed) model (2.17) may be replaced by their global versions $\vartheta(Q)$ and $X^*(Q)$ if there exists a constant $\delta_0 > 0$ such that for each $Q \in \mathcal{P}_{\mathcal{F}_{\mathcal{U}}}$ with $d_{\mathcal{F}_{\mathcal{U}}}(P, Q) < \delta_0$ either of the following conditions is satisfied: (a) The model (2.17) is convex and $X^*_{\mathcal{U}}(Q)$ is a CLM set, (b) the constraint set of (2.17) is contained in some bounded set $\mathcal{V} \subset \mathbb{R}^m$ not depending on Q, and it holds that $\mathcal{V} \subseteq \mathcal{U}$.

In case of a fixed constraint set, i.e., $d = 0$, we derive an extension of Theorem 9 by using a probability distance that is based on divided differences of the functions $x \mapsto \int_\Xi F_0(x, \xi)d(P - Q)(\xi)$ around the solution set of (1.1). For some nonempty, bounded, open subset $\mathcal{U}$ of $\mathbb{R}^m$ we consider the following set of probability measures

$$\hat{\mathcal{P}}_{\mathcal{F}_{\mathcal{U}}} := \Big\{ Q \in \mathcal{P}_{\mathcal{F}_{\mathcal{U}}} : \exists C_Q > 0 \quad \text{such that} \int_\Xi \frac{F_0(x, \xi) - F_0(\overline{x}, \xi)}{\|x - \overline{x}\|}\, \mathrm{d}Q(\xi) \le C_Q,$$
$$\forall x, \overline{x} \in X \cap \operatorname{cl}\mathcal{U},\ x \ne \overline{x} \Big\}$$

and the distance

$$\hat{d}_{\mathcal{F}_\mathcal{U}}(P, Q) := \sup\left\{\int_\Xi \frac{F_0(x, \xi)-F_0(\overline{x}, \xi)}{\|x-\overline{x}\|} d(P-Q)(\xi) : x, \overline{x} \in X \cap \operatorname{cl} \mathcal{U},\ x \neq \overline{x}\right\}$$

which is well defined and finite on $\hat{\mathcal{P}}_{\mathcal{F}_\mathcal{U}}$. The following result has been inspired by Section 4.4.1 in Bonnans and Shapiro (2000).

Theorem 12. *Let $d = 0$, $P \in \hat{\mathcal{P}}_{\mathcal{F}_\mathcal{U}}$, $X^*(P)$ be nonempty and $\mathcal{U} \subset \mathbb{R}^m$ be a bounded and open neighbourhood of $X^*(P)$. Then the estimate*

$$\sup_{x \in X^*_\mathcal{U}(Q)} d(x, X^*(P)) \leq (\psi_P^r)^{-1}(\hat{d}_{\mathcal{F}_\mathcal{U}}(P, Q))$$

is valid for any $Q \in \hat{\mathcal{P}}_{\mathcal{F}_\mathcal{U}}$, where $\psi_P^r(0) = 0$, $\psi_P^r(\tau) := \frac{\psi_P(\tau)}{\tau}$ for each $\tau > 0$ and $\psi_P(\cdot)$ is the growth function given by (2.22).

If, moreover, $(\psi_P^r)^{-1}$ is continuous at $\tau = 0$, there exists a constant $\delta > 0$ such that $X^*_\mathcal{U}(Q)$ is a CLM set relative to $\mathcal{U}$ whenever $\hat{d}_{\mathcal{F}_\mathcal{U}}(P, Q) < \delta$.

If, in particular, the original problem (1.1) has quadratic growth, i.e., $\psi_P(\tau) \geq \gamma\tau^2$ for some $\gamma > 0$, there exists a constant $\delta > 0$ such that the inclusion

$$\emptyset \neq X^*_\mathcal{U}(Q) \subseteq X^*(P) + \frac{1}{\gamma}\hat{d}_{\mathcal{F}_\mathcal{U}}(P, Q)\mathbb{B}$$

holds whenever $\hat{d}_{\mathcal{F}_\mathcal{U}}(P, Q) < \delta$.

Proof. Let $Q \in \hat{\mathcal{P}}_{\mathcal{F}_\mathcal{U}}$, $x \in X^*_\mathcal{U}(Q)$ and $\overline{x} \in X^*(P)$ be such that $\|x - \overline{x}\| = d(x, X^*(P)) > 0$. We denote $f_Q(y) := \int_\Xi F_0(y, \xi)\, \mathrm{d}Q(\xi)$ for each $y \in X$, and have $f_Q(x) \leq f_Q(\overline{x})$ and $f_P(x) - f_P(\overline{x}) \geq \psi_P(d(x, X^*(P))) = \psi_P(\|x - \overline{x}\|)$. This leads to the following estimate

$$\begin{aligned}
\psi_P^r(\|x-\overline{x}\|) &= \frac{1}{\|x-\overline{x}\|}\psi_P(\|x-\overline{x}\|) \leq \frac{1}{\|x-\overline{x}\|}(f_P(x) - f_P(\overline{x})) \\
&\leq \frac{1}{\|x-\overline{x}\|}(f_P(x) - f_Q(x) + f_Q(\overline{x}) - f_P(\overline{x})) \\
&= \frac{1}{\|x-\overline{x}\|}((f_P - f_Q)(x) - (f_P - f_Q)(\overline{x})) \\
&\leq \hat{d}_{\mathcal{F}_\mathcal{U}}(P, Q),
\end{aligned}$$

which completes the first part. Since $\mathcal{U}$ is open, there exists an $\varepsilon > 0$ such that the ε-enlargement $\{x \in \mathbb{R}^m : d(x, X^*(P)) \leq \varepsilon\}$ of $X^*(P)$ is contained in $\mathcal{U}$.

Let $\delta > 0$ be chosen such that $(\psi_P^r)^{-1}(\delta) \le \varepsilon$. Then $d(x, X^*(P)) \le \varepsilon$ and, thus, $x \in \mathcal{U}$ holds for each $x \in X_{\mathcal{U}}^*(Q)$, completing the second part.

Finally, it remains to remark that quadratic growth implies $\psi_P^r(\tau) \ge \gamma\tau$ for any $\tau > 0$ and some $\gamma > 0$. □

Compared to the estimate in Theorem 9 based on function values of the function F_0, the above bound uses divided difference information of F_0 relative to x and leads to Lipschitz-type results in case of quadratic growth.

While the growth behaviour of the objective function is important for the quantitative stability of solution sets even for convex models, the situation is much more advantageous for ε-approximate solution sets. For convex models (1.1) with a fixed constraint set (i.e., $d = 0$), we will see that the latter sets behave Lipschitz continuously with respect to changes of probability distributions measured in terms of the distance $d_{\mathcal{F}_{\mathcal{U}}}$, but for a larger set $\mathcal{U}$ compared with stability results for solution sets. To state the result, let

$$\mathbb{D}_\rho(C,D) := \inf\{\eta \ge 0 \colon C \cap \rho\mathbb{B} \subset D + \eta\mathbb{B},\ D \cap \rho\mathbb{B} \subset C + \eta\mathbb{B}\} \tag{2.25}$$

$$\mathbb{D}_\infty(C,D) := \inf\{\eta \ge 0 \colon C \subset D + \eta\mathbb{B},\ D \subset C + \eta\mathbb{B}\} \tag{2.26}$$

denote the ρ-distance ($\rho \ge 0$) and the Pompeiu–Hausdorff distance, respectively, of nonempty closed subsets C, D of $\mathbb{R}^m$.

Theorem 13. *Let $d = 0$, F_0 be a random lower semicontinuous convex function, X be closed convex, $P \in \mathcal{P}_{\mathcal{F}_{\mathcal{U}}}$ and $X^*(P)$ be nonempty and bounded. Then there exist constants $\rho > 0$ and $\overline{\varepsilon} > 0$ such that the estimate*

$$\mathbb{D}_\infty(X_\varepsilon^*(P), X_\varepsilon^*(Q)) \le \frac{2\rho}{\varepsilon} d_{\mathcal{F}_{\mathcal{U}}}(P, Q)$$

holds for $\mathcal{U} := (\rho + \overline{\varepsilon})\mathbb{B}$ and any $\varepsilon \in (0, \overline{\varepsilon})$, $Q \in \mathcal{P}_{\mathcal{F}_{\mathcal{U}}}$ such that $d_{\mathcal{F}_{\mathcal{U}}}(P, Q) < \varepsilon$.

Proof. First we choose $\rho_0 > 0$ such that $X^*(P)$ is contained in the open ball $\mathcal{U}_{\rho_0}$ around the origin in $\mathbb{R}^m$ with radius ρ_0 and that $\vartheta(P) \ge -\rho_0 + 1$. Applying Theorem 5 with $\mathcal{U}_{\rho_0}$ as the bounded open neighbourhood of $X^*(P)$, we obtain some constant $\varepsilon_0 > 0$ such that $X^*(Q)$ is nonempty and contained in $\mathcal{U}_{\rho_0}$ and $\vartheta(Q) \ge \rho_0$ holds whenever $Q \in \mathcal{P}_{\mathcal{F}_{\mathcal{U}_{\rho_0}}}$ and $d_{\mathcal{F}_{\mathcal{U}_{\rho_0}}}(P,Q) < \varepsilon_0$. Now, let $\rho > \rho_0$, $\overline{\varepsilon} := \min\{\varepsilon_0, \rho - \rho_0, 1\}$ and $\mathcal{U} := (\rho + \overline{\varepsilon})\mathbb{B}$.

For any $Q \in \mathcal{P}_{\mathcal{F}_{\mathcal{U}}}$ we set again $f_Q(x) := \int_\Xi F_0(x,\xi)\,\mathrm{d}Q(\xi)$ for each $x \in \mathbb{R}^m$. Furthermore, we denote by $\hat{d}_\rho^+$ the auxiliary epi-distance of f_P and f_Q introduced in Proposition 7.61 in Rockafellar and Wets (1998):

$$\hat{d}_\rho^+(f_P, f_Q) := \inf\{\eta \ge 0 \colon \inf_{y \in x+\eta\mathbb{B}} f_Q(y) \le \max\{f_P(x), -\rho\} + \eta,\ \inf_{y \in x+\eta\mathbb{B}} f_P(y) \le \max\{f_Q(x), -\rho\} + \eta,\ \forall x \in \rho\mathbb{B}\}.$$

From Theorem 7.69 in Rockafellar and Wets (1998) we conclude that the estimate

$$\mathbb{D}_\rho(X_\varepsilon^*(P), X_\varepsilon^*(Q)) \le \frac{2\rho}{\varepsilon}\hat{d}^+_{\rho+\varepsilon}(f_P, f_Q)$$

is valid for $\varepsilon \in (0, \bar{\varepsilon})$ if $\hat{d}^+_{\rho+\varepsilon}(f_P, f_Q) < \varepsilon$. Furthermore, we may estimate the auxiliary epi-distance $\hat{d}^+_{\rho+\varepsilon}(f_P, f_Q)$ from above by the uniform distance $d_{\mathcal{F}_\mathcal{U}}(P, Q)$ (cf. also Example 7.62 in Rockafellar and Wets (1998)).

It remains to note that the level sets $X_\varepsilon^*(P)$ and $X_\varepsilon^*(Q)$ are also bounded, since f_P and f_Q are lower semicontinuous and convex, and their solution sets are nonempty and bounded, respectively. Hence, we may choose the constant ρ large enough such that the equality $\mathbb{D}_\rho(X_\varepsilon^*(P), X_\varepsilon^*(Q)) = \mathbb{D}_\infty(X_\varepsilon^*(P), X_\varepsilon^*(Q))$ holds. This completes the proof. □

Most of the results in this and the previous section illuminate the role of the distance $d_{\mathcal{F}_\mathcal{U}}$ as a *minimal information* (*m.i.*) pseudometric for stability, i.e., as a pseudometric processing the minimal information of problem (1.1) and implying quantitative stability of its optimal values and solution sets. Furthermore, notice that all results remain valid when enlarging the set $\mathcal{F}_\mathcal{U}$ and, thus, bounding $d_{\mathcal{F}_\mathcal{U}}$ from above by another distance, and when reducing the set $\mathcal{P}_{\mathcal{F}_\mathcal{U}}$ to a subset on which such a distance is defined and finite.

Such a distance d_{id} bounding $d_{\mathcal{F}_\mathcal{U}}$ from above will be called an *ideal probability metric* associated with (1.1) if it has ζ-structure (1.9) generated by some class of functions $\mathcal{F} = \mathcal{F}_{\mathrm{id}}$ from Ξ to $\overline{\mathbb{R}}$ such that $\mathcal{F}_{\mathrm{id}}$ contains the functions $CF_j(x, \cdot)$ for each $x \in X \cap \mathrm{cl}\,\mathcal{U}$, $j = 0, \ldots, d$, and some normalizing constant $C > 0$, and such that any function in $\mathcal{F}_{\mathrm{id}}$ shares typical analytical properties with some function $F_j(x, \cdot)$.

In our applications of the general analysis in Section 3 we clarify such typical analytical properties. Here, we only mention that typical functions $F_j(x, \cdot)$ in stochastic programming are nondifferentiable, but piecewise locally Lipschitz continuous with discontinuities at boundaries of polyhedral sets. More precisely, function classes $\mathcal{F}$ contained in

$$\mathrm{span}\,\{F\chi_B \colon F \in \mathcal{F},\ B \in \mathcal{B}\}, \tag{2.27}$$

where $\mathcal{F} \subseteq \mathcal{F}_p(\Xi)$, $\mathcal{B} \subseteq \mathcal{B}_{\mathrm{ph}_k}(\Xi)$ for some $p \ge 1$ and $k \in \mathbb{N}$, are candidates for an ideal class $\mathcal{F}_{\mathrm{id}}$. The extremal cases, namely, $\mathcal{F}_p(\Xi)$ and $\mathcal{F}_\mathcal{B}$, are discussed in Section 2.1. To get an idea of how to associate an ideal metric with a stochastic program, we consider the p-th order Fortet–Mourier metric ζ_p introduced in Section 2.1. Then the following result is an immediate consequence of the general ones.

Corollary 14. *Let $d = 0$ and assume that*

(i) *$X^*(P)$ is nonempty and $\mathcal{U}$ is an open, bounded neighbourhood of $X^*(P)$,*

(ii) *X is convex and $F_0(\cdot, \xi)$ is convex on $\mathbb{R}^m$ for each $\xi \in \Xi$,*
(iii) *there exist constants $L > 0$, $p \geq 1$ such that $\frac{1}{L} F_0(x, \cdot) \in \mathcal{F}_p(\Xi)$ for each $x \in X \cap \operatorname{cl} U$.*

Then there exists a constant $\delta > 0$ such that

$$|\vartheta(P) - \vartheta(Q)| \leq L\zeta_p(P, Q) \quad \text{and}$$
$$\emptyset \neq X^*(Q) \subseteq X^*/(P) + \Psi_P(L\zeta_p(P, Q))\mathbb{B}$$

whenever $Q \in \mathcal{P}_p(\Xi)$ and $\zeta_p(P, Q) < \delta$. Here, the function Ψ_P is given by (2.23).

Proof. The assumptions of Theorem 5 are satisfied. Hence, the result is a consequence of the Theorems 5 and 9 and the fact that (iii) is equivalent to

$$|F_0(x, \xi) - F_0(x, \tilde{\xi})| \leq L \max\left\{1, \|\xi\|, \|\tilde{\xi}\|\right\}^{p-1} \|\xi - \tilde{\xi}\|$$

for each $\xi, \tilde{\xi} \in \Xi$ and $x \in X \cap \operatorname{cl} \mathcal{U}$, and, thus, it implies $d_{\mathcal{F}_\mathcal{U}}(P, Q) \leq L\zeta_p(P, Q)$ for all $P, Q \in \mathcal{P}_p(\Xi)$. Furthermore, due to the convexity assumption (ii) the localized optimal values $\vartheta_\mathcal{U}$ and solution sets $X^*_\mathcal{U}$ may be replaced by ϑ and X^*, respectively, if Q is close to P (see Remark 11). □

Example 15. (newsboy continued) In case of minimal expected costs the set $\mathcal{F}_\mathcal{U}$ is a specific class of piecewise linear functions of the form $\{(r - c)x + c\max\{0, x - \cdot\} : x \in X \cap \operatorname{cl} \mathcal{U}\}$. Furthermore, $\int_\Xi F_0(x, \xi)\, \mathrm{d}P(\xi)$ is also piecewise linear and Corollary 14 applies with $L := c$, $p := 1$ and a linear function Ψ_P. Hence, the solution set $X^*(\cdot)$ behaves upper Lipschitzian at $P_1(\mathbb{N})$ with respect to ζ_1, i.e.,

$$\sup_{x \in X^*(Q)} d(x, X^*(P)) \leq c\zeta_1(P, Q) = c \int_{\mathbb{R}} \left| F_P(r) - F_Q(r) \right| \mathrm{d}r = c \sum_{k \in \mathbb{N}} \left| \sum_{i=1}^{k} (\pi_i - \tilde{\pi}_i) \right|.$$

Here, we made use of an explicit representation of the Kantorovich metric on $\mathcal{P}(\mathbb{R})$ (Section 5.4 in Rachev (1991)), and F_P and F_Q are the probability distribution functions of the measures $P = \sum_{k \in N} \pi_k \delta_k$ and $Q = \sum_{k \in N} \tilde{\pi}_k \delta_k$, respectively.

2.4 Mean-risk models

The expectation functional appearing in the basic model (1.1) is certainly not the only statistical parameter of interest of the (real-valued) cost or constraint functions F_j, $j = 0, \ldots, d$, with respect to P. *Risk functionals* or *risk measures* are regarded as statistical parameters of probability measures in $\mathcal{P}(\mathbb{R})$, i.e., they are mappings from subsets of $\mathcal{P}(\mathbb{R})$ to $\mathbb{R}$. When risk functionals

are used in the context of the model (1.1), they are evaluated at the probability distributions $P[F_j(x,\cdot)]^{-1}$ for $x \in X$ and $j = 0,\ldots,d$. Practical risk management in decision making under uncertainty often requires to minimize or bound several risk functionals of the underlying distributions. Typical examples for risk functionals are (standard semi-) deviations, excess probabilities, value-at-risk, conditional value-at-risk etc. Some risk measures are defined as infima of certain (simple) stochastic optimization models (e.g., value-at-risk, conditional value-at-risk). Other measures are given as the expectation of a nonlinear function and, hence, their optimization fits into the framework of model (1.1) (e.g., expected utility functions, excess probabilities).

We refer to Section 4 of Pflug (2003) for an introduction to risk functionals and various examples, to Artzner et al. (1999), Delbaen (2002), Föllmer and Schied (2002) for a theory of coherent and convex risk measures, to Ogryczak and Ruszczyński (1999) for the relations to stochastic dominance and to Rockafellar and Uryasev (2002) for the role of the conditional value-at-risk.

Now, we assume that risk functionals $\mathbb{F}_j$, $j = 0,\ldots,d$ are given. In addition to the mean-risk model (1.2) we denote by Q a perturbation of the original probability measure P and consider the perturbed model

$$\min\{\mathbb{F}_0(Q[F_0(x,\cdot)]^{-1}) : x \in X, \mathbb{F}_j(Q[F_j(x,\cdot)]^{-1}) \le 0, j = 1,\ldots,d\}. \tag{2.28}$$

To have all risk functionals $\mathbb{F}_j$ well defined, we assume for simplicity that they are given on the subset $\mathcal{P}_{\mathrm{b}}(\mathbb{R})$ of all probability measures in $\mathcal{P}(\mathbb{R})$ having bounded support. Then both models, (1.2) and (2.28), are well defined if we assume that all functions $F_j(x,\cdot)$ are bounded. Furthermore, we will need a continuity property of risk functionals.

A risk functional $\mathbb{F}$ on $\mathcal{P}_{\mathrm{b}}(\mathbb{R})$ is called *Lipschitz continuous w.r.t. to a class* $\mathcal{H}$ of measurable functions from $\mathbb{R}$ to $\mathbb{R}$ if the estimate

$$|\mathbb{F}(G) - \mathbb{F}(\tilde{G})| \le \sup_{H\in\mathcal{H}} \left| \int_{\mathbb{R}} H(r)d(G-\tilde{G})(r) \right| \tag{2.29}$$

is valid for all $G,\tilde{G} \in \mathcal{P}_{\mathrm{b}}(\mathbb{R})$. The following examples and Proposition 8 in Pflug (2003) show that many risk functionals satisfy such a Lipschitz property.

Example 16. We consider the *conditional value-at-risk* of a probability distribution $G \in \mathcal{P}_{\mathrm{b}}(\mathbb{R})$ at level $p \in (0,1)$, which is defined by

$$\mathrm{CVaR}_p(G) := \inf\left\{ r + \frac{1}{1-p}\int_{\mathbb{R}} \max\{0,\xi - r\}\, \mathrm{d}G(\xi) : r \in \mathbb{R} \right\}.$$

Hence, $\mathrm{CVaR}_p(G)$ is the optimal value of a stochastic program with recourse (see Section 3.1). Clearly, the estimate

$$|\mathrm{CVaR}_p(G) - \mathrm{CVaR}_p(\tilde{G})| \le \frac{1}{1-p} \sup_{r\in\mathbb{R}} \left| \int_{\mathbb{R}} \max\{0,\xi - r\} d(G - \tilde{G})(\xi) \right|$$

is valid for all $G,\tilde{G} \in \mathcal{P}_{\mathrm{b}}(\mathbb{R})$. Hence, the conditional value-at-risk is Lipschitz continuous w.r.t. the class $\mathcal{H} := \{\max\{0, \cdot - r\} : r \in \mathbb{R}\}$.

The *value-at-risk* of $G \in \mathcal{P}_{\mathrm{b}}(\mathbb{R})$ at level $p \in (0,1)$ is given by

$$\mathrm{VaR}_p(G) := \inf\{r \in \mathbb{R} : G(\xi \le r) \ge p\}.$$

Thus, $\mathrm{VaR}_p(G)$ is the optimal value of a chance constrained stochastic program. In Section 3.3 it is shown that the metric regularity of the mapping $r \mapsto \{y \in \mathbb{R} : G(\xi \le r) \ge p - y\}$ at pairs $(\bar{r}, 0)$ with $\bar{r} \in X^*(G)$ is indispenable for Lipschitz continuity properties of the optimal value. If the metric regularity property is satisfied for the measure G and the level p, we obtain, from Theorem 39 the estimate

$$|\mathrm{VaR}_p(G) - \mathrm{VaR}_p(\tilde{G})| \le L d_{\mathrm{K}}(G,\tilde{G}) = \sup_{r\in\mathbb{R}} \left| \int_{\mathbb{R}} L\chi_{(-\infty,r]}(\xi) d(G - \tilde{G})(\xi) \right|$$

for some constant $L > 0$ and sufficiently small Kolmogorov distance $d_{\mathrm{K}}(G,\tilde{G})$. Hence, the corresponding class of functions is $\mathcal{H} := \{L\chi_{(-\infty,r]} : r \in \mathbb{R}\}$. We note that the metric regularity requirement may lead to serious complications when using the value-at-risk in stochastic programming models because $\mathrm{VaR}_p(\cdot)$ has to be evaluated at measures depending on x.

Example 17. The *upper semi-deviation* $sd_+(G)$ of a measure $G \in \mathcal{P}_{\mathrm{b}}(\mathbb{R})$, which is defined by

$$sd_+(G) := \int_{\mathbb{R}} \max\left\{0, \xi - \int_{\mathbb{R}} u \, \mathrm{d}G(u)\right\} \mathrm{d}G(\xi),$$

is Lipschitz continuous w.r.t. the class $\mathcal{H} := \{\max\{0, \cdot - r\} + \cdot : r \in \mathbb{R}\}$.

The examples indicate that typical Lipschitz continuity classes $\mathcal{H}$ of risk functionals contain products of some functions in $\mathcal{F}_k(\mathbb{R})$ for some $k \in \mathbb{N}$ and of characteristic functions $\chi_{(-\infty,r]}$ for some $r \in \mathbb{R}$. Hence, their structure is strongly related to that of the ideal function classes (2.27) for stability.

To state our main stability result for the model (1.2), let $\mathcal{X}(P)$, $\vartheta(P)$, $X^*(P)$ denote the following more general quantities in this section:

$$\mathcal{X}(P) := \{x \in X : \mathbb{F}_j(P[F_j(x,\cdot)]^{-1}) \le 0, j = 1,\dots,d\}$$

$$\vartheta(P) := \inf\{\mathbb{F}_0(P[F_0(x,\cdot)]^{-1}) : x \in \mathcal{X}(P)\},$$

$$X^*(P) := \{x \in \mathcal{X}(P) : \mathbb{F}_0(P[F_0(x,\cdot)]^{-1}) = \vartheta(P)\}.$$

The localized notions $\vartheta_{\mathcal{U}}(P)$ and $X^*_{\mathcal{U}}(P)$ are defined accordingly.

Theorem 18. *For each $j = 0,\dots,d$, let the function F_j be uniformly bounded and the risk functional $\mathbb{F}_j$ be Lipschitz continuous on $\mathcal{P}_{\mathrm{b}}(\mathbb{R})$ w.r.t. some class $\mathcal{H}_j$ of measurable functions from $\mathbb{R}$ to $\mathbb{R}$. Let $P \in \mathcal{P}(\Xi)$ and assume that*

(i) *$X^*(P) \ne \phi$ and $\mathcal{U} \subseteq \mathbb{R}^m$ is an open bounded neighbourhood of $X^*(P)$,*
(ii) *if $d \ge 1$, the function $x \mapsto \mathbb{F}_0(P[F_0(x,\cdot)]^{-1})$ is Lipschitz continuous on $X \cap \mathrm{cl}\,\mathcal{U}$,*
(iii) *the mapping $x \mapsto \{y \in \mathbb{R}^d : x \in X, \mathbb{F}_j(P[F_j(x,\cdot)]^{-1}) \le y_j, j = 1,\dots,d\}$ from $\mathbb{R}^m$ to $\mathbb{R}^d$ is metrically regular at each pair $(\overline{x},0)$ with $\overline{x} \in X^*(P)$.*

Then there exist constants $L > 0$ and $\delta > 0$ such that the estimates

$$|\vartheta(P) - \vartheta_{\mathcal{U}}(Q)| \le L d_{\mathcal{F}^{\mathcal{H}}_{\mathcal{U}}}(P,Q)$$

$$\emptyset \ne X^*_{\mathcal{U}}(Q) \subseteq X^*(P) + L\Psi_P(d_{\mathcal{F}^{\mathcal{H}}_{\mathcal{U}}}(P,Q))\mathbb{B}$$

are valid whenever $Q \in \mathcal{P}(\Xi)$ and $d_{\mathcal{F}^{\mathcal{H}}_{\mathcal{U}}}(P,Q) < \delta$. Here, Ψ_P is given by (2.23) and the distance $d_{\mathcal{F}^{\mathcal{H}}_{\mathcal{U}}}$ is defined by

$$d_{\mathcal{F}^{\mathcal{H}}_{\mathcal{U}}}(P,Q) := \sup_{\substack{j=0,\dots,d \\ x \in X \cap \mathrm{cl}\,\mathcal{U} \\ H_j \in \mathcal{H}_j}} \left| \int_{\Xi} H_j(F_j(x,\xi))(P-Q)(\mathrm{d}\xi) \right|.$$

Proof. We proceed as in the proofs of Theorems 5 and 9, but now we use the distance

$$\hat{d}_{\mathbb{F}}(P,Q) := \sup_{\substack{j=0,\dots,d \\ x \in X \cap \mathrm{cl}\,\mathcal{U}}} |\mathbb{F}_j(P[F_j(x,\cdot)]^{-1}) - \mathbb{F}_j(Q[F_j(x,\cdot)]^{-1})|$$

instead of $d_{\mathcal{F}_\mathcal{U}}$. In this way we obtain constants $L > 0$, $\delta > 0$ and the estimates

$$|\vartheta(P) - \vartheta_\mathcal{U}(Q)| \le L\hat{d}_\mathbb{F}(P, Q)$$
$$\emptyset \ne X^*_\mathcal{U}(Q) \subseteq X^*(P) + L\Psi_P(\hat{d}_\mathbb{F}(P, Q))\mathbb{B}$$

for each $Q \in \mathcal{P}(\Xi)$ such that $\hat{d}_\mathbb{F}(P, Q) < \delta$. It remains to appeal to the estimate

$$\hat{d}_\mathbb{F}(P, Q) \le \sup_{\substack{j=0,\dots,d \\ x\in X\cap \mathrm{cl}\,\mathcal{U}}} \sup_{H_j\in\mathcal{H}_j} \left| \int_\mathbb{R} H_j(r) d((P-Q)[F_j(x,\cdot)]^{-1})(r) \right| = d_{\mathcal{F}^\mathcal{H}_\mathcal{U}}(P, Q),$$

which is a consequence of the Lipschitz continuity (2.29) of the risk functionals $\mathbb{F}_j$, $j = 0, \dots, d$. □

The result implies that stability properties of the mean-risk model (1.2) containing risk functionals $\mathbb{F}_j$ with Lipschitz continuity classes $\mathcal{H}_j$, $j = 0, \dots, d$, depend on the class

$$\mathcal{F}^\mathcal{H}_\mathcal{U} := \{H_j(F_j(x,\cdot)) \colon x \in X \cap \mathrm{cl}\,\mathcal{U},\ H_j \in \mathcal{H}_j,\ j = 0, \dots, d\}$$

instead of $\mathcal{F}_\mathcal{U}$ in case of model (1.1). Hence, the stability behaviour may change considerably when replacing the expectation functionals in (1.1) by other risk functionals. For example, the newsboy model based on minimal expected costs behaves stable at all $P \in \mathcal{P}_1(\mathbb{N})$ (Example 15), but the minimum risk variant of the model (see Example 1) may become unstable.

Example 19. (newsboy continued) We consider the chance constrained model (1.3) whose solution set is $X^*(P) = \{(k,0)\}$ with the maximal k such that $\sum_{i=k}^\infty \pi_i \ge p$ in its first component. We assume that equality $\sum_{i=k}^\infty \pi_i = p$ and $\pi_k > 0$ holds. To establish instability, we consider the approximations $P_n := \sum_{i=1_{(n)}}^\infty \pi_i^{(n)} \delta_i$ of P, where $\pi_i^{(n)} := \pi_i$ for all $i \notin \{k-1, k\}$ and $\pi_{k-1}^{(n)} := \pi_{k-1} + \frac{1}{n}$, $\pi_k^{(n)} := \pi_k - \frac{1}{n}$ for sufficiently large $n \in \mathbb{N}$ such that $\pi_k - \frac{1}{n} > 0$. Then the perturbed solution set is $X^*(P_n) = \{(k-1,0)\}$ for any sufficiently large n. On the other hand, we obtain for the Kolmogorov distance $d_K(P, P_n) = \frac{1}{n}$, i.e., weak convergence of (P_n) to P. Furthermore, the model (1.3) is stable with respect to the metric d_K at each $P = \sum_{i=1}^\infty \pi_i \delta_i \in \mathcal{P}(\mathbb{N})$ such that $\sum_{i=1}^k \pi_i \ne 1 - p$ for each $k \in \mathbb{N}$. The latter fact is a consequence of Theorem 5 as the metric regularity condition is satisfied (see also Remark 2.5 in Römisch and Schultz (1991b)).

However, if the conditional value-at-risk or the upper semi-deviation are incorporated into the objective of (mixed-integer) two-stage stochastic programs, their ideal function classes and, thus, their ideal metrics (see Sections 3.1 and 3.2) do not change. These observations are immediate consequences of the following more general conclusion of the previous theorem.

Corollary 20. *Let $d = 0$. We consider the stochastic programming model*

$$\min\{\mathbb{F}_0(P[F_0(x, \cdot)]^{-1}): x \in X\}, \tag{2.30}$$

where F_0 is uniformly bounded and the risk functional $\mathbb{F}_0$ is Lipschitz continuous on $\mathcal{P}_b(\mathbb{R})$ w.r.t. some class $\mathcal{H}_0$.

Let $P \in \mathcal{P}(\Xi)$, $X^*(P) \neq \emptyset$ and $\mathcal{U}$ be an open bounded neighbourhood of $X^*(P)$.

Assume that $\{F_0(x,\cdot): x \in X \cap \mathrm{cl}\,\mathcal{U}\}$ is contained in some class $\mathcal{F}_c$ of functions from Ξ to $\mathbb{R}$ and $H \circ F \in L_0\mathcal{F}_c$ holds for all $H \in \mathcal{H}_0$, $F \in \mathcal{F}_c$ and some positive constant L_0.

Then there exist constants $L > 0$ and $\delta > 0$ such that the estimates

$$\begin{aligned} |\vartheta(P) - \vartheta_{\mathcal{U}}(Q)| &\leq L d_{\mathcal{F}_c}(P, Q) \\ \emptyset \neq X^*_{\mathcal{U}}(Q) &\subseteq X^*(P) + L\Psi_P(d_{\mathcal{F}_c}(P, Q))\mathbb{B} \end{aligned}$$

are valid whenever $Q \in \mathcal{P}(\Xi)$ and $d_{F_c}(P, Q) < \delta$.

Proof. Clearly, we have in that case $d_{\mathcal{F}^{\mathcal{H}}_{\mathcal{U}}}(P,Q) \leq L_0 d_{\mathcal{F}_c}(P,Q)$. □

Important examples for $\mathcal{H}_0$ and $\mathcal{F}_c$ are multiples of $\mathcal{F}_1(\mathbb{R})$ and of $\mathcal{F}_p(\Xi)$ (for $p \geq 1$) and $\{F\mathcal{X}_B: F \in \mathcal{F}_1(\Xi), B \in \mathcal{B}\}$, respectively.

3 Stability of two-stage and chance constrained programs

3.1 Linear two-stage models

We consider the linear two-stage stochastic program with fixed recourse

$$\min\left\{\langle c, x\rangle + \int_{\Xi} \langle q(\xi), y(\xi)\rangle \, \mathrm{d}P(\xi): Wy(\xi) = h(\xi) - T(\xi)x,\ y(\xi) \geq 0,\ x \in X\right\}, \tag{3.31}$$

where $c \in \mathbb{R}^m$, $X \subseteq \mathbb{R}^m$ and $\Xi \subseteq \mathbb{R}^s$ are convex polyhedral, W is an $(r, \overline{m})$-matrix, $P \in \mathcal{P}(\Xi)$, and the vectors $q(\xi) \in \mathbb{R}^{\overline{m}}$, $h(\xi) \in \mathbb{R}^r$ and the (r, m)-matrix $T(\xi)$ depend affine linearly on $\xi \in \Xi$. The latter assumption covers many practical situations. At the same time, it avoids the inclusion of all components of the recourse costs, the technology matrix and the right-hand side into ξ, because this could lead to serious restrictions when imposing additional conditions on P. We define the function $F_0: \mathbb{R}^m \times \Xi \to \overline{\mathbb{R}}$ by

$$F_0(x, \xi) = \begin{cases} \langle c, x\rangle + \Phi(q(\xi), h(\xi) - T(\xi)x), & h(\xi) - T(\xi)x \in \mathrm{pos}\ W,\ q(\xi) \in D \\ +\infty, & \text{otherwise} \end{cases}$$

where pos $W = \{Wy\colon y \in \mathbb{R}^{\overline{m}}_+\}$, $D = \{u \in \mathbb{R}^{\overline{m}}\colon \{z \in \mathbb{R}^r\colon W'z \le u\} \ne \emptyset\}$ (with W' denoting the transpose of the matrix W) and $\Phi(u,t) = \inf\{\langle u, y\rangle\colon Wy = t,\ y \ge 0\}$ ($(u,t) \in \mathbb{R}^{\overline{m}} \times \mathbb{R}^r$). Then problem (3.31) may be rewritten equivalently as a minimization problem with respect to the first stage decision x, namely,

$$\min\left\{\int_\Xi F_0(x,\xi)\,\mathrm{d}P(\xi)\colon x \in X\right\}. \tag{3.32}$$

In order to utilize the general stability results of Section 2, we need a characterization of the continuity and growth properties of the function F_0. As a first step we recall some well-known properties of the function Φ, which were derived in Walkup and Wets (1969a).

Lemma 21. *The function Φ is finite and continuous on the $(\overline{m}+r)$-dimensional polyhedral cone $D \times \operatorname{pos} W$ and there exist $(r,\overline{m})$-matrices C_j and $(\overline{m}+r)$-dimensional polyhedral cones $\mathcal{K}_j$, $j = 1,\dots,N$, such that*

$$\bigcup_{j=1}^N \mathcal{K}_j = D \times \operatorname{pos} W, \quad \operatorname{int}\mathcal{K}_i \cap \operatorname{int}\mathcal{K}_j = \emptyset,\ i \ne j,$$
$$\Phi(u,t) = \langle C_j u, t\rangle, \quad \text{for each } (u,t) \in \mathcal{K}_j,\ j = 1,\dots,N.$$

Moreover, $\Phi(u,\cdot)$ is convex on pos W *for each $u \in D$, and $\Phi(\cdot,t)$ is concave on D for each $t \in$* pos W.

To have problem (3.32) well defined we introduce the following assumptions:

(A1) For each $(x,\xi) \in X \times \Xi$ it holds that $h(\xi) - T(\xi)x \in \operatorname{pos} W$ and $q(\xi) \in D$.
(A2) $P \in \mathcal{P}_2(\Xi)$, i.e., $\int_\Xi \|\xi\|^2\,\mathrm{d}P(\xi) < \infty$.

Condition (A1) sheds some light on the role of the set Ξ. Due to the affine linearity of $q(\cdot)$, $h(\cdot)$ and $T(\cdot)$ the polyhedrality assumption on Ξ is not restrictive. (A1) combines the two usual conditions: *relatively complete recourse* and *dual feasibility*. It implies that $X \times \Xi \subseteq \operatorname{dom} F_0$.

Proposition 22. *Let* (A1) *be satisfied. Then F_0 is a random convex function. Furthermore, there exist constants $L > 0$, $\hat{L} > 0$ and $K > 0$ such that the following holds for all $\xi, \tilde{\xi} \in \Xi$ and $x, \tilde{x} \in X$ with* $\max\{\|x\|, \|\tilde{x}\|\} \le r$:

$$|F_0(x,\xi) - F_0(x,\tilde{\xi})| \le Lr \max\{1, \|\xi\|, \|\tilde{\xi}\|\}\|\xi - \tilde{\xi}\|,$$
$$|F_0(x,\xi) - F_0(\tilde{x},\xi)| \le \hat{L} \max\{1, \|\xi\|^2\}\|x - \tilde{x}\|,$$
$$|F_0(x,\xi)| \le Kr \max\{1, \|\xi\|^2\}.$$

Proof. From Lemma 21 and (A1) we conclude that F_0 is continuous on $\operatorname{dom} F_0$ and, hence, on $X \times \Xi$. This implies that F_0 is a random lower semi-continuous function (cf. Example 14.31 in Rockafellar and Wets, 1998). It is a random convex function since the properties of Φ in Lemma 21 imply that $F_0(\cdot, \xi)$ is convex for each $\xi \in \Xi$. In order to verify the Lipschitz property of F_0, let $x \in X$ with $\|x\| \le r$ and consider, for each $j = 1, \ldots, N$, and $\xi \in \Xi_j$ the function

$$g_j(\xi) := F_0(x, \xi) = \Phi(q(\xi), h(\xi) - T(\xi)x) = \langle C_j q(\xi), h(\xi) - T(\xi)x \rangle,$$

where the sets $\Xi_j := \{\xi \in \Xi\colon (q(\xi), h(\xi) - T(\xi)x) \in \mathcal{K}_j\}$ are polyhedral, and C_j and $\mathcal{K}_j$ are the matrices and the polyhedral cones from Lemma 21, respectively. Since $q(\cdot)$, $h(\cdot)$ and $T(\cdot)$ depend affine linearly on ξ, the function g_j depends quadratically on ξ and linearly on x. Hence, there exists a constant $L_j > 0$ such that g_j satisfies the following Lipschitz property:

$$|g_j(\xi) - g_j(\tilde{\xi})| \le L_j r \max\{1, \|\xi\|, \|\tilde{\xi}\|\}\|\xi - \tilde{\xi}\| \quad \text{for all } \xi, \tilde{\xi} \in \Xi_j.$$

Now, let $\xi, \tilde{\xi} \in \Xi$, assume that $\xi \in \Xi_i$ and $\tilde{\xi} \in \Xi_k$ for some $i, k \in \{1, \ldots, N\}$ and consider the line segment $[\xi, \tilde{\xi}] = \{\xi(\lambda) = (1 - \lambda)\xi + \lambda\tilde{\xi}\colon \lambda \in [0, 1]\}$. Since $[\xi, \tilde{\xi}] \subseteq \Xi$, there exist indices i_j, $j = 1, \ldots, l$, such that $i_1 = i$, $i_l = k$, $[\xi, \tilde{\xi}] \cap \Xi_{i_j} \ne \emptyset$ for each $j = 1, \ldots, l$ and $[\xi, \tilde{\xi}] \subseteq \cup_{j=1}^{l} \Xi_{i_j}$. Furthermore, there exist increasing numbers $\lambda_{i_j} \in [0, 1]$ for $j = 0, \ldots, l-1$ such that $\xi(\lambda_{i_0}) = \xi(0) = \xi$, $\xi(\lambda_{i_j}) \in \Xi_{i_j} \cap \Xi_{i_{j+1}}$ and $\xi(\lambda) \notin \Xi_{i_j}$ if $\lambda_{i_j} < \lambda \le 1$. Then we obtain

$$\begin{aligned}
|F_0(x, \xi) - F_0(x, \tilde{\xi})| &= |g_{i_1}(\xi) - g_{i_l}(\tilde{\xi})| \\
&\le \sum_{j=0}^{l-1} |g_{i_{j+1}}(\xi(\lambda_{i_j})) - g_{i_{j+1}}(\xi(\lambda_{i_{j+1}}))| \\
&\le \sum_{j=0}^{l-1} L_{i_{j+1}} r \max\{1, \|\xi\|, \|\tilde{\xi}\|\}\|\xi(\lambda_{i_j}) - \xi(\lambda_{i_{j+1}})\| \\
&\le \max_{j=1,\ldots,N} L_j r \max\{1, \|\xi\|, \|\tilde{\xi}\|\} \sum_{j=0}^{l-1} \|\xi(\lambda_{i_j}) - \xi(\lambda_{i_{j+1}})\| \\
&\le \max_{j=1,\ldots,N} L_j r \max\{1, \|\xi\|, \|\tilde{\xi}\|\}\|\xi - \tilde{\xi}\|,
\end{aligned}$$

where we have used for the last three estimates that $\|\xi(\lambda)\| \le \max\{\|\xi\|, \|\tilde{\xi}\|\}$ for each $\lambda \in [0, 1]$ and $|\lambda - \tilde{\lambda}|\|\xi - \tilde{\xi}\| = \|\xi(\lambda) - \xi(\tilde{\lambda})\|$ holds for all λ, $\tilde{\lambda} \in [0, 1]$. Lipschitz continuity of F_0 with respect to x is shown in Theorem 10 of Kall (1976) and in Theorem 7.7 of Wets (1974). In particular, the second estimate

of the proposition is a consequence of those results. Furthermore, from Lemma 21 we conclude the estimate

$$|F_0(x,\xi)| \le \sup_{\|x\|\le r}\left\{|\langle c, x\rangle| + \max_{j=1,\dots,N}|\langle C_j q(\xi), h(\xi) - T(\xi)x\rangle|\right\}$$
$$\le \|c\|r + \left(\max_{j=1,\dots,N}\|C_j\|\right)\|q(\xi)\|(\|h(\xi)\| + \|T(\xi)\|r)$$

for any pair $(x,\xi)\in X\times\Xi$ with $\|x\|\le r$. Then the third estimate follows again from the fact that $q(\cdot)$, $h(\cdot)$ and $T(\cdot)$ depend affine linearly on ξ. □

The estimate in Proposition 22 implies that, for any $r>0$, any nonempty bounded $\mathcal{U}\subseteq\mathbb{R}^m$ and some $\rho>0$, it holds that

$$\int_\Xi \inf_{\substack{x\in X\\ \|x\|\le r}} F_0(x,\xi)\,\mathrm{d}Q(\xi) \ge -Kr\left(1+\int_\Xi \|\xi\|^2\,\mathrm{d}Q(\xi)\right) > -\infty,$$
$$\sup_{x\in X\cap\mathcal{U}}\left|\int_\Xi F_0(x,\xi)\,\mathrm{d}Q(\xi)\right| \le K\rho\left(1+\int_\Xi \|\xi\|^2\,\mathrm{d}Q(\xi)\right) < \infty,$$

if $Q\in\mathcal{P}(\Xi)$ has a finite second order moment. Hence, for any nonempty bounded $\mathcal{U}\subseteq\mathbb{R}^m$ the set of probability measures $\mathcal{P}_{\mathcal{F}_\mathcal{U}}$ contains the set of measures on Ξ having finite second order moments, i.e.,

$$\mathcal{P}_{\mathcal{F}_\mathcal{U}} \supseteq \left\{Q\in\mathcal{P}(\Xi)\colon \int_\Xi \|\xi\|^2\,\mathrm{d}Q(\xi) < \infty\right\} = \mathcal{P}_2(\Xi).$$

The following stability results for optimal values and solution sets of the two-stage problem (3.32) are now a direct consequence of the results of Section 2.

Theorem 23. *Let* (A1) *and* (A2) *be satisfied and let* $X^*(P)$ *be nonempty and* $\mathcal{U}$ *be an open, bounded neighbourhood of* $X^*(P)$.

Then there exist constants $L>0$ *and* $\delta>0$ *such that*

$$|\vartheta(P) - \vartheta(Q)| \le L\zeta_2(P,Q)$$

$$\emptyset \ne X^*(Q) \subseteq X^*(P) + \Psi_P(L\zeta_2(P,Q))\mathbb{B}$$

whenever $Q\in\mathcal{P}_2(\Xi)$ *and* $\zeta_2(P,Q)<\delta$, *where* Ψ_P *is given by* (2.23).

Proof. The result is a consequence of Corollary 14 with $p=2$. The assumptions (ii) and (iii) of Corollary 14 are verified in Proposition 22. □

Theorem 24. *Let* (A1) *and* (A2) *be satisfied and let* $X^*(P)$ *be nonempty and bounded. Then there exist constants* $\overline{L} > 0$ *and* $\overline{\varepsilon} > 0$ *such that the estimate*

$$\mathbb{D}_\infty(X_\varepsilon^*(P), X_\varepsilon^*(Q)) \leq \frac{\overline{L}}{\varepsilon}\, \zeta_2(P, Q)$$

holds for any $\varepsilon \in (0, \overline{\varepsilon})$ *and* $Q \in \mathcal{P}_2(\Xi)$ *such that* $\zeta_2(P, Q) < \delta$. *Here,* $\mathbb{D}_\infty$ *denotes the Pompeiu–Hausdorff distance* (2.26).

Proof. Since the assumptions of Theorem 13 are satisfied, we conclude that there exist constants $\rho > 0$ and $\overline{\varepsilon} > 0$ such that

$$\mathbb{D}_\infty(X_\varepsilon^*(P), X_\varepsilon^*(Q)) \leq \frac{2\rho}{\varepsilon}\, d_{\mathcal{F}_\mathcal{U}}(P, Q)$$

holds for $\mathcal{U} := (\rho + \overline{\varepsilon})\mathbb{B}$ and any $\varepsilon \in (0, \overline{\varepsilon})$, $Q \in \mathcal{P}_{\mathcal{F}_\mathcal{U}}$ such that $d_{\mathcal{F}_\mathcal{U}}(P, Q) < \varepsilon$. Proposition 22 implies the estimate $d_{\mathcal{F}_\mathcal{U}}(P, Q) \leq L(\rho + \overline{\varepsilon})\zeta_2(P, Q)$, for some constant $L > 0$, which completes the proof. □

The theorems establish the quantitative stability of $\vartheta(\cdot)$ and $X^*(\cdot)$ and the Lipschitz stability of $X_\varepsilon^*(\cdot)$ with respect to ζ_2 in case of two-stage models with fixed recourse for fairly general situations. In case that either only the recourse costs or only the technology matrix and right-hand side are random, both results are valid for $(\mathcal{P}_1(\Xi), \zeta_1)$ instead of $(\mathcal{P}_2(\Xi), \zeta_2)$. We verify this observation for the corresponding conclusion of Theorem 23.

Corollary 25. *Let either only* $q(\cdot)$ *or only* $T(\cdot)$ *and* $h(\cdot)$ *be random and* (A1) *be satisfied. Let* $P \in \mathcal{P}_1(\Xi)$, $X^*(P)$ *be nonempty and* $\mathcal{U}$ *be an open, bounded neighbourhood of* $X^*(P)$. *Then there exist constants* $L > 0$, $\delta > 0$ *such that*

$$|\vartheta(P) - \vartheta(Q)| \leq L\zeta_1(P, Q)$$

$$\emptyset \neq X^*(Q) \subseteq X^*(P) + \Psi_P(L\zeta_1(P, Q))\mathbb{B}$$

whenever $Q \in \mathcal{P}_1(\Xi)$ *and* $\zeta_1(P, Q) < \delta$, *where* Ψ_P *is given by* (2.23).

Proof. By inspecting the proof of Proposition 22 one observes that now the function F_0 satisfies the following continuity and growth properties for all $\xi, \tilde{\xi} \in \Xi$ and $x, \tilde{x} \in X$ with $\max\{\|x\|, \|\tilde{x}\|\} \le r$:

$$|F_0(x, \xi) - F_0(x, \tilde{\xi})| \le Lr\|\xi - \tilde{\xi}\|,$$
$$|F_0(x, \xi)| \le Kr \max\{1, \|\xi\|\}.$$

Hence, the set $\mathcal{P}_{\mathcal{F}_U}$ contains $\mathcal{P}_1(\Xi)$ and Corollary 14 applies with $p = 1$. □

Next we provide some examples of recourse models showing that, in general, the estimate for solution sets in Theorem 23 is the best possible one and that $X^*(\cdot)$ is not lower semicontinuous at P if $X^*(P)$ is not a singleton.

All examples exploit the specific structure provided by the *simple recourse* condition, i.e., $\overline{m} = 2s$, $q = (q_+, q_-)$ and $W = (I, -I)$, where $q_+, q_- \in \mathbb{R}^s$ and I is the (s, s)-identity matrix. Then pos $W = \mathbb{R}^s$ holds and, hence, (A1) is satisfied iff $q \in D$, which is equivalent to the condition $q_+ + q_- \ge 0$, and

$$\Phi(q, t) = \sup\{\langle t, u\rangle : -q_- \le u \le q_+\}.$$

Example 26. Let $m = s = r = 1$, $\overline{m} = 2$, $c = 0$, $W = (1, -1)$, $X = [-1, 1]$, $\Xi = \mathbb{R}$, $q(\xi) = (1, 1)$, $T(\xi) = 1$, $h(\xi) = \xi$, $\forall \xi \in \Xi$. Let $P \in \mathcal{P}(\mathbb{R})$ be the uniform distribution on the interval $[-\frac{1}{2}, \frac{1}{2}]$. Then $\vartheta(P) = 1$, $X^*(P) = \{0\}$, and quadratic growth

$$\int_\Xi F_0(x, \xi)\, \mathrm{d}P(\xi) = \int_{-1/2}^{1/2} |\xi - x|\, \mathrm{d}\xi = \frac{1}{4} + x^2 = \vartheta(P) + d(x, X^*(P))^2$$

holds for each $x \in [-\frac{1}{2}, \frac{1}{2}]$. Let us consider the following perturbations $P_n \in \mathcal{P}(\mathbb{R})$ of P for $n > 4$ given by

$$P_n = \left(\frac{1}{2} - \varepsilon_n\right)(P_{ln} + P_{rn}) + \varepsilon_n(\delta_{-\varepsilon_n} + \delta_{\varepsilon_n}),$$

where $\varepsilon = n^{-1/2}$, P_{ln} and P_{rn} are the uniform distributions on $[-\frac{1}{2}, -\varepsilon_n)$ and $(\varepsilon_n, \frac{1}{2}]$, respectively, and δ_r is the measure placing unit mass at r. Using the explicit representation of ζ_1 in case of probability distributions on $\mathbb{R}$ (see Chapter 5.4 of Rachev, 1991), we obtain

$$\zeta_1(P, P_n) = \int_{-\infty}^{\infty} |P((-\infty, \xi]) - P_n((-\infty, \xi])|\, \mathrm{d}\xi = \frac{1}{n} = \varepsilon_n^2.$$

Furthermore, it holds that $\vartheta(P_n) = \frac{1}{2}(\varepsilon_n^2 + \frac{1}{4})$, $X^*(P_n) = [-\varepsilon_n, \varepsilon_n]$ and, hence, $|\vartheta(P) - \vartheta(P_n)| = \frac{1}{2}\varepsilon_n^2$ and $\sup_{x \in X^*(P_n)} d(x, X^*(P)) = \varepsilon_n$ for each $n \in \mathbb{N}$. Hence, the estimate in Theorem 23 is best possible.

Next we consider the distribution $\hat{P} = \frac{1}{2}(\delta_{-1/2} + \delta_{1/2})$. Then we have $\vartheta(\hat{P}) = \frac{1}{2}$ and $X^*(\hat{P}) = [-\frac{1}{2}, \frac{1}{2}]$ and the linear growth condition

$$\int_\Xi F_0(x, \xi)\, \mathrm{d}\hat{P}(\xi) = \int_\Xi |\xi - x|\, \mathrm{d}\hat{P}(\xi) = \frac{1}{2}\left(\left|x + \frac{1}{2}\right| + \left|x - \frac{1}{2}\right|\right)$$
$$\geq \vartheta(\hat{P}) + d(x, X^*(\hat{P}))$$

for each $x \in X$. Consider the perturbations $\hat{P}_n = (1 - \frac{1}{n})\hat{P} + \frac{1}{n}\delta_0$ $(n \in \mathbb{N})$ of $\hat{P}$. Then

$$\zeta_1(\hat{P}, \hat{P}_n) = \int_{-\infty}^{\infty} |\hat{P}((-\infty, \xi]) - \hat{P}_n((-\infty, \xi])|\, \mathrm{d}\xi = \frac{1}{2n},$$

holds for each $n \in \mathbb{N}$, where we have again used the explicit representation of ζ_1 in case of probability measures on $\mathbb{R}$. Furthermore, it holds that $\vartheta(\hat{P}_n) = (1 - \frac{1}{n})\frac{1}{2}$ and $X^*(\hat{P}_n) = \{0\}$ for each $n \in \mathbb{N}$. Hence, we have $\sup_{x \in X^*(\hat{P})} d(x, X^*(\hat{P}_n)) = \frac{1}{2}$.

Next we consider models with a stochastic technology matrix and recourse costs, respectively, and show that in such cases $X^*(\cdot)$ is also not lower semicontinuous at P, in general.

Example 27. Let $m = s = r = 1$, $\overline{m} = 2$, $c = 0$, $W = (1, -1)$, $X = [0, 1]$, $\Xi = \mathbb{R}_+$, $h(\xi) = 0$, $\forall \xi \in \Xi$.

In the first case, we set $q(\xi) = (1, 1)$ and $T(\xi) = -\xi$, $\forall \xi \in \Xi$.

In the second case, we set $q(\xi) = (\xi, \xi)$ and $T(\xi) = -1$, $\forall \xi \in \Xi$.

In both cases (A1) is satisfied. We consider $P = \delta_0$ and $P_n = \delta_{1/n}$, i.e., the unit masses at 0 and $\frac{1}{n}$, respectively, for each $n \in \mathbb{N}$. Clearly, (P_n) converges with respect to the metric ζ_1 to P in $\mathcal{P}_1(\mathbb{R})$. Furthermore, in both cases

$$\int_\Xi F_0(x, \xi)\, \mathrm{d}P_n(\xi) = \int_\Xi \xi x\, \mathrm{d}P_n(\xi) = \frac{x}{n}$$

holds for each $x \in X$. Then $X^*(P) = X$ and $X^*(P_n) = \{0\}$ for any $n \in \mathbb{N}$, which implies $\sup_{x \in X^*(P)} d(x, X^*(P_n)) = 1$.

The examples show that continuity properties of $X^*(\cdot)$ at P in terms of the Pompeiu–Hausdorff distance cannot be achieved in general unless $X^*(P)$ is a singleton. Nevertheless, we finally establish such quantitative stability results for models where the technology matrix is fixed, i.e., $T(\xi) \equiv T$, and a specific

nonuniqueness of $X^*(P)$ is admitted. For their derivation we need an argument that decomposes the original two-stage stochastic program into another two-stage program with decisions taken from $T(X)$ and a parametric linear program not depending on P.

Lemma 28. *Let* (A1) *be satisfied and let* $Q \in \mathcal{P}_2(\Xi)$ *be such that* $X^*(Q)$ *is nonempty. Then we have*

$$\begin{aligned}\vartheta(Q) &= \inf\left\{\pi(\chi) + \int_\Xi \Phi(q(\xi), h(\xi) - \chi)\, \mathrm{d}Q(\xi) \colon \chi \in T(X)\right\} \\ &= \pi(Tx) + \int_\Xi \Phi(q(\xi), h(\xi) - Tx)\, \mathrm{d}Q(\xi), \quad \forall x \in X^*(Q), \\ X^*(Q) &= \sigma(Y^*(Q)),\end{aligned}$$

where

$$\begin{aligned}Y^*(Q) &:= \arg\min\left\{\pi(\chi) + \int_\Xi \Phi(q(\xi), h(\xi) - \chi)\, \mathrm{d}Q(\xi) \colon \chi \in T(X)\right\}, \\ \pi(\chi) &:= \inf\{\langle c, x\rangle \colon x \in X,\ Tx = \chi\}, \\ \sigma(\chi) &:= \arg\min\{\langle c, x\rangle \colon x \in X,\ Tx = \chi\} \quad (\chi \in T(X)).\end{aligned}$$

Moreover, π *is convex polyhedral on* $T(X)$ *and* σ *is a polyhedral set-valued mapping which is Lipschitz continuous on* $T(X)$ *with respect to the Pompeiu–Hausdorff distance*.

Proof. Let $\overline{x} \in X^*(Q)$. We set $\Phi_Q(\chi) := \int_\Xi \Phi(q(\xi), h(\xi) - \chi)\, \mathrm{d}Q(\xi)$ and have

$$\vartheta(Q) = \langle c, \overline{x}\rangle + \Phi_Q(T\overline{x}) \geq \inf\{\pi(\chi) + \Phi_Q(\chi) \colon \chi \in T(X)\}.$$

For the converse inequality, let $\varepsilon > 0$ and $\overline{\chi} \in T(X)$ be such that

$$\pi(\overline{\chi}) + \Phi_Q(\overline{\chi}) \leq \inf\{\pi(\chi) + \Phi_Q(\chi) \colon \chi \in T(X)\} + \frac{\varepsilon}{2}.$$

Then there exists an $\overline{x} \in X$ such that $T\overline{x} = \overline{\chi}$ and $\langle c, \overline{x}\rangle \leq \pi(\overline{\chi}) + \frac{\varepsilon}{2}$. Hence,

$$\begin{aligned}\vartheta(Q) &\leq \langle c, \overline{x}\rangle + \Phi_Q(T\overline{x}) \leq \pi(\overline{\chi}) + \Phi_Q(\overline{\chi}) + \frac{\varepsilon}{2} \\ &\leq \inf\{\pi(\chi) + \Phi_Q(\chi) \colon \chi \in T(X)\} + \varepsilon.\end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, the first statement is verified. In particular, $x \in \sigma(Tx)$ and $Tx \in Y^*(Q)$ for any $x \in X^*(Q)$. Hence, it holds that $X^*(Q) \subseteq \sigma(Y^*(Q))$. Conversely, let $x \in \sigma(Y^*(Q))$. Then $x \in \sigma(\chi)$ for some $\chi \in Y^*(Q)$. Thus $Tx = \chi$ and $\langle c, x\rangle = \pi(\chi) = \pi(Tx)$, implying

$$\begin{aligned}\langle c, x\rangle + \Phi_Q(Tx) &= \pi(Tx) + \Phi_Q(Tx) = \inf\{\pi(\chi) + \Phi_Q(\chi)\colon \chi \in T(X)\}\\ &= \vartheta(Q) \quad \text{and} \quad x \in X^*(Q).\end{aligned}$$

Furthermore, π is clearly convex and polyhedral, and the properties of σ are well known (cf. Walkup and Wets, 1969b).

Theorem 29. *Let* (A1), (A2) *be satisfied*, $X^*(P)$ *be nonempty and* $\mathcal{U}$ *be an open bounded neighbourhood of* $X^*(P)$. *Furthermore, assume that* $T(X^*(P))$ *is a singleton. Then there exist constants* $L > 0$ *and* $\delta > 0$ *such that*

$$\mathbb{D}_\infty(X^*(P), X^*(Q)) \le L\Psi_P(L\zeta_2(P, Q))$$

whenever $Q \in \mathcal{P}_2(\Xi)$ *and* $\zeta_2(P, Q) < \delta$, *where* Ψ_P *is given by* (2.23) *and* $\mathbb{D}_\infty$ *denotes the Pompeiu–Hausdorff distance.*

Proof. Let χ^* be the single element belonging to $T(X^*(P))$. We use the notation of Lemma 28 and conclude that $Y^*(P) = \{\chi^*\}$. Let $\mathcal{V}$ denote a neighbourhood of χ^* such that $T^{-1}(\mathcal{V}) \subset \mathcal{U}$ and consider the growth function

$$\psi_P^*(\tau) := \min\{\pi(\chi) + \Phi_P(\chi) - \vartheta(P)\colon \|\chi - \chi^*\| \ge \tau,\ \chi \in T(X) \cap \mathcal{V}\}$$

and the associated function $\Psi_P^*(\eta) := \eta + (\psi_P^*)^{-1}(2\eta)$ of the stochastic program $\inf\{\pi(\chi) + \Phi_P(\chi)\colon \chi \in T(X)\}$. Applying Corollary 14 to the latter program yields the estimate

$$\sup_{\chi \in Y^*(Q)} d(\chi, Y^*(P)) = \sup_{\chi \in Y^*(Q)} \|\chi - \chi^*\| \le \Psi_P^*(L_*\zeta_2(P, Q))$$

for some $L_* > 0$ and small $\zeta_2(P, Q)$. Since $X^*(P) = \sigma(\chi^*)$ and $X^*(Q) = \sigma(Y^*(Q))$ hold due to Lemma 28 and the set-valued mapping σ is Lipschitz continuous on $T(X)$ with respect to $\mathbb{D}_\infty$ (with some constant $L_\sigma > 0$), we obtain

$$\begin{aligned}\mathbb{D}_\infty(X^*(P), X^*(Q)) = \mathbb{D}_\infty(\sigma(\chi^*), \sigma(Y^*(Q)) &\le \sup_{\chi \in Y^*(Q)} \mathbb{D}_\infty(\sigma(\chi^*), \sigma(\chi))\\ &\le L_\sigma \sup_{\chi \in Y^*(Q)} \|\chi^* - \chi\| \le L_\sigma \Psi_P^*(L_*\zeta_2(P, Q)).\end{aligned}$$

It remains to explore the relation between the two growth functions ψ_P and ψ_P^*, and the associated functions Ψ_P and Ψ_P^*, respectively. Let $\tau \in \mathbb{R}_+$ and $\chi_\tau \in T(X) \cap \mathcal{V}$ such that $\|\chi_\tau - \chi^*\| \geq \tau$ and $\psi_P^*(\tau) = \pi(\chi_\tau) + \Phi_P(\chi_\tau) - \vartheta(P)$. Let $x_\tau \in X$, $\tilde{x}_\tau \in X^*(P)$ be such that $Tx_\tau = \chi_\tau$, $\pi(\chi_\tau) = cx_\tau$ and $d(x_\tau, X^*) = \|x_\tau - \tilde{x}_\tau\|$. Hence, we obtain $x_\tau \in \mathcal{U}$, $\psi_P^*(\tau) = cx_\tau + \Phi_P(Tx_\tau) - \vartheta(P)$ and

$$\tau \leq \|\chi_\tau - \chi^*\| = \|Tx_\tau - T\tilde{x}_\tau\| \leq \|T\| d(x_\tau, X^*),$$

where $\|T\|$ denotes the matrix norm of T. If $\|T\| \neq 0$, we conclude that $\psi_P^*(\tau) \geq \psi_P(\frac{\tau}{\|T\|})$ holds for any $\tau \in \mathbb{R}_+$ and, hence, we have $(\psi_P^*)^{-1}(\eta) \leq \|T\| \psi_P^{-1}(\eta)$ and $\Psi_P^*(\eta) \leq \max\{1, \|T\|\} \Psi_P(\eta)$ for any $\eta \in \mathbb{R}_+$. This implies

$$\mathbb{D}_\infty(X^*(P), X^*(Q)) \leq \max\{1, \|T\|\} L_\sigma \Psi_P(L_* \zeta_2(P, Q)),$$

and, thus, the desired estimate. In case of $\|T\| = 0$, the solution set $X^*(P)$ is equal to $\arg\min\{\langle c, x\rangle : x \in X\}$ and, consequently, does not change if P is perturbed. Hence, the result is correct in the latter case, too. □

Theorem 30. *Let* (A1), (A2) *be satisfied*, $X^*(P)$ *be nonempty*, $\mathcal{U}$ *be an open bounded neighbourhood of* $X^*(P)$ *and* $T(X^*(P))$ *be a singleton. Assume that the function* $(\psi_P^r)^{-1}$ *is continuous at* $\tau = 0$, *where* $\psi_P^r(0) = 0$, $\psi_P^r(\tau) := \frac{1}{\tau} \psi_P(\tau)$ *for each* $\tau > 0$ *and* $\psi_P(\cdot)$ *is the growth function given by* (2.22).

Then there exists constants $L > 0$ *and* $\delta > 0$ *such that the estimate*

$$\mathbb{D}_\infty(X^*(P), X^*(Q)) \leq L(\psi_P^r)^{-1}(\hat{d}_{\Phi_\mathcal{U}}(P, Q)) \tag{3.33}$$

is valid for each $Q \in \mathcal{P}_2(\Xi)$ *with* $\hat{d}_{\Phi_\mathcal{U}}(P, Q) < \delta$. *Here, we denote*

$$\hat{d}_{\Phi_\mathcal{U}}(P, Q) := \sup\left\{ \left| \int_\Xi \frac{\Phi(q(\xi), h(\xi) - Tx) - \Phi(q(\xi), h(\xi) - T\overline{x})}{\|x - \overline{x}\|} \mathrm{d}(P - Q)(\xi) \right| : x, \overline{x} \in X \cap \mathrm{cl}\, \mathcal{U}, x \neq \overline{x} \right\}.$$

If the two-stage model (3.31) *has quadratic growth, the estimate* (3.33) *asserts Lipschitz continuity with respect to* $\hat{d}_{\Phi_\mathcal{U}}$.

Proof. Using the same notation as in the previous proof we conclude again that

$$\mathbb{D}_\infty(X^*(P), X^*(Q)) \leq L_\sigma \sup_{\chi \in Y^*(Q)} \|\chi^* - \chi\|.$$

If T is the null matrix, the result is true since $X^*(Q)$ does not depend on Q. Otherwise, we denote by $\|T\|$ the matrix norm of T, argue as in the proofs of the Theorem 12 and 29 and arrive at the estimate

$$\psi_P\left(\frac{1}{\|T\|}\,\|\chi-\chi^*\|\right)\le\psi_P^*(\|\chi-\chi^*\|)\le\Phi_P(\chi)-\Phi_Q(\chi)-(\Phi_P(\chi^*)-\Phi_Q(\chi^*))$$

for each $\chi\in Y^*(Q)$, where $\Phi_P(\chi):=\int_\Xi \Phi(q(\xi),h(\xi)-\chi)\,\mathrm{d}P(\xi)$. The latter estimate implies (3.33). □

Remark 31. In all cases, where the original and perturbed solution sets $X^*(P)$ and $X^*(Q)$ are convex and an estimate of the form

$$\mathbb{D}_\infty(X^*(P),\,X^*(Q))\le\phi(d(P,\,Q))\quad\text{whenever } Q\in\mathcal{P}_d,\ d(P,\,Q)<\delta$$

is available for some (pseudo) metric d on a set of probability measures $\mathcal{P}_d$ and some function ϕ from $\mathbb{R}_+$ to $\mathbb{R}_+$, this estimate may be complemented by a quantitative continuity property of a countable dense family of selections. Namely, there exists a family $\{x_k^*(Q)\}_{k\in\mathbb{N}}$ of selections of $X^*(Q)$ such that

$$X^*(Q)=\mathrm{cl}\left(\bigcup_{k\in\mathbb{N}}x_k^*(Q)\right)$$
$$\|x_k^*(P)-x_k^*(Q)\|\le L_k\phi(d(P,\,Q))\quad\text{whenever } Q\in\mathcal{P}_d,\ d(P,\,Q)<\delta$$

for some constant $L_k>0$ and any $k\in\mathbb{N}$. To derive this conclusion, let us first recall the notion of a generalized Steiner point of a convex compact set $C\subset\mathbb{R}^m$ (see Dentcheva, 2000). It is given by $\mathrm{St}_\alpha(C):=\int_{\mathbb{B}}\mu(\partial\sigma_C(x))\alpha(\mathrm{d}x)$, where $\sigma_C(\cdot)$ is the support function of C, i.e., $\sigma_C(x):=\sup_{y\in C}\langle x,y\rangle$, $\partial\sigma_C(x)$ is the convex subdifferential of σ_C at x and $\mu(\partial\sigma_C(x))$ its norm-minimal element. Furthermore, α is a probability measure on $\mathbb{B}$ having a C^1-density with respect to the Lebesgue measure. A generalized Steiner selection $\mathrm{St}_\alpha(\cdot)$ is Lipschitz continuous (with a Lipschitz constant depending on α) on the set of all nonempty convex compact subsets of $\mathbb{R}^m$ equipped with the distance $\mathbb{D}_\infty$. Furthermore, there exists a countable family $\{\alpha_k\}_{k\in\mathbb{N}}$ of probability measures on $\mathbb{R}$, each having a C^1-density with respect to the Lebesgue measure, such that the corresponding family of generalized Steiner selections $\{\mathrm{St}_{\alpha_k}(C)\}_{k\in\mathbb{N}}$ is dense in C. Both results are proved in Dentcheva (2000). By combining these two arguments for the countable family $\{x_k^*(Q):=\mathrm{St}_{\alpha_k}(X^*(Q))\}_{k\in\mathbb{N}}$ of selections to the convex compact sets $X^*(Q)$ the desired result follows.

The previous Theorems 29 and 30 extend the main results of Römisch and Schultz (1993, 1996) and Shapiro (1994) to the case of a general growth condition. The crucial assumption of both results is that $T(X^*(P))$ is a singleton.

The latter condition is satisfied, for example, if the expected recourse function $\Phi_P(\cdot) := \int_\Xi \Phi(q(\xi), h(\xi) - \cdot)\, dP(\xi)$ is strictly convex on a convex neighbourhood of $T(X^*(P))$.

The situation simplifies in case of random right-hand sides only, i.e., $q(\xi) \equiv q$ and $h(\xi) = \xi$. Then the distance $\hat{d}_{\Phi_\mathcal{U}}$ can be bounded above by a discrepancy w.r.t. certain polyhedral cones. Namely,

$$\hat{d}_{\Phi_\mathcal{U}}(P, Q) \le \hat{L} \sup\{|(P - Q)(Tx + B_i(\mathbb{R}^s_+))| : x \in \operatorname{cl} \mathcal{U},\ i = 1, \dots, \ell\},$$

holds, where $\hat{L} > 0$ is some constant and B_i, $i = 1, \dots, \ell$, are certain nonsingular submatrices of the recourse matrix W (Römisch and Schultz, 1996). In this case, verifiable sufficient conditions for the strict and strong convexity of the expected recourse function Φ_P are also available (Schultz, 1994). Namely, the function Φ_P is strictly convex on any open convex subset of the support of P if P has a density on $\mathbb{R}^s$ and the set $\{z \in \mathbb{R}^s : W'z < q\}$ is nonempty. It is strongly convex if, in addition to the conditions implying strict convexity, the density of P is bounded away from zero on the corresponding convex neighbourhood. Furthermore, the model (3.31) has quadratic growth if the function Φ_P is strongly convex on some open convex neighbourhood of $T(X^*(P))$. The latter fact was proved in Dentcheva and Römisch (2000) by exploiting the Lipschitz continuity of the mapping σ in Lemma 28. The Lipschitz continuity result of Theorem 30 in case of quadratic growth forms the basis of the following differential stability result for optimal values and solution sets proved in Dentcheva and Römisch (2000).

Theorem 32. *Let* (A1), (A2) *be satisfied*, $X^*(P)$ *be nonempty and bounded, and* $T(X^*(P))$ *be a singleton, i.e.*, $T(X^*(P)) = \{\chi^*\}$. *Let* $Q \in \mathcal{P}(\Xi)$.

Then the function ϑ *is Gateaux directionally differentiable at* P *in direction* $Q - P$ *and it holds*

$$\vartheta'(P;\ Q - P) := \lim_{t \to 0+} \frac{1}{t}(\vartheta(P + t(Q - P)) - \vartheta(P)) = \Phi_Q(\chi^*) - \Phi_P(\chi^*).$$

If, in addition, model (3.31) *has quadratic growth and* Φ_P *is twice continuously differentiable at* $\{\chi^*\}$, *then the second-order Gateaux directional derivative of* ϑ *at* P *in direction* $Q - P$ *exists and we have*

$$\begin{aligned}\vartheta''(P;\ Q - P) &:= \lim_{t \to 0+} \frac{1}{t^2}(\vartheta(P + t(Q - P)) - \vartheta(P) - t\vartheta'(P;\ Q - P)) \\ &= \inf\left\{\frac{1}{2}\langle \nabla^2 \Phi_P(\chi^*)Tx,\ Tx\rangle + (\Phi_Q - \Phi_P)'(\chi^*;\ Tx) : x \in S(\overline{x})\right\},\end{aligned}$$

where $S(\overline{x}) = \{x \in T_X(\overline{x})\colon cx + \langle \nabla \Phi_P(\chi^*), Tx\rangle = 0\}$ *and* $T_X(\overline{x})$ *is the tangent cone to* X *at some* $\overline{x} \in X^*(P)$. *The directional derivative* $(\Phi_Q - \Phi_P)'(\chi^*;\ Tx)$ *of* $\Phi_Q - \Phi_P$ *exists since both functions are convex and* Φ_P *is differentiable.*

The first-order Gateaux directional derivative of the set-valued mapping $X^*(\,\cdot\,)$

$$(X^*)'(P, \overline{x};\ Q - P) := \lim_{t \to 0+} \frac{1}{t}(X^*(P + t(Q - P)) - \overline{x})$$

at the pair $(P, \overline{x})$, $\overline{x} \in X^*(P)$, *in direction* $Q - P$ *exists and coincides with* $\arg\min\{\frac{1}{2}\langle \nabla^2 \Phi_P(\chi^*)Tx, Tx\rangle + (\Phi_Q - \Phi_P)'(\chi^*;\ Tx)\colon x \in S(\overline{x})\}$.

3.2 *Mixed-integer two-stage models*

Next we allow for mixed-integer decisions in both stages and consider the stochastic program

$$\min\left\{\langle c, x\rangle + \int_{\Xi} \Phi(h(\xi) - T(\xi)x)\, \mathrm{d}P(\xi)\colon x \in X\right\}, \tag{3.34}$$

where

$$\Phi(t) := \min\{\langle q, y\rangle + \langle \overline{q}, \overline{y}\rangle\colon Wy + \overline{W}\overline{y} = t,\ y \in \mathbb{Z}_+^{\hat{m}},\ \overline{y} \in \mathbb{R}_+^{\overline{m}}\} \quad (t \in \mathbb{R}^r), \tag{3.35}$$

$c \in \mathbb{R}^m$, X is a closed subset of $\mathbb{R}^m$, Ξ a polyhedron in $\mathbb{R}^s$, $q \in \mathbb{R}^{\hat{m}}$, $\overline{q} \in \mathbb{R}^{\overline{m}}$, W and $\overline{W}$ are $(r, \hat{m})$- and $(r, \overline{m})$-matrices, respectively, $h(\xi) \in \mathbb{R}^r$ and the (r, m)-matrix $T(\xi)$ are affine linear functions of $\xi \in \mathbb{R}^s$, and $P \in \mathcal{P}(\Xi)$.

Basic properties of Φ like convexity and continuity on $\operatorname{dom}\Phi$ in the purely linear case cannot be maintained for reasonable problem classes. Since Φ is discontinuous, in general, it is interesting to characterize its continuity regions. Similarly, as for the two-stage models without integrality requirements in the previous section, we need some conditions to have the model (3.34) well-defined:

(B1) The matrices W and $\overline{W}$ have only rational elements.
(B2) For each pair $(x, \xi) \in X \times \Xi$ it holds that $h(\xi) - T(\xi)x \in \mathcal{T}$, where

$$\mathcal{T} := \{t \in \mathbb{R}^r\colon t = Wy + \overline{W}\overline{y},\ y \in \mathbb{Z}_+^{\hat{m}},\ \overline{y} \in \mathbb{R}_+^{\overline{m}}\}.$$

(B3) There exists an element $u \in \mathbb{R}^r$ such that $W'u \le q$ and $\overline{W}'u \le \overline{q}$.
(B4) $P \in \mathcal{P}_1(\Xi)$, i.e., $\int_{\Xi} \|\xi\|\, \mathrm{d}P(\xi) < +\infty$.

The conditions (B2) and (B3) mean *relatively complete recourse* and *dual feasibility*, respectively. We note that condition (B3) is equivalent to $\Phi(0) = 0$, and that (B2) and (B3) imply $\Phi(t)$ to be finite for all $t \in \mathcal{T}$ (see Proposition 1 in Louveaux and Schultz, 2003). In the context of this section, the following properties of the value function Φ on $\mathcal{T}$ are important.

Lemma 33. *Assume* (B1)–(B3). *Then there exists a countable partition of* $\mathcal{T}$ *into Borel subsets* $\mathcal{B}_i$, *i.e.*, $\mathcal{T} = \cup_{i\in\mathbb{N}} \mathcal{B}_i$ *such that*

(1) *each of the sets has a representation* $\mathcal{B}_i = \{b_i + \operatorname{pos} \overline{W}\} \setminus \cup_{j=1}^{N_0} \{b_{ij} + \operatorname{pos} \overline{W}\}$, *where* b_i, $b_{ij} \in \mathbb{R}^r$ *for* $i \in \mathbb{N}$ *and* $j = 1, \ldots, N_0$. *Moreover, there exists an* $N_1 \in \mathbb{N}$ *such that for any* $t \in \mathcal{T}$ *the ball* $\mathbb{B}(t, 1)$ *in* $\mathbb{R}^r$ *is intersected by at most* N_1 *different subsets* $\mathcal{B}_i$;
(2) *the restriction* $\Phi|_{\mathcal{B}_i}$ *of* Φ *to* $\mathcal{B}_i$ *is Lipschitz continuous with a constant* $L_\Phi > 0$ *that does not depend on* i.

Furthermore, the function Φ *is lower semicontinuous and piecewise polyhedral on* $\mathcal{T}$ *and there exist constants* $a, b > 0$ *such that it holds for all* $t, \tilde{t} \in \mathcal{T}$:

$$|\Phi(t) - \Phi(\tilde{t})| \le a\|t - \tilde{t}\| + b.$$

Part (i) of the lemma was proved in Section 5.6 of Bank et al. (1982) and in Lemma 2.5 of Schultz (1996), (ii) was derived as Lemma 2.3 in Schultz (1996) and the remaining properties of Φ were established in Blair and Jeroslow (1977). Compared to Lemma 21 for optimal value functions of linear programs without integrality requirements, the representation of Φ is now given on countably many (possibly unbounded) Borel sets. This requires to incorporate the tail behaviour of P and leads to the following representation of the function $F_0(x, \xi) := \langle c, x\rangle + \Phi(h(\xi) - T(\xi)x)$ for each pair (x, ξ) in $X \times \Xi$.

Proposition 34. *Assume* (B1)–(B3) *and let* $\mathcal{U}$ *be an open bounded subset of* $\mathbb{R}^m$. *For each* $R \ge 1$ *and* $x \in X \cap \operatorname{cl}\mathcal{U}$ *there exist disjoint Borel subsets* $\Xi_{j,x}^R$ *of* Ξ, $j = 1, \ldots, \nu$, *whose closures are polyhedra with a uniformly bounded number of faces such that the function*

$$F_0(x, \xi) = \sum_{j=0}^{\nu} (\langle c, x\rangle + \Phi(h(\xi) - T(\xi)x)\chi_{\Xi_{j,x}^R}(\xi) \quad ((x, \xi) \in X \times \Xi)$$

is Lipschitz continuous with respect to ξ *on each* $\Xi_{j,x}^R$, $j = 1, \ldots, \nu$, *with some uniform Lipschitz constant. Here,* $\Xi_{0,x}^R := \Xi \setminus \cup_{j=1}^{\nu} \Xi_{j,x}^R$ *is contained in* $\{\xi \in \mathbb{R}^s : \|\xi\| > R\}$ *and* ν *is bounded by a multiple of* R^r.

Proof. Since $h(\cdot)$ and $T(\cdot)$ are affine linear functions, there exists a constant $C_2 > 0$ such that the estimate $\|h(\xi) - T(\xi)x\|_\infty \le C_2 \max\{1, \|\xi\|\}$ holds for each

pair in $X \cap \mathrm{cl}\,\mathcal{U}$. Let $R > 0$ and $\mathcal{T}_R := \mathcal{T} \cap RC_2\mathbb{B}_\infty$, where $\mathbb{B}_\infty$ refers to the closed unit ball in $\mathbb{R}^r$ with respect to the norm $\|\cdot\|_\infty$. Now, we partition the ball $RC_2\mathbb{B}_\infty$ into disjoint Borel sets whose closures are $\mathbb{B}_\infty$-balls with radius 1, where possible gaps are filled with maximal balls of radius less than 1. Then the number of elements in this partition of $RC_2\mathbb{B}_\infty$ is bounded above by $(2RC_2)^r$. From Lemma 33 (i) we know that each element of this partition is intersected by at most N_1 subsets $\mathcal{B}_i$ (for some $N_1 \in \mathbb{N}$). Another consequence of Lemma 33 (i) is that each $\mathcal{B}_i$ splits into disjoint Borel subsets whose closures are polyhedra. Moreover, the number of such subsets can be bounded from above by a constant not depending on i. Hence, there exist a number $\nu \in \mathbb{N}$ and disjoint Borel subsets $\{B_j : j = 1, \ldots, \nu\}$ such that their closures are polyhedra, their union contains $\mathcal{T}_R$, and ν is bounded above by κR^r, where the constant $\kappa > 0$ is independent of R. Now, let $x \in X \cap \mathrm{cl}\,\mathcal{U}$ and consider the following disjoint Borel subsets of Ξ:

$$\Xi^R_{j,x} := \{\xi \in \Xi : h(\xi) - T(\xi)x \in B_j\} \quad (j = 1, \ldots, \nu),$$
$$\Xi^R_{0,x} := \Xi \setminus \bigcup_{j=1}^{\nu} \Xi^R_{j,x} \subseteq \{\xi \in \Xi : \|h(\xi) - T(\xi)x\|_\infty > RC_2\} \subseteq \{\xi \in \Xi : \|\xi\| > R\}.$$

For each $j = 1, \ldots, \nu$ the closures of the sets B_j are polyhedra with a number of faces that is bounded above by some number not depending on j, ν and R. Hence, the same is true for the closures of the sets $\Xi^R_{j,x}$, i.e., for $\{\xi \in \Xi : h(\xi) - T(\xi)x \in \mathrm{cl}\, B_j\}$, where, moreover, the corresponding number $k \in \mathbb{N}$ does not depend on $x \in X \cap \mathrm{cl}\,\mathcal{U}$. Finally, we conclude from Lemma 33 that there exists a constant $L_1 > 0$ (which does not depend on $x \in X \cap \mathrm{cl}\,\mathcal{U}$, $j = 1, \ldots, \nu$ and $R > 0$) such that the function $F_0(x, \cdot)|_{\Xi^R_{j,x}} = cx + \Phi|_{B_j}(h(\cdot) - T(\cdot)x)$ is Lipschitz continuous with constant L_1. □

For further structural properties of model (3.34) we refer to Louveaux and Schultz (2003). In order to state stability results for model (3.34), we consider the following probability metrics with ζ-structure on $\mathcal{P}_1(\Xi)$ for every $k \in \mathbb{N}$:

$$\zeta_{1,\mathrm{ph}_k}(P, Q) := \sup\left\{\left|\int_B f(\xi)(P - Q)(\mathrm{d}\xi)\right| : f \in \mathcal{F}_1(\Xi),\ B \in \mathcal{B}_{\mathrm{ph}_k}(\Xi)\right\}$$
$$= \sup\left\{\left|\int_\Xi f(\xi)\chi_B(\xi)(P - Q)(\mathrm{d}\xi)\right| : f \in \mathcal{F}_1(\Xi),\ B \in \mathcal{B}_{\mathrm{ph}_k}(\Xi)\right\}. \quad (3.36)$$

Here, $\mathcal{B}_{\mathrm{ph}_k}(\Xi)$ and $\mathcal{F}_1(\Xi)$ denote the sets of polyhedra in Ξ and of Lipschitz continuous functions from Ξ to $\mathbb{R}$ introduced in Section 2.1.

Theorem 35. *Let the conditions* (B1)–(B4) *be satisfied*, $X^*(P)$ *be nonempty and* $\mathcal{U} \subseteq \mathbb{R}^m$ *be an open bounded neighbourhood of* $X^*(P)$.

Then there exist constants $L>0$, $\delta>0$ *and* $k\in\mathbb{N}$ *such that*

$$|\vartheta(P)-\vartheta_{\mathcal{U}}(Q)|\le L\phi_P(\zeta_{1,\mathrm{ph}_k}(P,Q))$$
$$\emptyset\ne X^*_{\mathcal{U}}(Q)\subseteq X^*(P)+\Psi_P(L\phi_P(\zeta_{1,\mathrm{ph}_k}(P,Q)))\mathbb{B}, \tag{3.37}$$

and $X^*_{\mathcal{U}}(Q)$ *is a CLM set of* (3.34) *relative to* $\mathcal{U}$ *whenever* $Q\in\mathcal{P}_1(\Xi)$ *and* $\zeta_{1,\mathrm{ph}_k}(P,Q)<\delta$. *Here, the function* ϕ_P *on* $\mathbb{R}_+$ *is defined by*

$$\phi_P(0)=0 \quad\text{and}\quad \phi_P(t):=\inf_{R\ge 1}\left\{R^r t+\int_{\{\xi\in\Xi:\ \|\xi\|>R\}}\|\xi\|\,\mathrm{d}P(\xi)\right\}\quad (t>0)$$

and continuous at $t=0$, *and the function* Ψ_P *is given by* (2.23).

If P *has a finite absolute moment of pth order for some* $p>1$, *the estimate* $\phi_P(t)\le Ct^{(p-1)/(p-1+r)}$ *holds for small* $t>0$ *and some constant* $C>0$.

Proof. Since the function Φ is lower semicontinuous on $\mathcal{T}$ (Lemma 33), F_0 is lower semicontinuous on $X\times\Xi$ and, hence, a random lower semicontinuous function (Example 14.31 in Rockafellar and Wets, 1998). Using Lemma 33 we obtain the estimate

$$|F_0(x,\xi)|\le\|c\|\,\|x\|+a(\|h(\xi)\|+\|T(\xi)\|\,\|x\|)+b$$

for each pair $(x,\xi)\in X\times\Xi$. Since $h(\xi)$ and $T(\xi)$ depend affine linearly on ξ, there exists a constant $C_1>0$ such that $|F_0(x,\xi)|\le C_1\max\{1,\|\xi\|\}$ holds for each pair $(x,\xi)\in(X\cap\mathrm{cl}\,\mathcal{U})\times\Xi$. Hence, $\mathcal{P}_{\mathcal{F}_{\mathcal{U}}}(\Xi)\supseteq\mathcal{P}_1(\Xi)$ and Theorems 5 and 9 apply with $d=0$ and the distance $d_{\mathcal{F}_{\mathcal{U}}}$ on $\mathcal{P}_1(\Xi)$.

From Proposition 34 we know that, for each $R\ge 1$ and $x\in X\cap\mathrm{cl}\,\mathcal{U}$, there exist Borel subsets $\Xi^R_{j,x}$, $j=1,\dots,\nu$, of Ξ such that the function $f^R_{j,x}(\cdot):=F_0(x,\cdot)|_{\Xi^R_{j,x}}$ is Lipschitz continuous on $\Xi^R_{j,x}$ with some Lipschitz constant $L_1>0$ (not depending on x, j and R). We extend each function $f^R_{j,x}(\cdot)$ to the whole of Ξ by preserving the Lipschitz constant L_1. Proposition 34 also implies that the closures of $\Xi^R_{j,x}$ are contained in $\mathcal{B}_{\mathrm{ph}_k}(\Xi)$ for some $k\in\mathbb{N}$, that the number ν is bounded above by κR^r, where the constant $\kappa>0$ is independent on R, and that $\Xi^R_{0,x}:=\Xi\setminus\cup_{j=1}^{\nu}\Xi^R_{j,x}$ is a subset of $\{\xi\in\Xi:\|\xi\|>R\}$.

For each $Q\in\mathcal{P}_1(\Xi)$ and $x\in X\cap\mathrm{cl}\,\mathcal{U}$ we obtain

$$\begin{aligned}\left|\int_\Xi F_0(x,\xi)\,\mathrm{d}(P-Q)(\xi)\right|&=\left|\sum_{j=0}^{\nu}\int_{\Xi^R_{j,x}}F_0(x,\xi)\,\mathrm{d}(P-Q)(\xi)\right|\\&\le\sum_{j=1}^{\nu}\left|\int_{\Xi^R_{j,x}}f^R_{j,x}(\xi)\,\mathrm{d}(P-Q)(\xi)\right|+I^R_x(P,Q)\\&\le\nu L_1\sup_{\substack{f\in\mathcal{F}_1(\Xi)\\ j=1,\dots,\nu}}\left|\int_\Xi f(\xi)\chi_{\Xi^R_{j,x}}\,\mathrm{d}(P-Q)(\xi)\right|+I^R_x(P,Q),\end{aligned}$$

where $I_x^R(P,Q) := |\int_{\Xi_{0,x}^R} F_0(x,\xi)\,\mathrm{d}(P-Q)(\xi)|$.

For each $\Xi_{j,x}^R$ we now consider a sequence of polyhedra $B_{j,x}^R$, which are contained in $\Xi_{j,x}^R$ and belong to $\mathcal{B}_{\mathrm{ph}_k}(\Xi)$, such that their characteristic functions $\chi_{B_{j,x}^R}$ converge pointwise to the characteristic function $\chi_{\Xi_{j,x}^R}$. Then the sequence consisting of the elements $|\int_\Xi f(\xi)\chi_{B_{j,x}^R}(\xi)\,\mathrm{d}(P-Q)(\xi)|$ converges to $|\int_\Xi f(\xi)\chi_{\Xi_{j,x}^R}(\xi)\,\mathrm{d}(P-Q)(\xi)|$ while each element is bounded by $\zeta_{1,\mathrm{ph}_k}(P,Q)$. Hence, the above estimate may be continued to

$$\left|\int_\Xi F_0(x,\xi)\,\mathrm{d}(P-Q)(\xi)\right| \le \kappa L_1 R^r \zeta_{1,\mathrm{ph}_k}(P,Q) + I_x^R(P,Q). \tag{3.38}$$

For the term $I_x^R(P,Q)$ we have

$$\begin{aligned} I_x^R(P,Q) &\le C_1 \int_{\{\xi\in\Xi:\ \|\xi\|>R\}} \|\xi\|\,\mathrm{d}(P+Q)(\xi) \\ &\le C_1 \int_{\{\xi\in\Xi:\ \|\xi\|_\infty > R/C_2\}} \|\xi\|\,\mathrm{d}(P+Q)(\xi) \end{aligned}$$

where we have used the estimate $|F_0(x,\xi)| \le C_1\|\xi\|$ for each pair $(x,\xi) \in (X\cap \mathrm{cl}\,\mathcal{U})\times\{\xi\in\Xi:\ \|\xi\|>R\}$ and $C_2>0$ is a norming constant such that $\|\xi\| \le C_2\|\xi\|_\infty$ holds for each $\xi\in\mathbb{R}^s$. Clearly, the set $\{\xi\in\Xi:\ \|\xi\|_\infty > \frac{R}{C_2}\}$ can be covered by 2^s intersections of Ξ by open halfspaces whose closures belong to $\mathcal{B}_{\mathrm{ph}_k}(\Xi)$. Hence, a similar argument as the one above yields the estimate

$$\int_{\{\xi\in\Xi:\ \|\xi\|_\infty > R/C_2\}} \|\xi\|\,\mathrm{d}Q(\xi) \le 2^s \zeta_{1,\mathrm{ph}_k}(P,Q) + \int_{\{\xi\in\Xi:\ \|\xi\|_\infty > R/C_2\}} \|\xi\|\,\mathrm{d}P(\xi).$$

Hence, from the previous estimates we obtain that

$$\begin{aligned} d_{\mathcal{F}_\mathcal{U}}(P,Q) &\le \kappa(L_1R^r + 2^sC_1)\zeta_{1,\mathrm{ph}_k}(P,Q) + 2C_1\int_{\{\xi\in\Xi:\ \|\xi\|_\infty > R/C_2\}} \|\xi\|\mathrm{d}P(\xi) \\ &\le CR^r\zeta_{1,\mathrm{ph}_k}(P,Q) + \int_{\{\xi\in\Xi:\ \|\xi\|>\alpha R\}} \|\xi\|\,\mathrm{d}P(\xi) \end{aligned}$$

for some constants $C>0$ and $\alpha\in(0,1)$, the latter depending on the norming constants of $\|\cdot\|$ and $\|\cdot\|_\infty$, respectively. Finally, we obtain

$$d_{\mathcal{F}_\mathcal{U}}(P,Q) \le \hat{C}\phi_P(\zeta_{1,\mathrm{ph}_k}(P,Q)), \tag{3.39}$$

where

$$\phi_P(0) := 0 \quad \text{and} \quad \phi_P(t) := \inf_{R \geq 1} \left\{ R^r t + \int_{\{\xi \in \Xi:\ \|\xi\| > R\}} \|\xi\| \, \mathrm{d}P(\xi) \right\} \quad (t > 0) \tag{3.40}$$

with some constant $\hat{C} > 0$. Now, the result is a consequence of the Theorem 5 and Theorem 9. If $\int_\Xi \|\xi\|^p \, \mathrm{d}P(\xi) < \infty$, it holds that $\int_{\{\xi \in \Xi:\ \|\xi\| > R\}} \|\xi\| \, \mathrm{d}P(\xi) \leq R^{1-p} \int_\Xi \|\xi\|^p \, \mathrm{d}P(\xi)$ by Markov's inequality. The desired estimate follows by inserting $R = t^{-1/(p+r-1)}$ for small $t > 0$ into the function whose infimum w.r.t. $R \geq 1$ is $\phi_P(t)$. □

In case that the underlying distribution P and its perturbations Q have supports in some bounded subset of $\mathbb{R}^s$, the stability result improves slightly.

Corollary 36. *Let the conditions* (B1)–(B3) *be satisfied and* Ξ *be bounded. Assume that* $P \in \mathcal{P}(\Xi)$, $X^*(P)$ *is nonempty and* $\mathcal{U} \subseteq \mathbb{R}^m$ *is an open bounded neighbourhood of* $X^*(P)$.

Then there exist constants $L > 0$, $\delta > 0$ *and* $k \in \mathbb{N}$ *such that*

$$\begin{aligned} &|\vartheta(P) - \vartheta_{\mathcal{U}}(Q)| \leq L \zeta_{1,\mathrm{ph}_k}(P, Q) \\ &\quad \emptyset \neq X^*_{\mathcal{U}}(Q) \subseteq X^*(P) + \Psi_P(L \zeta_{1,\mathrm{ph}_k}(P, Q)) \mathbb{B}, \end{aligned}$$

and $X^*_{\mathcal{U}}(Q)$ *is a CLM set of* (3.34) *relative to* $\mathcal{U}$ *whenever* $Q \in \mathcal{P}(\Xi)$ *and* $\zeta_{1,\mathrm{ph}_k}(P, Q) < \delta$.

Proof. Since Ξ is bounded, we have $\mathcal{P}_1(\Xi) = \mathcal{P}(\Xi)$. Moreover, the function $\phi_P(t)$ can be estimated by $R^r t$ for some sufficiently large $R > 0$. Hence, Theorem 35 implies the assertion. □

Remark 37. Since $\Xi \in \mathcal{B}_{\mathrm{ph}_k}(\Xi)$ for some $k \in \mathbb{N}$, we obtain from (3.36) by choosing $B := \Xi$ and $f \equiv 1$, respectively,

$$\max\{\zeta_1(P, Q), \alpha_{\mathrm{ph}_k}(P, Q)\} \leq \zeta_{1,\mathrm{ph}_k}(P, Q) \tag{3.41}$$

for large k and all $P, Q \in \mathcal{P}_1(\Xi)$. Here, α_{ph_k} denotes the polyhedral discrepancy (see Section 2.1). Hence, convergence with respect to ζ_{1,ph_k} implies weak convergence, convergence of first order absolute moments and convergence with respect to the polyhedral discrepancy α_{ph_k}. The converse is also true. The latter observation is a consequence of the estimate

$$\zeta_{1,\mathrm{ph}_k}(P, Q) \leq C_s \alpha_{\mathrm{ph}_k}(P, Q)^{1/(s+1)} \quad (P, Q \in \mathcal{P}(\Xi)) \tag{3.42}$$

for some constant $C_s > 0$. It is valid for bounded $\Xi \subset \mathbb{R}^s$ and can be derived by using the technique in the proof of Proposition 3.1 in Schultz (1996). In view of (3.41) and (3.42) the metric ζ_{1,ph_k} is stronger than α_{ph_k} in general, but in case of bounded Ξ both metrize the same topology on $\mathcal{P}(\Xi)$.

For more specific models (3.34), improvements of the above results are possible. The potential of such improvements consists in exploiting specific recourse structures, i.e., in additional information on the shape of the sets $\mathcal{B}_i$ in Lemma 33 and on the behaviour of the (value) function Φ on these sets. These considerations may lead to stability results with respect to probability metrics that are (much) weaker than ζ_{1,ph_k}. To illustrate such an improvement, let us consider the case of pure integer recourse where Φ is given by

$$\Phi(t) = \min\{\langle q, y\rangle : Wy \geq t,\ y \in \mathbb{Z}^{\hat{m}}\}, \tag{3.43}$$

the technology matrix is fixed and the right-hand side is fully stochastic, i.e., $T(\xi) \equiv T$ and $h(\xi) \equiv \xi$. This situation fits into the general model (3.34) by setting $\overline{q} = 0$, $\overline{m} = r$ and $\overline{W} = -I_r$, with I_r denoting the (r, r)-identity matrix. For such models Schultz (1996) observed that stability holds with respect to the Kolmogorov metric d_{K} on $\mathcal{P}(\Xi)$.

Corollary 38. *Let Φ be given by (3.43), $T(\xi) \equiv T$, $h(\xi) \equiv \xi$ and Ξ be bounded. Furthermore, let the conditions* (B1)–(B3) *be satisfied with $\mathcal{T} = \mathbb{R}^s$. Assume that $P \in \mathcal{P}(\Xi)$, $X^*(P)$ is nonempty and $\mathcal{U} \subseteq \mathbb{R}^m$ is an open bounded neighbourhood of $X^*(P)$. Then there exist constants $L > 0$ and $\delta > 0$ such that*

$$|\vartheta(P) - \vartheta_{\mathcal{U}}(Q)| \leq L d_{\mathrm{K}}(P, Q)$$
$$\emptyset \neq X^*_{\mathcal{U}}(Q) \subseteq X^*(P) + \Psi_P(L d_{\mathrm{K}}(P, Q))\mathbb{B},$$

and $X^*_{\mathcal{U}}(Q)$ is a CLM set of (3.34) relative to $\mathcal{U}$ whenever $Q \in \mathcal{P}(\Xi)$ and $d_{\mathrm{K}}(P, Q) < \delta$. Here, the function Ψ_P is given by (2.23).

Proof. The assumptions imply that Φ is even constant on $\mathcal{B}_i$ for each $i \in \mathbb{N}$ and the continuity regions of Φ are rectangular (see Schultz, 1996). Without loss of generality the set Ξ may be chosen to be rectangular. Then the sets $\Xi^R_{j,x}$ in Proposition 34 are also bounded rectangular sets and $F_0(x, \cdot)$ is constant on each $\Xi^R_{j,x}$. Hence, the estimate (3.38) takes the form

$$\left| \int_{\Xi} F_0(x, \xi)\, \mathrm{d}(P - Q)(\xi) \right| \leq \kappa L_1 R^s \alpha_{\mathrm{box}}(P, Q),$$

where $\alpha_{\text{box}}(P, Q) := \sup\{|P(B) - Q(B)| : B \text{ is a box in } \mathbb{R}^s\}$. Finally, we use the known estimate

$$\alpha_{\text{box}}(P, Q) \leq C d_{\text{K}}(P, Q)$$

for some constant $C > 0$ and derive the result from Theorem 35. □

3.3 Linear chance constrained programs

In this section, we study consequences of the general stability analysis of Section 2 for linear chance constrained stochastic programs of the form

$$\min\{\langle c, x\rangle : x \in X,\ P(\{\xi \in \Xi : T(\xi)x \geq h(\xi)\}) \geq p\}, \tag{3.44}$$

where $c \in \mathbb{R}^m$, X and Ξ are polyhedra in $\mathbb{R}^m$ and $\mathbb{R}^s$, respectively, $p \in (0, 1)$, $P \in \mathcal{P}(\Xi)$ and the right-hand side $h(\xi) \in \mathbb{R}^r$ and the (r, m)-matrix $T(\xi)$ depend affine linearly on $\xi \in \Xi$.

We set $d = 1$, $F_0(x, \xi) = \langle c, x\rangle$, $F_1(x, \xi) = p - \chi_H(x)(\xi)$ where $H(x) = \{\xi \in \Xi : T(\xi)x \geq h(\xi)\}$ and $\chi_{H(x)}$ its characteristic function, and observe that the program (3.44) is a special case of the general stochastic program (1.1). We note that the set $H(x)$ is polyhedral for each $x \in X$. In fact, these sets are given as the finite intersection of r closed half-spaces. Furthermore, the multifunction H from $\mathbb{R}^m$ to $\mathbb{R}^s$ has a closed graph and, hence, the mapping $(x, \xi) \mapsto \chi_{H(x)}(\xi)$ from $\mathbb{R}^m \times \Xi$ to $\mathbb{R}$ is upper semicontinuous. This implies that F_1 is lower semicontinuous on $\mathbb{R}^m \times \Xi$ and, hence, a random lower semicontinuous function (Example 14.31 in Rockafellar and Wets, 1998). Moreover, we have $p - 1 \leq F_1(x, \xi) \leq p$ for any pair (x, ξ). By specifying the general class of probability measures and the minimal information probability metric in Section 2.2 we obtain

$$\begin{aligned}
\mathcal{P}_{\mathcal{F}_{\mathcal{U}}}(\Xi) &= \left\{ Q \in \mathcal{P}(\Xi) : \sup_{x \in X \cap \text{cl}\,\mathcal{U}} \max_{j=0,1} \left| \int_\Xi F_j(x, \xi)\, \mathrm{d}Q(\xi) \right| < \infty \right\} = \mathcal{P}(\Xi) \\
d_{\mathcal{F}_{\mathcal{U}}}(P, Q) &= \sup_{x \in X \cap \text{cl}\,\mathcal{U}} \max_{j=0,1} \left| \int_\Xi F_j(x, \xi)(P - Q)(\mathrm{d}\xi) \right| \\
&= \sup_{x \in X \cap \text{cl}\,\mathcal{U}} \left| P(H(x)) - Q(H(x)) \right|
\end{aligned}$$

for each $P, Q \in \mathcal{P}(\Xi)$ and any nonempty, open and bounded subset $\mathcal{U}$ of $\mathbb{R}^m$. Due to the polyhedrality of the sets $H(x)$ for any $x \in \mathbb{R}^m$, the polyhedral discrepancies α_{ph_k} on $\mathcal{P}(\Xi)$ for every $k \in \mathbb{N}$ (see Section 2.1) or related discrepancies appear as natural candidates for suitable probability metrics in case of model (3.44). The following result is an immediate consequence of the general methodology in Section 2.

Theorem 39. *Let $P \in \mathcal{P}(\Xi)$ and assume that*

(i) *$X^*(P) \neq \emptyset$ and $\mathcal{U} \subseteq \mathbb{R}^m$ is an open bounded neighbourhood of $X^*(P)$,*
(ii) *the mapping $x \mapsto \{y \in \mathbb{R}\colon P(\{\xi \in \Xi\colon T(\xi)x \geq h(\xi)\}) \geq p - y\}$ is metrically regular at each pair $(\overline{x}, 0)$ with $\overline{x} \in X^*(P)$.*

Then there exist constants $L > 0$, $\delta > 0$ and $k \in \mathbb{N}$ such that

$$|\vartheta(P) - \vartheta_{\mathcal{U}}(Q)| \leq L\alpha_{\mathrm{ph}_k}(P, Q)$$

$$\emptyset \neq X^*_{\mathcal{U}}(Q) \subseteq X^*(P) + \Psi_P(L\alpha_{\mathrm{ph}_k}(P, Q))\mathbb{B},$$

and $X^*_{\mathcal{U}}(Q)$ is a CLM set of (3.44) relative to $\mathcal{U}$ whenever $Q \in \mathcal{P}(\Xi)$ and $\alpha_{\mathrm{ph}_k}(P, Q) < \delta$. Here, the function Ψ_P is given by (2.23).

Proof. All sets $H(x)$ are polyhedra in $\mathbb{R}^s$ given by r linear inequalities. Hence, the number of faces of $H(x)$ is bounded by some $k \in \mathbb{N}$ not depending on $x \in \mathbb{R}^m$. Since all assumptions of Theorem 5 are satisfied for the special situation considered here, the result follows from the Theorems 5 and 9 by taking into account the estimate $d_{\mathcal{F}_{\mathcal{U}}}(P, Q) \leq \alpha_{\mathrm{ph}_k}(P, Q)$. □

We show that Theorem 39 applies to many chance constrained models known from the literature. First we discuss the metric regularity property (ii) of the original probabilistic constraint in (3.44). The following example shows that condition (ii) is indispensable for Theorem 39 to hold.

Example 40. Let $P \in \mathcal{P}(\mathbb{R})$ have a distribution function F_P which is continuously differentiable and satisfies $F_P(x) = x^{2s+1} + p$ for all x in a neighbourhood of $x = 0$ and some $p \in (0, 1)$ and $s \in \mathbb{N}$. Let us consider the model

$$\min\{x\colon x \in \mathbb{R},\ P(\xi \leq x) = F_P(x) \geq p\}.$$

Then the condition $\nabla F_P(\overline{x}) \neq 0$ is necessary and sufficient for the metric regularity at $\overline{x}$ with $F_P(\overline{x}) = p$ (Example 9.44 in Rockafellar and Wets, 1998). Clearly, this condition is violated at the minimizer $\overline{x} = 0$. To show that the result gets lost, we consider the measures $P_n = (1 - \frac{1}{n})P + \frac{1}{n}\,\delta_{1/n}$, $n \in \mathbb{N}$. The sequence (P_n) converges weakly to P and, thus, it converges with respect to the Kolmogorov metric d_{K} as P is continuous. Then $|\vartheta(P) - \vartheta(P_n)| = (\frac{p}{n-1})^{1/(2s+1)} =: x_n$, but $d_{\mathrm{K}}(P, P_n) \geq |F_P(x_n) - F_{P_n}(x_n)| = \frac{p}{n-1}$.

When looking for general conditions implying (ii), one has to resort to results for nonconvex and nondifferentiable situations. The function

$$g(x) := P(\{\xi \in \Xi\colon T(\xi)x \geq h(\xi)\})$$

from $\mathbb{R}^m$ into $\mathbb{R}$ is known to be upper semicontinuous (Proposition 3.1 in Römisch and Schultz, 1991c). However, g happens to be nondifferentiable or even discontinuous not only in cases where the probability distribution P is discrete, but even if $T(\xi)$ is nonstochastic and P is continuous.

Example 41. Let P be the standard normal distribution with distribution function Φ. First let $T(\xi) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $h(\xi) = \begin{pmatrix} \xi \\ 0 \end{pmatrix}$ for each $\xi \in \mathbb{R}$. Then

$$g(x) = P(\{\xi \in \mathbb{R}: x \geq \xi, x \geq 0\}) = \begin{cases} 0, & x < 0 \\ \Phi(x), & x \geq 0 \end{cases}.$$

Secondly, let $T(\xi) = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ and $h(\xi) = \begin{pmatrix} \xi \\ \xi \end{pmatrix}$ for each $\xi \in \mathbb{R}$. Then we have

$$g(x) = P(\{\xi \in \mathbb{R}: x \geq \xi, \ -x \geq \xi\}) = \Phi(\min\{-x, x\}).$$

We also refer to Example 9 in Henrion and Römisch (1999) for a probability distribution P having a (bounded) continuous density on $\Xi = \mathbb{R}^2$, but a probability distribution function (i.e., g in case of $T(\xi) = I$ and $h(\xi) = \xi$) that is not locally Lipschitz continuous.

Hence, one has to go back to tools from nonsmooth analysis in general. For example, if the function g is locally Lipschitz continuous on $\mathbb{R}^m$, condition (ii) is satisfied if the constraint qualification

$$\partial(-g)(\overline{x}) \cap (-N_X(\overline{x})) = \emptyset \tag{3.45}$$

holds at each $\overline{x} \in X^*(P)$ with $g(\overline{x}) = p$ (Corollary 4.2 in Mordukhovich, 1994b). Here, the symbol ∂ stands for the Mordukhovich subdifferential (cf. Mordukhovich, 1994a) and $N_X(\overline{x}) := \{x^* \in \mathbb{R}^m: \langle x^*, x - \overline{x}\rangle \leq 0, \forall x \in X\}$ is the normal cone to the polyhedral set X at $\overline{x} \in X$.

For more specific structures of probabilistic constraints, even in case of a stochastic matrix $T(\xi)$, the situation may become much more comfortable if P is a multivariate normal distribution. To demonstrate this, we consider the case $\Xi = \mathbb{R}^{m+1}$, $T(\xi)x = \sum_{i=1}^m \xi_i x_i$, i.e., $T(\xi)$ consists of one single row, and $h(\xi) = \xi_{m+1}$. Then $H(x)$ takes the form

$$H(x) = \left\{\xi \in \mathbb{R}^{m+1}: \sum_{i=1}^m \xi_i x_i \geq \xi_{m+1}\right\} \tag{3.46}$$

for each $x \in \mathbb{R}^m$, i.e., the sets $H(x)$ are closed half-spaces in $\mathbb{R}^{m+1}$.

Corollary 42. *Let P be a normal distribution on $\mathbb{R}^{m+1}$ with mean $\mu \in \mathbb{R}^{m+1}$ and nonsingular covariance matrix $\Sigma \in \mathbb{R}^{(m+1)\times(m+1)}$, H be given by (3.46) and $p \in (\frac{1}{2}, 1)$. Let $X^*(P)$ be nonempty and $\mathcal{U} \subseteq \mathbb{R}^m$ be an open bounded neighbourhood of $X^*(P)$. Assume that there exists an $\hat{x} \in X$ such that $P(H(\hat{x})) > p$.*

Then there are constants $L > 0$ and $\delta > 0$ such that

$$|\vartheta(P) - \vartheta_{\mathcal{U}}(Q)| \leq L\alpha_h(P, Q)$$

$$\emptyset \neq X^*_{\mathcal{U}}(Q) \subseteq X^*(P) + \Psi_P(L\alpha_h(P, Q))\mathbb{B}$$

holds and $X^*_{\mathcal{U}}(Q)$ is a CLM set for (3.44) relative to $\mathcal{U}$ for each $Q \in \mathcal{P}(\Xi)$ with $\alpha_h(P, Q) < \delta$. Here, the function Ψ_P is given by (2.23) and α_h is the half-space discrepancy (see Section 2.1).

Proof. For any $x \in \mathbb{R}^m$, we set $x' := (x_1, \ldots, x_m, -1)$ and $\sigma(x) := \langle \Sigma x', x' \rangle^{1/2}$. Let Φ denote the standard normal distribution function and ϕ the standard normal density. Then $\langle \xi, x' \rangle$ is normal with mean $\langle \mu, x' \rangle$ and standard deviation $\sigma(x') > 0$ (due to the nonsingularity of Σ), and

$$g(x) = P(\{\xi \in \mathbb{R}^{m+1} : \langle \xi, x' \rangle \geq 0\}) = \Phi\left(\frac{\langle \mu, x' \rangle}{\sigma(x')}\right)$$

holds for any $x \in \mathbb{R}^m$. Further, the function

$$\hat{g}(x) := \langle \mu, x' \rangle - \Phi^{-1}(p)\sigma(x') = [\Phi^{-1}(g(x)) - \Phi^{-1}(p)]\sigma(x')$$

is concave on $\mathbb{R}^m$ due to $\Phi^{-1}(p) > 0$ and continuously differentiable on $\mathbb{R}^m$ with gradient

$$\nabla \hat{g}(x) = \frac{\sigma(x')}{\phi(g(x))} \nabla g(x) + [\Phi^{-1}(g(x)) - \Phi^{-1}(p)] \nabla \sigma(x') \begin{pmatrix} I_m \\ 0 \end{pmatrix}.$$

Let $\overline{x} \in X$ be such that $g(\overline{x}) = p$ and $\hat{x} \in X$ be the element having the property $P(H(\hat{x})) > p$ or, equivalently, $\hat{g}(\hat{x}) > 0$. Then the concavity of $\hat{g}$ implies $\langle \nabla \hat{g}(\overline{x}), \hat{x} - \overline{x} \rangle > 0$ and, thus, $\nabla \hat{g}(\overline{x}) \notin N_X(\overline{x})$. Due to the equation $\nabla \hat{g}(\overline{x}) = \frac{\sigma(\overline{x}')}{\phi(g(\overline{x}))} \nabla g(\overline{x})$, we conclude $\nabla g(\overline{x}) \notin N_X(\overline{x})$. Hence, the constraint qualification (3.45) and, thus, condition (ii) of Theorem 39 are satisfied. □

For the remainder of this section we assume that the technology matrix $T(\cdot)$ is fixed, i.e., $T(\xi) \equiv T$. We will show that the constraint qualification of Corollary 42, i.e., $P(H(\hat{x})) > p$ for some $\hat{x} \in X$, implies condition (ii) of Theorem 39 for any r-concave probability distribution.

To recall the notion of r-concavity, we introduce first the generalized mean function m_r on $\mathbb{R}_+ \times \mathbb{R}_+ \times [0,1]$ for $r \in [-\infty, \infty]$ by

$$m_r(a, b; \lambda) := \begin{cases} (\lambda a^r + (1-\lambda)b^r)^{1/r}, & r \in (0, \infty) \text{ or } r \in (-\infty, 0),\ ab > 0, \\ 0, & ab = 0,\ r \in (-\infty, 0), \\ a^\lambda b^{1-\lambda}, & r = 0, \\ \max\{a, b\}, & r = \infty, \\ \min\{a, b\}, & r = -\infty. \end{cases} \tag{3.47}$$

A measure $P \in \mathcal{P}(\mathbb{R}^s)$ is called *r-concave* for some $r \in [-\infty, \infty]$ (cf. Prekopa, 1995) if the inequality

$$P(\lambda B_1 + (1-\lambda)B_2) \geq m_r(P(B_1), P(B_2);\ \lambda)$$

holds for all $\lambda \in [0,1]$ and all convex Borel subsets B_1, B_2 of $\mathbb{R}^s$ such that $\lambda B_1 + (1-\lambda)B_2$ is Borel. For $r = 0$ and $r = -\infty$, P is also called *logarithmic concave* and *quasi-concave*, respectively. Since $m_r(a, b; \lambda)$ is increasing in r if all the other variables are fixed, the sets of all r-concave probability measures are increasing if r is decreasing. It is known that $P \in \mathcal{P}(\mathbb{R}^s)$ is r-concave for some $r \in [-\infty, 1/s]$ if P has a density f_P such that

$$f_P(\lambda z + (1-\lambda)\tilde{z}) \geq m_{r(s)}(f_P(z), f_P(\tilde{z});\ \lambda), \tag{3.48}$$

where $r(s) = r(1-rs)^{-1}$, holds for all $\lambda \in [0,1]$ and $z, \tilde{z} \in \mathbb{R}^s$. Let us mention that many multivariate probability distributions are r-concave for some $r \in (-\infty, \infty]$, e.g., the uniform distribution (on some bounded convex set), the (nondegenerate) multivariate normal distribution, the Dirichlet distribution, the multivariate Student and Pareto distributions (see Prekopa, 1995).

The key observation of r-concave measures in the context of probabilistic constraints is the following one.

Lemma 43. *Let H be a multifunction from $\mathbb{R}^m$ to $\mathbb{R}^s$ with closed convex graph and P be r-concave for some $r \in [-\infty, \infty]$. Then the function $g := P(H(\cdot))$ from $\mathbb{R}^m$ to $\mathbb{R}$ has the property*

$$g(\lambda x + (1-\lambda)\tilde{x}) \geq m_r(g(x), g(\tilde{x});\ \lambda)$$

for each $x, \tilde{x} \in \mathbb{R}^m$ and $\lambda \in [0,1]$.

Proof. In particular, $H(x)$ is a closed convex subset of $\mathbb{R}^s$ for any $x \in \mathbb{R}^m$. Let $x, \tilde{x} \in \mathbb{R}^m$ and $\lambda \in [0,1]$. Then the set $\lambda H(x) + (1-\lambda)H(\tilde{x})$ is also closed and convex and it holds that $\lambda H(x) + (1-\lambda)H(\tilde{x}) \subseteq H(\lambda x + (1-\lambda)\tilde{x})$. Using the r-concavity of P this implies

$$g(\lambda x + (1-\lambda)\tilde{x}) \geq m_r(P(H(x)), P(H(\tilde{x})); \ \lambda) = m_r(g(x), g(\tilde{x}); \ \lambda).$$

□

Corollary 44. *Let $T(\xi) \equiv T$ and P be r-concave for some $r \in (-\infty, \infty]$. Let $X^*(P)$ be nonempty and $\mathcal{U} \subset \mathbb{R}^m$ be an open bounded neighbourhood of $X^*(P)$. Assume that there exists an element $\hat{x} \in X$ such that $P(H(\hat{x})) > p$ holds.*

Then there are constants $L > 0$, $\delta > 0$ and $k \in \mathbb{N}$ such that

$$|\vartheta(P) - \vartheta_{\mathcal{U}}(Q)| \leq L\alpha_{\mathrm{ph}_k}(P, Q)$$
$$\emptyset \neq X^*_{\mathcal{U}}(Q) \subseteq X^*(P) + \Psi_P(L\alpha_{\mathrm{ph}_k}(P, Q))\mathbb{B},$$

and $X^*_{\mathcal{U}}(Q)$ is a CLM set for (3.44) relative to $\mathcal{U}$ whenever $Q \in \mathcal{P}(\Xi)$ and $\alpha_{\mathrm{ph}_k}(P, Q) < \delta$. Here, the function Ψ_P is given by (2.23).

Proof. We assume without loss of generality that $r < 0$. Again we have to verify the metric regularity condition (ii) of Theorem 39. To this end, we use the function $\hat{g}(\cdot) := p^r - g^r(\cdot)$ instead of $g(\cdot) := P(H(\cdot))$. Since P is r-concave, the function $\hat{g}(\cdot)$ is concave on $\mathbb{R}^m$. We consider the set-valued mapping $\Gamma(x) := \{v \in \mathbb{R} : x \in X, \hat{g}(x) \geq v\}$ from $\mathbb{R}^m$ to $\mathbb{R}$. Its graph is closed and convex. Let $\overline{x} \in X$ with $g(\overline{x}) = p$, i.e., $\hat{g}(\overline{x}) = p^r$. As there exists an $\hat{x} \in X$ such that $g(\hat{x}) > p$, i.e., $\hat{g}(\hat{x}) > 0$, the element $v = 0$ belongs to the interior of the range of Γ. Hence, the Robinson–Ursescu Theorem (Theorem 9.48 in Rockafellar and Wets, 1998) implies the existence of constants $a > 0$ and $\varepsilon > 0$ such that

$$d(x, \Gamma^{-1}(v)) \leq a d(v, \Gamma(x)) \leq a \max\{0, v - \hat{g}(x)\}$$

holds whenever $x \in X$, $\|x - \overline{x}\| \leq \varepsilon$ and $|v| \leq \varepsilon$. For $x \in X$ with $\|x - \overline{x}\| \leq \varepsilon$ and sufficiently small $|y|$ we obtain

$$d(x, \mathcal{X}_y(P)) = d(x, \Gamma^{-1}(p^r - (p-y)^r)) \leq a \max\{0, g^r(x) - (p-y)^r\}$$

Finally, it remains to use that the function $v \mapsto v^r$ is locally Lipschitz continuous on $(0, +\infty)$. □

The above result improves in case of $h(\xi) \equiv \xi$ and, hence, $g(x) = F_P(Tx)$, where F_P is the distribution function of P. Then the polyhedral discrepancy α_{ph_k} can be replaced by the Kolmogorov distance d_{K}.

The next result provides a sufficient condition for (ii) in situations where P is not quasiconcave, but has a density on $\mathbb{R}^s$. Here, metric regularity is implied by a growth condition of $g(\cdot) = F_P(T\cdot)$ (see Henrion and Römisch, 1999).

Corollary 45. *Let $T(\xi) \equiv T$, $h(\xi) \equiv \xi$, $P \in \mathcal{P}(\mathbb{R}^s)$ have a density f_P, $X^*(P)$ be nonempty and $\mathcal{U} \subseteq \mathbb{R}^m$ be an open bounded neighbourhood of $X^*(P)$. Assume the following two conditions for each $\overline{x} \in X^*(P)$:*

(i) $(T\overline{x} + \operatorname{bd} \mathbb{R}^s_-) \cap \{\xi \in \mathbb{R}^s : \exists \varepsilon > 0 \text{ such that } f_P(\eta) \geq \varepsilon, \forall \eta \in \xi + \varepsilon \mathbb{B}\} \neq \emptyset$,
(ii) *there exists an $\hat{x} \in X$ such that $T\hat{x} > T\overline{x}$ holds componentwise.*

Then there are constants $L > 0$ and $\delta > 0$ such that

$$|\vartheta(P) - \vartheta_{\mathcal{U}}(Q)| \leq L\, d_K(P, Q)$$
$$\emptyset \neq X^*_{\mathcal{U}}(Q) \subseteq X^*(P) + \Psi_P(L\, d_{\mathrm{K}}(P, Q))\mathbb{B},$$

and $X^*_{\mathcal{U}}(Q)$ is a CLM set of (3.44) relative to $\mathcal{U}$ whenever $Q \in \mathcal{P}(\Xi)$ and $d_{\mathrm{K}}(P,Q) < \delta$. Here, the function Ψ_P is given by (2.23).

The essential condition (i) says that, for each $\chi \in T(X^*(P))$, the boundary of the cell $\chi + \mathbb{R}^s_-$ meets the strict positivity region of the density of P somewhere. This implies a suitable growth behaviour of the distribution function F_P at elements of $T(X^*(P))$, and hence metric regularity.

Finally, we study the growth function ψ_P of (3.44) and derive conditions implying quadratic growth near solution sets in case of $h(\xi) \equiv \xi$ and a logarithmic concave measure P. The first step of our analysis consists in a reduction argument that decomposes problem (3.44) into two auxiliary problems. The first one is a stochastic program with modified objective and probabilistic constraints (with decisions taken in $\mathbb{R}^s$) whereas the second one represents a parametric linear program. The argument is similar to Lemma 28 for two-stage models and was proved in Henrion and Römisch (1999).

Lemma 46. *Let $Q \in \mathcal{P}(\mathbb{R}^s)$ and $\mathcal{U} \subseteq \mathbb{R}^m$ be a nonempty open set such that its closure is a polytope. Then we have*

$$\vartheta_{\mathcal{U}}(Q) = \inf\{\pi_{\mathcal{U}}(y) : y \in T(X_{\mathcal{U}}),\, F_Q(y) \geq p\} \quad \text{and} \quad X^*_{\mathcal{U}}(Q) = \sigma_{\mathcal{U}}(Y_{\mathcal{U}}(Q)),$$

where

$$X_{\mathcal{U}} = X \cap \operatorname{cl} \mathcal{U},$$
$$Y_{\mathcal{U}}(Q) = \arg\min\{\pi_{\mathcal{U}}(y) : y \in T(X_{\mathcal{U}}),\, F_Q(y) \geq p\},$$
$$\pi_{\mathcal{U}}(y) = \inf\{\langle c, x\rangle : Tx = y,\, x \in X_U\},$$
$$\sigma_{\mathcal{U}}(y) = \arg\min\left\{\langle c, x\rangle : Tx = y,\, x \in X_{\mathcal{U}}\right\} \quad (y \in T(X_{\mathcal{U}})).$$

Here, $\pi_{\mathcal{U}}$ is convex polyhedral on $T(X_{\mathcal{U}})$ and $\sigma_{\mathcal{U}}$ is Lipschitz continuous on $T(X_{\mathcal{U}})$ with respect to the Pompeiu–Hausdorff distance on $\mathbb{R}^s$.

Theorem 47. *Let $T(\xi) \equiv T$, $h(\xi) \equiv \xi$, $P \in \mathcal{P}(\mathbb{R}^s)$ be logarithmic concave and $X^*(P)$ be nonempty and bounded. Assume that*

(i) $X^*(P) \cap \arg\min\{\langle c,x\rangle : x \in X\} = \emptyset$;
(ii) *there exists an $\overline{x} \in X$ such that $F_P(T\overline{x}) > p$;*
(iii) *$\log F_P$ is strongly concave on some convex neighbourhood $\mathcal{V}$ of $T(X^*(P))$.*

Then there exist $L > 0$ and $\delta > 0$ and a neighbourhood $\mathcal{U}$ of $X^*(P)$ such that

$$\mathbb{D}_\infty(X^*(P), X^*_{\mathcal{U}}(Q)) \leq L d_{\mathrm{K}}(P, Q)^{1/2}$$

holds whenever $Q \in \mathcal{P}(\mathbb{R}^s)$ and $d_{\mathrm{K}}(P,Q) < \delta$. Here, $\mathbb{D}_\infty$ denotes the Pompeiu–Hausdorff distance on subsets of $\mathbb{R}^m$ and d_{K} the Kolmogorov metric on $\mathcal{P}(\mathbb{R}^s)$.

Proof. Let $\mathcal{U}_0 \subseteq \mathbb{R}^m$ be an open convex set such that $X^*(P) \subseteq \mathcal{U}_0$ and $T(\mathcal{U}_0) \subseteq \mathcal{V}$. For each $x \in X^*(P)$ select $\varepsilon(x) > 0$ such that the polyhedron $x + \varepsilon(x)\mathbb{B}_\infty$ (with $\mathbb{B}_\infty$ denoting the closed unit ball w.r.t. the norm $\|\cdot\|_\infty$ on $\mathbb{R}^m$) is contained in $\mathcal{U}_0$. Since $X^*(P)$ is compact, finitly many of these balls cover $X^*(P)$. The closed convex hull $\overline{\mathcal{U}}$ of their union is a polyhedron with $X^*(P) \subseteq \mathcal{U} \subset \overline{\mathcal{U}} \subseteq \mathcal{U}_0$, where $\mathcal{U} = \operatorname{int} \overline{\mathcal{U}}$. With the notations of Lemma 46 we consider the problem

$$\min\{\pi_{\mathcal{U}}(y) : y \in T(X_{\mathcal{U}}),\ \hat{g}(y) := \log p - \log F_P(y) \leq 0\}.$$

According to Lemma 46 the solution set $Y_{\mathcal{U}}(P)$ of this problem fulfils $X^*(P) = X^*_{\mathcal{U}}(P) = \sigma_{\mathcal{U}}(Y_{\mathcal{U}}(P))$. Let $y_* \in Y_{\mathcal{U}}(P)$ and $\overline{y} = T\overline{x}$ with $\overline{x} \in X$ from (ii). Then the logarithmic concavity of P implies for any $\lambda \in (0,1]$:

$$\begin{aligned}\hat{g}(\lambda\overline{y} + (1-\lambda)y_*) &= \log p - \log F_P(\lambda\overline{y} + (1-\lambda)y_*) \\ &\leq \log p - \lambda \log F_P(\overline{y}) - (1-\lambda)\log F_P(y_*) \\ &\leq \lambda(\log p - \log F_P(\overline{y})) < 0.\end{aligned}$$

Thus, we may choose $\hat{\lambda} \in (0,1]$ such that $\hat{y} = \hat{\lambda}\overline{y} + (1-\hat{\lambda})y_*$ belongs to $T(X_{\mathcal{U}})$ and has the property $\hat{g}(\hat{y}) < 0$. This constraint qualification implies the existence of a Kuhn–Tucker coefficient $\lambda_* \geq 0$ such that

$$\pi_{\mathcal{U}}(y_*) = \min\{\pi_{\mathcal{U}}(y) + \lambda_* \hat{g}(y) : y \in T(X_{\mathcal{U}})\} \quad \text{and} \quad \lambda_* \hat{g}(y_*) = 0.$$

In case $\lambda_* = 0$, this would imply $y_* \in \arg\min\{\pi_{\mathcal{U}}(y): y \in T(X_{\mathcal{U}})\}$ and, hence, the existence of some $x_* \in X^*(P)$ with $\langle c, x_*\rangle = \pi_{\mathcal{U}}(Tx_*) = \min\{\langle c, x\rangle : Tx = y_*, x \in X_{\mathcal{U}}\}$. Hence, condition (i) would be violated due to $x^* \in \operatorname{int}\mathcal{U}$. Thus $\lambda_* > 0$ and $\pi_V + \lambda_*\hat{g}$ is strongly convex on $T(X_{\mathcal{U}})$. Hence, y_* is the unique minimizer of $\pi_V + \lambda_*\hat{g}$ and the growth property

$$\rho\|y - y_*\|^2 \le \pi_{\mathcal{U}}(y) + \lambda_*\hat{g}(y) - \pi_{\mathcal{U}}(y^*) \tag{3.49}$$

holds for some $\rho > 0$ and all $y \in T(X_{\mathcal{U}})$.

As the assumptions of Corollary 44 are satisfied, the set-valued mapping $X^*_{\mathcal{U}}(\cdot)$ is upper semicontinuous at P and $X^*_{\mathcal{U}}(Q) \neq \emptyset$ is a complete local minimizing set if $d_{\mathrm{K}}(P,Q)$ is sufficiently small. Hence, there exists a $\delta > 0$ such that $\emptyset \neq X^*_{\mathcal{U}}(Q) \subseteq \mathcal{U}$ for all $Q \in \mathcal{P}(\mathbb{R}^s)$ with $d_{\mathrm{K}}(P.Q) < \delta$. With the notations from Lemma 46 and using the fact that $Y_U(P) = \{y_*\}$ and $X^*(P) = X^*_{\mathcal{U}}(P) = \sigma_{\mathcal{U}}(y_*)$ we obtain

$$\mathbb{D}_\infty(X^*(P), X^*_{\mathcal{U}}(Q)) = \mathbb{D}_\infty(\sigma_{\mathcal{U}}(y_*), \sigma_{\mathcal{U}}(Y_{\mathcal{U}}(Q))) \le \hat{L} \sup_{y \in Y_{\mathcal{U}}(Q)} \|y - y_*\|,$$

where $\hat{L} > 0$ is the Lipschitz constant of σ_U (cf. Lemma 46). Using (3.49) and $Y_{\mathcal{U}}(Q) \subseteq T(X_{\mathcal{U}})$, the above chain of inequalities extends to

$$\begin{aligned}
\mathbb{D}_\infty(X^*(P), X^*_{\mathcal{U}}(Q)) &\le \frac{\hat{L}}{\rho^{1/2}} \sup_{y \in Y_{\mathcal{U}}(Q)} [\pi_{\mathcal{U}}(y) + \lambda_*\hat{g}(y) - \pi_{\mathcal{U}}(y_*)]^{1/2} \\
&= \frac{\hat{L}}{\rho^{1/2}} [\vartheta_{\mathcal{U}}(Q) - \vartheta(P) + \lambda_*(\log p - \log F_P(y))]^{1/2} \\
&\le \frac{\hat{L}}{\rho^{1/2}} [\vartheta_{\mathcal{U}}(Q) - \vartheta(P) + \lambda_*(\log F_Q(y) - \log F_P(y))]^{1/2} \\
&\le \frac{\hat{L}}{\rho^{1/2}} \left[\left(L + \frac{\lambda_*}{p}\right) d_{\mathrm{K}}(P, Q)\right]^{1/2},
\end{aligned}$$

where $L > 0$ is the constant from Theorem 39 and $\frac{1}{p}$ the Lipschitz constant of $\log(\cdot)$ on $[p, 1]$. This completes the proof. □

A slightly more general version of the result for r-concave measures was proved in Henrion and Römisch (1999). The assumptions (i)–(iii) imposed in Theorem 47 concern the original problem. The conditions (i) and (ii) mean that the probability level p is not chosen too low and too high, respectively. Condition (i) expresses the fact that the presence of the probabilistic constraint $F_P(Tx) \ge p$ moves the solution set $X^*(P)$ away from the one obtained without imposing that constraint. Recent results in Henrion and Römisch (2002) show that assumption (i) is not necessary for Theorem 47 to hold. Assumption (iii) is decisive for the desired growth condition of the

objective function around $X^*(P)$. In contrast to the global concavity of $\log F_P$, (iii) requires the strong concavity of $\log F_P$ as a local property around $T(X^*(P))$. Since general sufficient criteria for (iii) are not available so far, we provide a few examples.

Example 48. (strong logarithmic concavity of measures) Let P be the uniform distribution on some bounded rectangle in $\mathbb{R}^s$ having the form $D = \times_{i=1}^s [a_i,b_i]$. Then $\log F_P(\xi) = \sum_{i=1}^s \log(\xi_i - a_i), \xi \in D$. Clearly, $\log(\cdot - a_i)$ is strongly concave on any closed subinterval of (a_i,b_i). Hence, $\log F_P(\cdot)$ is strongly concave on any closed convex subset of $\operatorname{int} D$.

Let P be the multivariate normal distribution on $\mathbb{R}^s$ having a nonsingular diagonal covariance matrix. A direct computation for the standard normal distribution function Φ on $\mathbb{R}$ shows that $\log\Phi$ is strongly concave on any bounded interval. Since $\log F_P$ is equal to the sum of logarithms of the marginal distribution functions, it is strongly concave on any bounded convex set in $\mathbb{R}^s$.

4 Approximations of stochastic programs

Many approximations of stochastic programs result from replacing the underlying probability distribution by some other measure, which typically leads to simpler models. Important examples are nonparametric statistical estimates (e.g., empirical ones) and scenario tree constructions using probability distribution information. Next we give an idea how the results of the previous sections may be used to design and to analyse approximations of stochastic programs. We begin with some glimpses into the analysis of empirical approximations and the relations to empirical process theory. A more far-reaching analysis is given in Pflug (2003) and Shapiro (2003).

4.1 A glimpse of empirical approximations

Let $P \in \mathcal{P}(\Xi)$ and $\xi_1,\xi_2,\ldots,\xi_n,\ldots$ be independent identically distributed Ξ-valued random variables on a probability space $(\Omega,\mathcal{A},\mathbb{P})$ having the common distribution P, i.e., $P = \mathbb{P}\xi_1^{-1}$. We consider the *empirical measures*

$$P_n(\omega) := \frac{1}{n}\sum_{i=1}^n \delta_{\xi_i(\omega)} \qquad (\omega \in \Omega;\ n \in \mathbb{N}),$$

where δ_ξ denotes the unit mass at $\xi \in \Xi$, and the empirical approximations of the stochastic program (1.1), i.e., the models that result from replacing P by $P_n(\cdot)$. These take the form

$$\min\left\{\sum_{i=1}^n F_0(x,\xi_i(\cdot)) \colon x \in X, \sum_{i=1}^n F_j(x,\xi_i(\cdot)) \le 0\ , j = 1,\ldots,d\right\}, \qquad (4.50)$$

where the factor $\frac{1}{n}$ in the objective and constraints has been removed. Since the objective and constraint functions F_j, $j = 0, \ldots, d$, are assumed to be random lower semicontinuous functions from $\mathbb{R}^m \times \Xi$ to $\overline{\mathbb{R}}$, the constraint set is closed-valued and measurable from Ω to $\mathbb{R}^m$ and, hence, the optimal value $\vartheta(P_n(\cdot))$ of (4.50) is measurable from Ω to $\overline{\mathbb{R}}$ and the solution set $X^*(P_n(\cdot))$ is a closed-valued measurable multifunction from Ω to $\mathbb{R}^m$ (see Chapter 14 and, in particular, Theorem 14.37 in Rockafellar and Wets (1998)). The same conclusion is valid for the localized concepts ϑ_U and X_U^* for any nonempty open subset $\mathcal{U}$ of $\mathbb{R}^m$.

Another measurability question arises when studying uniform convergence properties of the *empirical process*

$$\left\{n^{\frac{1}{2}}(P_n(\cdot) - P)F = n^{-\frac{1}{2}}\sum_{i=1}^{n}(F(\xi_i(\cdot)) - PF)\right\}_{F\in\mathcal{F}},$$

indexed by some class $\mathcal{F}$ of functions that are integrable with respect to P. Here, we set $QF := \int_\Xi F(\xi)\,\mathrm{d}Q(\xi)$ for any $Q \in \mathcal{P}(\Xi)$ and $F \in \mathcal{F}$. Since the suprema $d_{\mathcal{F}}(P_n(\cdot),P) = \sup_{F\in\mathcal{F}}|P_n(\cdot)F - PF|$ may be non-measurable functions from Ω to $\overline{\mathbb{R}}$, we introduce a condition on $\mathcal{F}$ that simplifies matters and is satisfied in most stochastic programming models. A class $\mathcal{F}$ of measurable functions from Ξ to $\overline{\mathbb{R}}$ is called *P-permissible* for some $P \in \mathcal{P}(\Xi)$ if there exists a countable subset $\mathcal{F}_0$ of $\mathcal{F}$ such that for each function $F \in \mathcal{F}$ there exists a sequence (F_n) in $\mathcal{F}_0$ converging pointwise to F and such that the sequence (PF_n) also converges to PF. Then

$$d_{\mathcal{F}}(P_n(\omega), P) = \sup_{F\in\mathcal{F}}|(P_n(\omega) - P)F| = d_{\mathcal{F}_0}(P_n(\omega), P)$$

holds for each $n \in \mathbb{N}$ and $\omega \in \Omega$, i.e., the analysis is reduced to a countable class and, in particular, $d_{\mathcal{F}}(P_n(\cdot), P)$ is a measurable function from Ω to $\overline{\mathbb{R}}$.

A P-permissible class $\mathcal{F}$ is called a *P-Glivenko–Cantelli class* if the sequence $(d_{\mathcal{F}}(P_n(\cdot), P))$ of random variables converges to 0 $\mathbb{P}$-almost surely. If $\mathcal{F}$ is P-permissible, the empirical process $\{n^{\frac{1}{2}}(P_n(\cdot) - P)F\}_{F\in\mathcal{F}}$ is called *uniformly bounded in probability with tail* $C_{\mathcal{F}}(\cdot)$ if the function $C_{\mathcal{F}}(\cdot)$ is defined on $(0,\infty)$ and decreasing to 0, and the estimate

$$\mathbb{P}(\{\omega : n^{\frac{1}{2}}d_{\mathcal{F}}(P_n(\omega), P) \geq \varepsilon\}) \leq C_{\mathcal{F}}(\varepsilon) \tag{4.51}$$

holds for each $\varepsilon > 0$ and $n \in \mathbb{N}$. Whether a given class $\mathcal{F}$ is a P-Glivenko–Cantelli class or the empirical process is uniformly bounded in probability depends on the size of the class $\mathcal{F}$ measured in terms of certain *covering numbers* or the corresponding *metric entropy numbers* defined as their logarithms (e.g., Dudley (1984), Pollard (1990), van der Vaart and Wellner (1996)).

To introduce these concepts, let $\mathcal{F}$ be a subset of the normed space $L_r(\Xi, P)$ for some $r \geq 1$ equipped with the usual norm $\|F\|_{P,r} = (P|F|^r)^{\frac{1}{r}}$. The covering number $N(\varepsilon, \mathcal{F}, L_r(\Xi, P))$ is the minimal number of open balls $\{G \in L_r(\Xi, P): \|G - F\|_{P,r} < \varepsilon\}$ needed to cover $\mathcal{F}$. A measurable function $F_{\mathcal{F}}$ from Ξ to $\overline{\mathbb{R}}$ is called an *envelope* of the class $\mathcal{F}$ if $|F(\xi)| \leq F_{\mathcal{F}}(\xi)$ holds for every $\xi \in \Xi$ and $F \in \mathcal{F}$. The following result provides criteria for the desired properties in terms of uniform covering numbers.

Theorem 49. *Let $\mathcal{F}$ be P-permissible with envelope $F_{\mathcal{F}}$. If $PF_{\mathcal{F}} < \infty$ and*

$$\sup_Q N(\varepsilon \|F_{\mathcal{F}}\|_{Q,1}, \mathcal{F}, L_1(Q)) < \infty, \tag{4.52}$$

then $\mathcal{F}$ is a P-Glivenko–Cantelli class. If $\mathcal{F}$ is uniformly bounded and there exist constants $r \geq 1$ and $R \geq 1$ such that

$$\sup_Q N(\varepsilon \|F_{\mathcal{F}}\|_{Q,2}, \mathcal{F}, L_2(Q)) \leq \left(\frac{R}{\varepsilon}\right)^r \tag{4.53}$$

holds for all $\varepsilon > 0$, then the empirical process indexed by $\mathcal{F}$ is uniformly bounded in probability with exponential tail $C_{\mathcal{F}}(\varepsilon) = (K(R)\varepsilon r^{-\frac{1}{2}})^r \exp(-2\varepsilon^2)$ with some constant $K(R)$ depending only on R.

The suprema in (4.52) and (4.53) are taken over all finitely discrete probability measures Q with $\|F_{\mathcal{F}}\|_{Q,1} = QF_{\mathcal{F}} > 0$ and $\|F_{\mathcal{F}}\|_{Q,2}^2 = QF_{\mathcal{F}}^2 > 0$, respectively.

For the proof we refer to Talagrand (1994), van der Vaart and Wellner (1996) and van der Vaart (1998). For studying entropic sizes of stochastic programs Pflug (1999, 2003) uses results of this type but with *bracketing numbers* instead of uniform covering numbers. He also studies situations where $\mathcal{F}$ is not uniformly bounded and shows that he *blow-up function* $n^{\frac{1}{2}}$ for $n \to \infty$ has to be replaced by some function converging to ∞ more slowly. Here, we use the concept of uniform covering numbers since they turn out to be useful for discontinuous functions.

The stability results of Section 2 directly translate into convergence results and rates, respectively, for empirical optimal values and solution sets.

Theorem 50. *Assume that the conditions (i)–(iii) of Theorem 5 are satisfied and that $\mathcal{F}_{\mathcal{U}}$ is P-permissible.*

If $\mathcal{F}_{\mathcal{U}}$ is a P-Glivenko–Cantelli class, the sequences

$$(|\vartheta(P) - \vartheta_{\mathcal{U}}(P_n(\cdot))|) \quad \text{and} \quad \left(\sup_{x \in X_{\mathcal{U}}^*(P_n(\cdot))} d(x, X^*(P))\right)$$

converge $\mathbb{P}$-almost surely to 0. *Furthermore, the set $X^*_{\mathcal{U}}(P_n(\omega))$ is a CLM set of* (4.50) *relative to $\mathcal{U}$ for sufficiently large $n \in \mathbb{N}$ and for $\mathbb{P}$-almost all $\omega \in \Omega$.*

If the empirical process indexed by $\mathcal{F}_{\mathcal{U}}$ is uniformly bounded in probability with tail $C_{\mathcal{F}_{\mathcal{U}}}(\cdot)$, the following estimates hold for each $\varepsilon > 0$ and each $n \in \mathbb{N}$:

$$\mathbb{P}(|\vartheta(P) - \vartheta_{\mathcal{U}}(P_n(\cdot))| > \varepsilon n^{-\frac{1}{2}}) \leq C_{\mathcal{F}_{\mathcal{U}}}\left(\min\left\{\delta, \frac{\varepsilon}{L}\right\}\right), \tag{4.54}$$

$$\mathbb{P}\left(\sup_{x \in X^*_{\mathcal{U}}(P_n(\cdot))} d(x, X^*(P)) > \varepsilon n^{-\frac{1}{2}}\right) \leq C_{\mathcal{F}_{\mathcal{U}}}(\min\{\delta, \hat{L}^{-1}\Psi_P^{-1}(\varepsilon)\}). \tag{4.55}$$

Proof. Let $L > 0$, $\hat{L} > 0$, $\delta > 0$ be the constants in Theorems 5 and 9. First, let $\mathcal{F}_{\mathcal{U}}$ be a P-Glivenko–Cantelli class and $A \in \mathcal{A}$ be such that $\mathbb{P}(A) = 0$ and $(d_{\mathcal{F}_{\mathcal{U}}}(P_n(\omega),P))$ converges to 0 for every $\omega \in \Omega\backslash A$. Let $\omega \in \Omega\backslash A$. Then $X^*_{\mathcal{U}}(P_n(\omega))$ is nonempty, since the objective function $\int_{\Xi} F_0(\cdot,\xi)\,\mathrm{d}P(\xi)$ is lower semicontinuous on X and the constraint set $X_{\mathcal{U}}(P_n(\omega))$ is compact due to Proposition 3. Let $n_0(\omega) \in \mathbb{N}$ be such that $d_{\mathcal{F}_{\mathcal{U}}}(P_n(\omega),P) < \delta$ holds for each $n \geq n_0(\omega)$. Due to the Theorems 5 and 9 the estimates

$$\begin{aligned} &|\vartheta(P) - \vartheta_{\mathcal{U}}(P_n(\omega))| \leq d_{\mathcal{F}_{\mathcal{U}}}(P_n(\omega),P) \\ &\sup_{x \in X^*_{\mathcal{U}}(P_n(\omega))} d(x,X^*(P)) \leq \Psi_P(\hat{L} d_{\mathcal{F}_{\mathcal{U}}}(P_n(\omega),P)) \end{aligned}$$

hold for $n \geq n_0(\omega)$. In particular, the sequences $(|\vartheta(P) - \vartheta_{\mathcal{U}}(P_n(\omega))|)$ and $(\sup_{x \in X^*_{\mathcal{U}}(P_n(\omega))} d(x, X^*(P)))$ converge to 0. Hence, $X^*_{\mathcal{U}}(P_n(\omega)) \subseteq \mathcal{U}$ and, thus, $X^*_{\mathcal{U}}(P_n(\omega))$ is a CLM set relative to $\mathcal{U}$ for sufficiently large $n \in \mathbb{N}$.

Now, let $\varepsilon > 0$ be arbitrary. The Theorems 5 and 9 also imply

$$\mathbb{P}(|\vartheta(P) - \vartheta_{\mathcal{U}}(P_n(\cdot))| > \varepsilon) \leq \mathbb{P}\left(d_{\mathcal{F}_{\mathcal{U}}}(P_n(\cdot), P) \geq \min\left\{\delta, \frac{\varepsilon}{L}\right\}\right), \tag{4.56}$$

$$\mathbb{P}\left(\sup_{x \in X^*_{\mathcal{U}}(P_n(\cdot))} d(x, X^*(P)) > \varepsilon\right) \leq \mathbb{P}(d_{\mathcal{F}_{\mathcal{U}}}(P_n(\cdot), P) \geq \min\{\delta, \hat{L}^{-1}\Psi_P^{-1}(\varepsilon)\}). \tag{4.57}$$

If the empirical process indexed by $\mathcal{F}_{\mathcal{U}}$ is uniformly bounded in probability with tail $C_{\mathcal{F}_{\mathcal{U}}}(\cdot)$, the estimates (4.56) and (4.57) may be continued by using (4.51) and, thus, lead to (4.54) and (4.55). □

The estimates (4.54) and (4.55) may be used to derive the speed of convergence in probability of optimal values and solution sets, respectively. Clearly, the speed depends on the asymptotic behaviour of the tail $C_{\mathcal{F}_{\mathcal{U}}}(\varepsilon)$ as

$\varepsilon \to \infty$ and of the function Ψ_P. For the situation of exponential tails, this is elaborated in Rachev and Römisch (2002).

Next we show how our analysis applies to two-stage stochastic programs with and without integrality requirements and to chance constrained models. It turns out that, under reasonable assumptions on all models, the empirical process indexed by $\mathcal{F}_{\mathcal{U}}$ is uniformly bounded in probability with exponential tails.

Example 51. (linear chance constrained models) A class $\mathcal{B}$ of Borel sets of $\mathbb{R}^s$ is called a *Vapnik-Červonenkis* (*VC*) *class* of index $r = r(\mathcal{B})$ if r is finite and equal to the smallest $n \in \mathbb{N}$ for which no set of cardinality $n+1$ is shattered by $\mathcal{B}$. $\mathcal{B}$ is said to shatter a subset $\{\xi_1, \ldots, \xi_l\}$ of cardinality l in $\mathbb{R}^s$ if each of its 2^l subsets is of the form $B \cap \{\xi_1, \ldots, \xi_l\}$ for some $B \in \mathcal{B}$. For VC classes $\mathcal{B}$ it holds that

$$N(\varepsilon, \{\chi_B : B \in \mathcal{B}\}, L_1(\Xi, Q)) \leq K\varepsilon^{-r}$$

for any $\varepsilon > 0$ and $Q \in \mathcal{P}(\Xi)$, and some constant $K > 0$ depending only on the index r (Theorem 2.6.4 in van der Vaart and Wellner (1996)).

For any polyhedral set $\Xi \subseteq \mathbb{R}^s$ and $k \in \mathbb{N}$ the class $\mathcal{B}_{\mathrm{ph}_k}(\Xi)$ is a VC class, since the class of all closed half spaces is VC and finite intersections of VC classes are again VC. The corresponding class of characteristic functions is permissible for P, since the set of all polyhedra in $\mathcal{B}_{\mathrm{ph}_k}(\Xi)$ having vertices at rational points in $\mathbb{R}^s$ plays the role of the countable subset in the definition of permissibility. Hence, Theorem 49 applies and the empirical process indexed by $\mathcal{F}_{\mathcal{U}} = \{\chi_{H(x)} : x \in X \cap \mathrm{cl}\,\mathcal{U}\}$, where $\mathcal{U}$ is a bounded open set containing $X^*(P)$, is uniformly bounded in probability with exponential tail $C_{\mathcal{F}_{\mathcal{U}}}(\varepsilon) = \hat{K}\varepsilon^r \exp(-2\varepsilon^2)$ for some index $r \in \mathbb{N}$ and some constant $\hat{K} > 0$. For example, from Theorem 50 we obtain for each $\varepsilon > 0$ and $n \in N$ the estimate

$$\mathbb{P}\left(\sup_{x \in X^*_{\mathcal{U}}(P_n(\cdot))} d(x, X^*(P)) > \varepsilon n^{-\frac{1}{2}} \right) \leq \hat{K}\varepsilon^r \exp(-2\min\{\delta, \hat{L}^{-1}\Psi_P^{-1}(\varepsilon)\}^2).$$

Example 52. (two-stage models without integrality) Let F_0 be defined as in Section 3.1 and let (A1) and (A2) be satisfied. Then, for each nonempty open and bounded subset $\mathcal{U}$ of $\mathbb{R}^m$, the class $\mathcal{F}_{\mathcal{U}} = \{F_0(x, \cdot) : x \in X \cap \mathrm{cl}\,\mathcal{U}\}$ is a subset of $L_1(\Xi, P)$. $\mathcal{F}_{\mathcal{U}}$ is also permissible for P, since any class $\{F_0(x, \cdot) : x \in X_c\}$ with X_c being a countable and dense subset of $X \cap \mathrm{cl}\,\mathcal{U}$ may be used as the countable subset of $\mathcal{F}_{\mathcal{U}}$ in the definition of permissibility. Proposition 22 implies that the function $F_{\mathcal{F}_{\mathcal{U}}}(\xi) := K\max\{1, \|\xi\|^2\}$ $(\xi \in \Xi)$ is an envelope of $\mathcal{F}_{\mathcal{U}}$ for sufficiently large $K > 0$. Furthermore, due to the Lipschitz continuity property of $F_0(\cdot, \xi)$ with Lipschitz constant $\hat{L}\max\{1, \|\xi\|^2\}$ (see Proposition 22), the uniform covering numbers of $\mathcal{F}_{\mathcal{U}}$ are bounded by the

covering numbers of $X \cap \operatorname{cl} U$ (see Theorem 2.7.11 in van der Vaart and Wellner (1996)). In particular, for each finitely discrete measure $Q \in \mathcal{P}(\Xi)$ and with $\hat{F}(\xi) := L\max\{1, \|\xi\|^2\}$ $(\xi \in \Xi)$ it holds that

$$N(\varepsilon \|\hat{F}\|_{Q,r}, \mathcal{F}_{\mathcal{U}}, L_r(\Xi, Q)) \leq N(\varepsilon, X \cap \operatorname{cl} \mathcal{U}, \mathbb{R}^m) \leq K\varepsilon^{-m}, \tag{4.58}$$

for each $\varepsilon > 0$, $r \geq 1$ and some constant $K > 0$ depending only on m and the diameter of $X \cap \operatorname{cl}\mathcal{U}$. Using (4.58) for $r = 1$, Theorem 49 implies that $\mathcal{F}_{\mathcal{U}}$ is a P-Glivenko-Cantelli class. If Ξ is bounded, $\mathcal{F}_{\mathcal{U}}$ is uniformly bounded and, using (4.58) for $r = 2$, Theorem 49 implies that the empirical process indexed by $\mathcal{F}_{\mathcal{U}}$ is uniformly bounded in probability with exponential tail.

Example 53. (mixed-integer two-stage models) Let F_0 be defined as in Section 3.2 and let (B1)–(B3) be satisfied and Ξ be bounded. Then, for each nonempty open and bounded subset $\mathcal{U}$ of $\mathbb{R}^m$, the class

$$F_{\mathcal{U}} = \left\{ F_0(x,\cdot) = \sum_{j=1}^{\nu} (\langle c, x\rangle + \Phi(h(\cdot) - T(\cdot)x)\chi_{\Xi^R_{j,x}}(\cdot) : x \in X \cap \operatorname{cl} \mathcal{U} \right\}$$

is a subset of $L_1(\Xi, P)$. This representation follows from Proposition 34 if $R > 0$ is chosen sufficiently large such that $\{\xi \in \Xi : \|h(\xi) - T(\xi)x\|_\infty > R\} = \emptyset$ for each $x \in X \cap \operatorname{cl}\mathcal{U}$. For each $X \cap \operatorname{cl}\mathcal{U}$ the sets $\Xi^R_{j,x}$ $(j = 1, \ldots, \nu)$ form a disjoint partition of Ξ into Borel sets whose closures are in $\mathcal{B}_{\mathrm{ph}_k}(\Xi)$ for some $k \in \mathbb{N}$. Furthermore, the function $\Phi(h(\cdot) - T(\cdot)x)$ is Lipschitz continuous on each of these sets with a uniform constant $L_1 > 0$. Let $F_0^j(x,\cdot)$ denote a Lipschitz extension of the function $\langle c, x\rangle + \Phi(h(\cdot) - T(\cdot)x)$ from $\Xi^R_{j,x}$ to $\mathbb{R}$ by preserving the Lipschitz constant L_1 $(j = 1, \ldots, \nu)$. Furthermore, let $\mathcal{F}^j_{\mathcal{U}} := \{F_0^j(x,\cdot) : x \in X \cap \operatorname{cl}\mathcal{U}\}$ and $\mathcal{G}^j_{\mathcal{U}} := \{\chi_{\Xi^R_{j,x}} : x \in X \cap \operatorname{cl}\mathcal{U}\}$ $(j = 1, \ldots, \nu)$.

Now, we use a permanence property of the uniform covering numbers (cf. Section 2.10.3 in van der Vaart and Wellner (1996)). Let $Q \in \mathcal{P}(\Xi)$ be discrete with finite support. Then the estimate

$$N(\varepsilon C_0, \mathcal{F}_{\mathcal{U}}, L_2(\Xi, Q)) \leq \prod_{j=1}^{\nu} N(\varepsilon C_j, \mathcal{F}^j_{\mathcal{U}}, L_2(\Xi, Q_j)) N(\varepsilon \hat{C}_j, \mathcal{G}^j_{\mathcal{U}}, L_2(\Xi, \hat{Q}_j)) \tag{4.59}$$

is valid, where C_0, $C_j > 1$, $\hat{C}_j$, $j = 1, \ldots, \nu$, are certain constants and Q_j, $\hat{Q}_j$, $j = 1, \ldots, \nu$, certain discrete measures having finite support. The constants depend on the bounds of the uniformly bounded classes $\mathcal{F}^j_{\mathcal{U}}$ and $\mathcal{G}^j_{\mathcal{U}}$, $j = 1, \ldots, \nu$. Since the latter classes satisfy the condition (4.53) (see Example 51 and Example 52), the estimate (4.59) implies that $\mathcal{F}_{\mathcal{U}}$ satisfies (4.53), too. Hence, we obtain the same estimates for mixed-integer two-stage models as in Example 52 for two-stage models without integrality requirements and in Example 51 for linear chance constrained models.

Example 54. (newsboy continued) According to Example 15, the class $\mathcal{F}_{\mathcal{U}}$ is of the form $\mathcal{F}_{\mathcal{U}} = \{F_0(x,\cdot) = (r-c)x + c\max\{0, x - \cdot\}: x \in X \cap \mathrm{cl}\,\mathcal{U}\}$ with envelope $F_{\mathcal{F}_{\mathcal{U}}}(\xi) = r \sup_{X \in X_n \mathrm{cl}\,\mathcal{U}}|x| + c|\xi|$ and a uniform Lipschitz constant c. Hence, $\mathcal{F}_{\mathcal{U}}$ is a subset of $L_1(\Xi,P)$ if $\int_{\Xi} |\xi| \mathrm{d}P(\xi) = \sum_{k \in N} \pi_k k < \infty$. As in Example 52 we obtain

$$N(\varepsilon c, \mathcal{F}_{\mathcal{U}}, L_2(\Xi,Q)) \leq N(\varepsilon, X \cap \mathrm{cl}\,\mathcal{U}, \mathbb{R}^m) \leq C\varepsilon^{-m}$$

for each finitely discrete measure $Q \in \mathcal{P}(\Xi)$ and, hence, Theorem 50 provides the rate of convergence of the solution sets $X^*_{\mathcal{U}}(P_n(\cdot))$ of (1.4) with linear Ψ_P.

4.2 Scenario generation and reduction

Most of the numerical solution approaches for stochastic programs resort to discrete approximations of the underlying probability measure P. Several approaches have been developed for the generation or construction of discrete approximations and are in use for solving applied stochastic programming models (see the overview by Dupačová et al. (2000) and the references therein). The quantitative stability results of Section 2.3 suggest another approach, namely, to construct approximations for the original measure P such that they are close to P with respect to the corresponding probability (pseudo) metric. Let $\mathcal{F}$ be a set of measurable functions from Ξ to $\mathbb{R}$ such that the stochastic programming model (1.1) is stable in the sense of the Theorems 5 and 9 with respect to the (pseudo) metric

$$d_{\mathcal{F}}(P,Q) = \sup_{F \in \mathcal{F}} \left| \int_{\Xi} F(\xi) d(P - Q)(\xi) \right|$$

or some other distance bounding $d_{\mathcal{F}}(P,Q)$ from above. This means that the optimal values and the solution sets of (1.1) behave continuously at P when perturbing P with respect to $d_{\mathcal{F}}$.

Then it is a natural requirement to construct approximate probability distributions such that they are best approximations to P in the sense of $d_{\mathcal{F}}$. For instance, the principle of *optimal scenario generation* with a prescribed number of scenarios may be formulated as follows:

Given $P \in \mathcal{P}(\Xi)$ and $M \in \mathbb{N}$, determine a discrete probability measure $Q^* \in \mathcal{P}(\Xi)$ having M scenarios such that

$$d_{\mathcal{F}}(P, Q^*) = \min \left\{ d_{\mathcal{F}}\left(P, \sum_{j=1}^{M} q_j \delta_{\xi_j}\right) : \sum_{j=1}^{M} q_j = 1, q_j \geq 0, \xi_j \in \Xi, j = 1, \ldots, M \right\}.$$

Further constraints could be incorporated into the minimization problem, e.g., constraints implying that the scenarios exhibit a tree structure. Unfortunately, it seems to be hopeless to solve this problem for general measures P, function classes $\mathcal{F}$, supports Ξ, and large numbers M of scenarios. However, it is of course a challenging problem to develop approaches for solving the best approximation problem for more specific situations, like e.g. for the unconstrained case $\Xi = \mathbb{R}^s$, discrete measures P (involving very many scenarios) and function classes that are relevant in Section 3. An approach for solving the best approximation problem in case of $\Xi = \mathbb{R}^s$ and $\mathcal{F} = \mathcal{F}_1(\mathbb{R}^s)$ is developed in Pflug (2001).

Another important problem consists in reducing a given discrete probability measure $P = \sum_{i=1}^N p_i \delta_{\xi_i}$ with a (very) large number N of scenarios to a measure containing n of the original scenarios with $n << N$. Similarly as in case of optimal scenario generation, the problem of *optimal scenario reduction* may be formulated in the form

$$\min\left\{ d_{\mathcal{F}}\Big(\sum_{i=1}^N p_i\delta_{\xi_i}, \sum_{j\in J} q_j \delta_{\xi_j}\Big) : J \subset \{1,\ldots,N\}, |J| = n, \sum_{j\in J} q_j = 1, q_j \geq 0 \right\}, \quad (4.60)$$

i.e., as a nonlinear mixed-integer program. Since its objective function is difficult to compute for general classes $\mathcal{F}$, solution methods for (4.60) are a challenging task. However, in the special case that $\mathcal{F} = \mathcal{F}_p(\Xi)$, for some $p \geq 1$, the objective function of (4.60) turns out to be the dual optimal value of the standard network flow problem (see Rachev and Rüschendorf (1998))

$$\min\left\{ \sum_{\substack{i=1\\ j\in J}}^N c_p(\xi_i,\xi_j)\|\xi_i - \xi_j\|\eta_{ij} : \eta_{ij} \geq 0, \sum_{i=1}^N \eta_{ij} - \sum_{j\in J} \eta_{ij} = q_j - p_i, \forall i,j \right\}$$

where $c_p(\xi_i,\xi_j) = \max\{1, \|\xi_i\|\xi_j\|\}^{p-1}$, $i,j = 1,\ldots,N$, and, hence, it is a polyhedral function of q. Furthermore, in case of $\mathcal{F} = \mathcal{F}_1(\Xi)$ problem (4.60) simplifies considerably.

Proposition 55. *Given $J \subset \{1,\ldots,N\}$ we have*

$$\min\left\{ d_{\mathcal{F}_1(\Xi)}\left(\sum_{i=1}^N p_i\delta_{\xi_i}, \sum_{j\in J} q_j\delta_{\xi_j}\right) : \sum_{j\in J} q_j = 1, q_j \geq 0 \right\} = \sum_{i\notin J} p_i \min_{j\in J}\|\xi_i - \xi_j\|.$$

Moreover, the minimum is attained at $\overline{q}_j = p_j + \sum_{i \in J_j} p_i$, *for each* $j \in J$, *where* $J_j := \{i \notin J : j = j(i)\}$ *and* $j(i) \in \arg\min_{j \in J} \|\xi_i - \xi_j\|$ *for each* $i \notin J$.

The proposition provides an explicit formula for the *redistribution* of the given probabilities p_i, $i = 1, \ldots, N$, to the scenarios with indices in J. For its proof we refer to Theorem 2 in Dupačová et al. (2003). Due to Proposition 55 the optimal scenario reduction problem (4.60) in case of $\mathcal{F} = \mathcal{F}_1(\Xi)$ takes the form: Given $P \in \mathcal{P}(\Xi)$ and $n \in \mathbb{N}$, determine a solution of

$$\min \left\{ \sum_{i \notin J} p_i \min_{j \in J} \|\xi_i - \xi_j\| : J \subset \{1, \ldots, N\}, |J| = n \right\} \tag{4.61}$$

and compute the optimal weights $\overline{q}$ according to the redistribution rule in Proposition 55. Notice that problem (4.61) means that the set $\{1, \ldots, N\}$ has to be covered by a subset J of $\{1, \ldots, N\}$ and by $\{1, \ldots, N\} \backslash J$ such that $|J| = n$ and the cover has minimal cost $\sum_{i \notin J} p_i \min_{j \in J} \|\xi_i - \xi_j\|$. Hence, problem (4.61) is of set-covering type and, thus, $\mathcal{NP}$-hard. However, the specific structure of the objective function allows the design of fast heuristic algorithms for its approximate solution (see Dupačová et al. (2003), Heitsch and Römisch (2003)). Depending on the size of the number n of remaining scenarios, the two basic ideas are *backward reduction* and *forward selection*, respectively. In the backward reduction heuristic an index set $J = \{l_1, \ldots, l_n\}$ is determined such that

$$l_i \in \arg\min_{l \in J_r^{[i-1]}} \sum_{k \in J_r^{[i-1]} \backslash \{l\}} p_k \min_{j \notin J_r^{[i-1]} \backslash \{l\}} \|\xi_k - \xi_j\| \quad (i = 1, \ldots, n),$$

where $J_r^{[0]} = \{1, \ldots, N\}$, $J_r^{[i]} = J_r^{[i-1]} \backslash \{l_i\}$, $i = 1, \ldots, n$. In the forward selection heuristic the index set $J = \{l_1, \ldots, l_n\}$ is chosen by an opposite strategy such that

$$l_i \in \arg\min_{l \notin J_s^{[i-1]}} \sum_{k \notin J_s^{[i-1]} \cup \{l\}} p_k \min_{j \in J_s^{[i-1]} \cup \{l\}} \|\xi_k - \xi_j\| \quad (i = 1, \ldots, n)$$

holds, where $J_s^{[0]} = \emptyset$, $J_s^{[i]} = J_s^{[i-1]} \cup \{l_i\}$, $i = 1, \ldots, n$. We refer to Heitsch and Römisch (2003) for a discussion of the complexity of both heuristics, for implementation issues and encouraging numerical results.

5 Bibliographical notes

The beginnings of approximation and estimation results in stochastic programming date back to the 1970s and the papers by Kall (1974) (see also the monograph Kall (1976)), Marti (1975, 1979) and Olsen (1976) on approximations, and the work of Kaňková (1977) and Wets (1979) on empirical estimation in stochastic programming. Surveys on stability were published by Dupačová (1990) and Schultz (2000). The notion of stability of stochastic programs appeared first in Bereanu (1975), in the context of the distribution problem, and in Kaňková (1978), where stability of minima of more general stochastic programming models was studied with respect to weak convergence of measures for the first time.

Later Dupačová (1984, 1987) and Wang (1985) studied the stability of stochastic programs with respect to changes of finite-dimensional parameters in the underlying probability distribution. Kall and Stoyan (1982), Salinetti (1983) and Römisch (1981, 1985) dealt with discrete approximations to stochastic programs. Further early work has been done in the surveys by Wets (1983, 1989) and in Friedrich and Tammer (1981) (on stability), Birge and Wets (1986) (on discrete approximation schemes), Römisch (1986b), Kall (1987), Robinson and Wets (1987), Römisch and Wakolbinger (1987), Vogel (1988), Kall, Ruszczyński and Frauendorfer (1988) (on discrete approximations), Dupačová and Wets (1988), Shapiro (1989) and Wang (1989). The landmark papers by Kall (1987) and by Robinson and Wets (1987) address qualitative stability results for optimal values and solution sets with respect to weak convergence of measures. This line of research was continued in the important work by Artstein and Wets (1994) and in Vogel (1992), Schultz (1992, 1995), Lucchetti and Wets (1993), Wang (1995) and Wets (1998), Zervos (1999) and Riis and Schultz (2002). Attempts to quantify such stability results using distances of probability measures were started in Römisch (1986b), Römisch and Wakolbinger (1987) and continued in Römisch and Schultz (1991a,b,c, 1993, 1996) and Artstein (1994), Kanková (1994b), Kanková (1998), Shapiro (1994), Fiedler and Römisch (1995), Schultz (1996), Henrion and Römisch (1999, 2000), Dentcheva (2000) and Rachev and Römisch (2002).

Most of the stability studies allow for general perturbations of the underlying probability measure and develop a general framework for both *discrete* and *statistical* approximations of stochastic programs. Nevertheless, these two kinds of approximations developed independently by exploiting their specific structures (e.g., bounding techniques on the one hand and asymptotic statistical arguments on the other hand). For (discrete) approximations we mention the work in Birge and Wets (1986), Kall et al. (1988), Lepp (1990), Birge and Qi (1995a,b), Frauendorfer (1992, 1996) and Kall (1998).

In parallel, statistical inference in stochastic programming models was studied intensively. After the early work by Kanková and Wets, many authors

contributed to this line of research on asymptotic properties of statistical estimators, e.g., their consistency, rates of convergence and limit theorems. We mention, in particular, the work of Dupačová and Wets (1988), Vogel (1988), Shapiro (1989, 1990, 1991, 1996, 2000), King (1989) and Kanková (1990, 1994), King and Wets (1991), Wets (1991), Ermoliev and Norkin (1991), Norkin (1992), King and Rockafellar (1993), Rubinstein and Shapiro (1993), Bouza (1994), Geyer (1994), Artstein and Wets (1995), Kaniovski et al. (1995), Lachout (1995), Pflug (1995, 1999), Robinson (1996), Gröwe (1997), Pflug et al. (1998a,b), Mak et al. (1999) and Shapiro and Homem-de-Mello (2000).

Another line of research on approximations of stochastic programs is based on the convergence (almost surely, in probability and in distribution) of measurable set-valued mappings and on the epi-convergence of integrands. Here, we mention the fundamental paper by Salinetti and Wets (1986) and the work of Salinetti (1981, 1983), Römisch (1986a), Vogel (1988, 1992, 1994, 1995) and Wets (1991), Hess (1996) and the recent papers by Korf and Wets (2000, 2001) and by Vogel and Lachout (2000).

Much is known on the stability of values and solutions of classical two-stage stochastic programs (Section 3.1). The situation is already different for the stability of solutions to chance constrained models and even more to mixed-integer two-stage models. The stability of multi-stage stochastic programs is widely open, especially in the mixed-integer case. Another open matter are the stability effects of incorporating risk functionals into stochastic programming models (cf. Section 2.4).The paper by Rachev and Römisch (2002) provides an important source for the material presented in this chapter, in particular, for the Sections 2.2, 2.3, and 4.1 and parts of the Sections 3.1 and 3.2. Some of the results are directly taken from that paper, namely, Theorems 5, 9, 23 and 39. Some other results represent modified or extended versions of those in Rachev and Römisch (2002) (e.g. Theorem 35 and Theorem 50). Theorems 13 and 24 are due to work in preparation by Römisch and Wets. Corollary 45 and Theorem 47 are taken from Henrion and Römisch (1999) and the Corollaries 42 and 44 from Römisch and Schultz (1991c). The Example 41 is due to Henrion and the notion of a Lipschitz continuous risk functional goes back to Pflug (2002).

Acknowledgements

This work was supported by the Deutsche Forschungsgemeinschaft, in particular, by the Schwerpunktprogramm *Online Optimization of Large Scale Systems*. The author wishes to thank Darinka Dentcheva (Stevens Institute of Technology) and René Henrion (Weierstrass Institute Berlin) for invaluable discussions, in particular, on Steiner selections of set-valued maps and on the stability of chance constrained models, respectively. Further thanks are due to

Georg Pflug (University of Vienna) and Roger Wets (University of California at Davis) for valuable conversations on empirical processes and risk functionals and on the stability of stochastic programs, respectively.

References

Artstein, Z. (1994). Sensitivity with respect to the underlying information in stochastic programs. *Journal of Computational and Applied Mathematics* 56, 127–136.

Artstein, Z., R.J.-B. Wets (1994). Stability results for stochastic programs and sensors, allowing for discontinuous objective functions. *SIAM Journal on Optimization* 4, 537–550.

Artstein, Z., R.J.-B. Wets (1995). Consistency of minimizers and the SLLN for stochastic programs. *Journal of Convex Analysis* 2, 1–17.

Artzner, P., F. Delbaen, J.-M. Eber, D. Heath (1999). Coherent measures of risk. *Mathematical Finance* 9, 203–228.

Attouch, H., R.J.-B. Wets (1993). Quantitative stability of variational systems II. A framework for nonlinear conditioning. *SIAM Journal on Optimization* 3, 359–381.

Bank, B., J. Guddat, D. Klatte, B. Kummer, K. Tammer (1982). *Non-Linear Parametric Optimization*, Akademie-Verlag, Berlin.

Bereanu, B. (1975). Stable stochastic linear programs and applications. *Mathematische Operationsforschung und Statistik* 6, 593–607.

Berge, C. (1963). *Topological spaces*, Oliver & Boyd, Edinburgh-London.

Billingsley, P. (1968). *Convergence of Probability Measures*, Wiley, New York.

Billingsley, P., F. Topsøe (1967). Uniformity in weak convergence. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* 7, 1–16.

Birge, J.R., L. Qi (1995a). Subdifferential convergence in stochastic programs. *SIAM Journal on Optimization* 5, 436–453.

Birge, J.R., L. Qi (1995b). Continuous approximation schemes for stochastic programs. *Annals of Operations Research* 56, 15–38.

Birge, J.R., R.J.-B Wets (1986). Designing approximation schemes for stochastic optimization problems, in particular for stochastic programs with recourse. *Mathematical Programming Study* 27, 54–102.

Blair, C.E., R.G. Jeroslow (1977). The value function of a mixed integer program. *Discrete Mathematics* 19, 121–138.

Bonnans, J.F., A. Shapiro (2000). *Perturbation Analysis of Optimization Problems*, Springer-Verlag, New York.

Bouza, C. (1994). Approximation of the value of the optimal solution in stochastic programming. *Revista Investigacion Operacional* 15, 13–33.

Delbaen, F. (2002). Coherent risk measures on general probability spaces, in: K. Sandmann, P.J. Schönbucher (eds.), *Advances in Finance and Stochastics*, Springer, Berlin, pp. 1–37.

Dentcheva, D. (2000). Regular Castaing representations of multifunctions with applications to stochastic programming. *SIAM Journal on Optimization* 10, 732–749.

Dentcheva, D., W. Römisch (2000). Differential stability of two-stage stochastic programs. *SIAM Journal on Optimization* 11, 87–112.

Dudley, R.M. (1984). A course on empirical processes, in: École d'Été de Probabilities de Saint-Flour XII – 1982, *Lecture Notes in Mathematics* 1097, Springer, Berlin, pp. 2–142.

Dudley, R.M. (1989). *Real Analysis and Probability*, Wadsworth & Brooks/Cole, Pacific Grove.

Dupačová, J. (1984). Stability in stochastic programming with recourse - estimated parameters. *Mathematical Programming* 28, 72–83.

Dupačová, J. (1987). Stochastic programming with incomplete information: A survey of results on postoptimization and sensitivity analysis. *Optimization* 18, 507–532.

Dupačová, J. (1990). Stability and sensitivity analysis for stochastic programming. *Annals of Operations Research* 27, 115–142.

Dupačová, J. (1994). Applications of stochastic programming under incomplete information. *Journal of Computational and Applied Mathematics* 56, 113–125.

Dupačová, J., G. Consigli, S.W. Wallace (2000). Scenarios for multistage stochastic programs. *Annals of Operations Research* 100, 25–53.

Dupačová, J., N. Gröwe-Kuska, W. Römisch (2003). Scenario reduction in stochastic programming: An approach using probability metrics. *Mathematical Programming*, *Ser. A* 95, 493–511.

Dupačová, J., R.J.-B. Wets (1988). Asymptotic behaviour of statistical estimators and of optimal solutions of stochastic optimization problems. *The Annals of Statistics* 16, 1517–1549.

Ermoliev, Y.M., V.I. Norkin (1991). Normalized convergence in stochastic optimization. *Annals of Operations Research* 30, 187–198.

Fiedler, O., W. Römisch (1995). Stability in multistage stochastic programming. *Annals of Operations Research* 56, 79–93.

Föllmer, H., A. Schied (2002). Convex risk measures and trading constraints. *Finance and Stochastics* 6, 429–447.

Fortet, R., E. Mourier (1953). Convergence de la répartition empirique vers la répartition théorique. *Ann. Sci. Ecole Norm. Sup.* 70, 266–285.

Frauendorfer, K. (1992 *Stochastic Two-Stage Programming, Lecture Notes in Economics and Mathematical Systems*, Vol. 392, Springer, Berlin.

Frauendorfer, K. (1996). Barycentric scenario trees in convex multistage stochastic programming. *Mathematical Programming* 75, 277–293.

Friedrich, B., K. Tammer (1981). Untersuchungen zur Stabilität einiger Ersatzprobleme der stochastischen Optimierung in bezug auf Änderungen des zugrunde gelegten Zufallsvektors, Wissenschaftliche Zeitschrift der Humboldt-Universität Berlin. *Math.-Naturwiss. Reihe* 30, 373–376.

Geyer, C.J. (1994). On the asymptotics of constrained M-estimation. *The Annals of Statistics* 22, 1993–2010.

Gröwe, N. (1997). Estimated stochastic programs with chance constraints. *European Journal of Operational Research* 101, 285–305.

Heitsch, H., W. Römisch (2003). Scenario reduction algorithms in stochastic programming. *Computational Optimization and Applications* 24, 187–206.

Henrion, R. (2000). Qualitative stability of convex programs with probabilistic constraints, in: V.H. Nguyen, J.-J. Strodiot, P. Tossings (eds.), *Optimization, Lecture Notes in Economics and Mathematical Systems*, Vol. 481, Springer, Berlin, pp. 164–180.

Henrion, R., W. Römisch (1999). Metric regularity and quantitative stability in stochastic programs with probabilistic constraints. *Mathematical Programming* 84, 55–88.

Henrion, R., W. Römisch (2000). Stability of solutions to chance constrained stochastic programs, in: J. Guddat, R. Hirabayashi, H.Th. Jongen, F. Twilt (eds.), *Parametric Optimization and Related Topics V,* Peter Lang Verlag, Frankfurt a.M., pp. 95–114.

Henrion, R., W. Römisch (2002). Hölder and Lipschitz stability of solution sets in programs with probabilistic constraints, Preprint No. 798, Weierstrass Institute for Applied Analysis and Stochastics Berlin, and submitted.

Hess, C. (1996). Epi-convergence of sequences of normal integrands and strong consistency of the maximum likelihood estimator. *The Annals of Statistics* 24, 1298–1315.

Kall, P. (1974). Approximations to stochastic programs with complete recourse. *Numerische Mathematik* 22, 333–339.

Kall, P. (1976). *Stochastic Linear Programming*, Springer, Berlin.

Kall, P. (1987). On approximations and stability in stochastic programming, in: J. Guddat, H.Th. Jongen, B. Kummer, F. Nozicka (eds.), *Parametric Optimization and Related Topics,* Akademie-Verlag, Berlin, pp. 387–407.

Kall, P., A. Ruszczyński, K. Frauendorfer (1988). Approximation techniques in stochastic programming, in: Y. Ermoliev, R.J.-B. Wets (eds.), *Numerical Techniques for Stochastic Optimization,* Springer, Berlin, pp. 33–64.

Kall, P. (1998). Bounds for and approximations to stochastic linear programs with recourse, in: K. Marti, P. Kall (eds.), *Stochastic Programming Methods and Technical Applications, Lecture Notes in Economics and Mathematical Systems*, Vol. 458, Springer, Berlin, pp. 1–21.

Kaniovski, Y.M., A.J. King, R.J.-B. Wets (1995). Probabilistic bounds (via large deviations) for the solutions of stochastic programming problems. *Annals of Operations Research* 56, 189–208.

Kanková, V. (1977). Optimum solution of a stochastic optimization problem with unknown parameters, in: Transactions of the 7th Prague Conference on *Information Theory, Statistical Decision Functions and Random Processes*, Academia, Prague, pp. 239–244.

Kanková, V. (1978). Stability in the stochastic programming. *Kybernetika* 14, 339–349.

Kanková, V. (1990). On the convergence rate of empirical estimates in chance constrained stochastic programming. *Kybernetika* 26, 449–461.

Kanková, V. (1994). A note on estimates in stochastic programming. *Journal of Computational and Applied Mathematics* 56, 97–112.

Kanková, V. (1994). On distribution sensitivity in stochastic programming, Research Report No. 1826, Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic.

Kaňková, V. (1998). A note on multifunctions in stochastic programming, in: K. Marti, P. Kall (eds.), *Stochastic Programming Methods and Technical Applications, Lecture Notes in Economics and Mathematical Systems*, Vol. 458, Springer, Berlin, pp. 154–168.

King, A.J. (1989). Generalized delta theorems for multivalued mappings and measurable selections. *Mathematics of Operations Research* 14, 720–736.

King, A.J., R.T. Rockafellar (1993). Asymptotic theory for solutions in statistical estimation and stochastic programming. *Mathematics of Operations Research* 18, 148–162.

King, A.J., R.J.-B. Wets (1991). Epi-consistency of convex stochastic programs. *Stochastics and Stochastics Reports* 34, 83–92.

Klatte, D. (1987). A note on quantitative stability results in nonlinear optimization, in: K. Lommatzsch (ed.), *Proceedings 19. Jahrestagung Mathematische Optimierung*, Humboldt-Universität Berlin, Sektion Mathematik. *Seminarbericht Nr*. 90, 77–86.

Klatte, D. (1994). On quantitative stability for non-isolated minima. *Control and Cybernetics* 23, 183–200.

Korf, L., R.J.-B. Wets (2000). Random lsc functions: scalarization, *Stochastic Programming E-Print Series* 7.

Korf, L., R.J-B. Wets (2001). Random lsc functions: An ergodic theorem. *Mathematics of Operations Research* 26, 421–445.

Lachout, P. (1995). On multifunction transforms of probability measures. *Annals of Operations Research* 56, 241–249.

Lepp, R. (1990). Approximations to stochastic programs with complete recourse. *SIAM Journal on Control and Optimization* 28, 382–394.

Louveaux, F., R. Schultz (2003). *Stochastic Integer Programming*, chapter 4, this volume.

Lucchetti, R., R.J.-B. Wets (1993). Convergence of minima of integral functionals, with applications to optimal control and stochastic optimization. *Statistics & Decisions* 11, 69–84.

Lucchetti, R., G. Salinetti, R.J.-B. Wets (1994). Uniform convergence of probability measures: Topological criteria. *Journal of Multivariate Analysis* 51, 252–264.

Mak, W.-K., D. Morton, R. Wood (1999). Monte Carlo bounding techniques for determining solution quality in stochastic programs. *Operations Research Letters* 24, 47–56.

Marti, K. (1975). Approximationen der Entscheidungsprobleme mit linearer Ergebnisfunktion und positiv homogener, subadditiver Verlustfunktion. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* 31, 203–233.

Marti, K. (1979). *Approximationen stochastischer Optimierungsprobleme*, Mathematical Systems in Economics Vol. 43, A. Hain, Königstein/Ts.

Mordukhovich, B.S. (1994a). Generalized differential calculus for nonsmooth and set-valued mappings. *Journal of Mathematical Analysis and Applications* 183, 250–288.

Mordukhovich, B.S. (1994b). Lipschitzian stability of constraint systems and generalized equations. *Nonlinear Analysis, Theory, Methods & Applications* 22, 173–206.

Norkin, V.I. (1992). Convergence of the empirical mean method in statistics and stochastic programming. *Cybernetics and Systems Analysis* 28, 253–264.
Ogryczak, W., A. Ruszczyński (1999). From stochastic dominance to mean-risk models: semideviation as risk measures. *European Journal of Operational Research* 116, 33–50.
Olsen, P. (1976). Discretizations of multistage stochastic programming problems. *Mathematical Programming Study* 6, 111–124.
Pflug, G. (1995). Asymptotic stochastic programs. *Mathematics of Operations Research* 20, 769–789.
Pflug, G. (1999). Stochastic programs and statistical data. *Annals of Operations Research* 85, 59–78.
Pflug, G. (2001). Scenario tree generation for multiperiod financial optimization by optimal discretization. *Mathematical Programming Series B* 89, 251–271.
Pflug, G. (2003). *Stochastic Optimization and Statistical Inference*, chapter 7, this volume.
Pflug, G., A. Ruszczynski, R. Schultz (1998a). On the Glivenko-Cantelli problem in stochastic programming: Linear recourse and extensions. *Mathematics of Operations Research* 23, 204–220.
Pflug, G., A. Ruszczynski, R. Schultz (1998b). On the Glivenko-Cantelli problem in stochastic programming: Mixed-integer linear recourse. *Mathematical Methods of Operations Research* 47, 39–49.
Pollard, D. (1990). *Empirical Processes: Theory and Applications*, NSF-CBMS Regional Conference Series in Probability and Statistics Vol. 2, Institute of Mathematical Statistics.
Prékopa, A. (1995). *Stochastic Programming*, Kluwer, Dordrecht.
Rachev, S.T. (1991). *Probability Metrics and the Stability of Stochastic Models*, Wiley, Chichester.
Rachev, S.T., W. Römisch (2002). Quantitative stability in stochastic programming: The method of probability metrics, *Mathematics of Operations Research*, 27, 792–818.
Rachev, S.T., L. Rüschendorf (1998). *Mass Transportation Problems*, *Vol. I, Theory*, Springer, New York.
Riis, M., R. Schultz (2002). Applying the minimum risk criterion in stochastic recourse programs. *Computational Optimization and Applications* (to appear).
Robinson, S.M. (1987). Local epi-continuity and local optimization. *Mathematical Programming* 37, 208–223.
Robinson, S.M. (1996). Analysis of sample-path optimization. *Mathematics of Operations Research* 21, 513–528.
Robinson, S.M., R.J-B Wets (1987). Stability in two-stage stochastic programming. *SIAM Journal on Control and Optimization* 25, 1409–1416.
Rockafellar, R.T., R.J.-B. Wets (1998). *Variational Analysis*, Springer, Berlin.
Rockafellar, R.T., S. Uryasev (2002). Conditional value-at-risk for general loss distributions. *Journal of Banking & Finance* 26, 1443–1471.
Römisch, W. (1981). On discrete approximations in stochastic programming, in: K. Lommatzsch (ed.), *Proceedings 13. Jahrestagung Mathematische Optimierung*, Humboldt-Universität Berlin, Sektion Mathematik. *Seminarbericht Nr.* 39, 166–175.
Römisch, W. (1985). An approximation method in stochastic optimization and control *Mathematical Control Theory*, Vol. 14, Banach Center Publications, PWN, Warsaw, 477–490.
Römisch, W. (1986a). On the convergence of measurable selections and an application to approximations in stochastic optimization. *Zeitschrift für Analysis und Anwendungen* 5, 277–288.
Römisch, W. (1986b). On convergence rates of approximations in stochastic programming, in: K. Lommatzsch (ed.), *Proceedings 17. Jahrestagung Mathematische Optimierung*, Humboldt-Universität Berlin, Sektion Mathematik. *Seminarbericht Nr.* 80, 82–91.
Römisch, W., R. Schultz (1991a). Distribution sensitivity in stochastic programming. *Mathematical Programming* 50, 197–226.
Römisch, W., R. Schultz (1991b). Distribution sensitivity for certain classes of chance-constrained models with application to power dispatch. *Journal of Optimization Theory and Applications* 71, 569–588.
Römisch, W., R. Schultz (1991c). Stability analysis for stochastic programs. *Annals of Operations Research* 30, 241–266.

Römisch, W., R. Schultz (1993). Stability of solutions for stochastic programs with complete recourse. *Mathematics of Operations Research* 18, 590–609.

Römisch, W., R. Schultz (1996). Lipschitz stability for stochastic programs with complete recourse. *SIAM Journal on Optimization* 6, 531–547.

Römisch, W., A. Wakolbinger (1987). Obtaining convergence rates for approximations in stochastic programming, in: Guddat, J., Jongen, H.T., Kummer, B., Nozicka, F. (eds.), *Parametric Optimization and Related Topics,* Akademie-Verlag, Berlin, pp. 327–343.

Rubinstein, R.Y., A. Shapiro (1993). *Discrete Event Systems, Sensitivity Analysis and Stochastic Optimization by the Score Function Method*, Wiley, Chichester.

Ruszczynski, A., A. Shapiro (2003). *Stochastic Programming Models*, chapter 1, this volume.

Salinetti, G. (1981). Convergence of measurable multifunctions: An application to stochastic optimization problems, in: Castellani, G., Mazzoleni, P. (eds.), *Mathematical Programming and its Economic Applications,* Angeli, Milano, pp. 757–774.

Salinetti, G. (1983). Approximations for chance-constrained programming problems. *Stochastics* 10, 157–179.

Salinetti, G., R.J.-B. Wets (1986). On the convergence in distribution of measurable multifunctions (random sets), normal integrands, stochastic processes and stochastic infima. *Mathematics of Operations Research* 11, 385–419.

Schultz, R. (1992). Continuity and stability in two-stage stochastic integer programming, in: K. Marti (ed.), *Stochastic Optimization, Numerical Methods and Technical Applications, Lecture Notes in Economics and Mathematical Systems*, Vol. 379, Springer, Berlin, pp. 81–92.

Schultz, R. (1994). Strong convexity in stochastic programs with complete recourse. *Journal of Computational and Applied Mathematics* 56, 3–22.

Schultz, R. (1995). On structure and stability in stochastic programs with random technology matrix and complete integer recourse. *Mathematical Programming* 70, 73–89.

Schultz, R. (1996). Rates of convergence in stochastic programs with complete integer recourse. *SIAM Journal on Optimization* 6, 1138–1152.

Schultz, R. (2000). Some aspects of stability in stochastic programming. *Annals of Operations Research* 100, 55–84.

Shapiro, A. (1989). Asymptotic properties of statistical estimators in stochastic programming. *The Annals of Statistics* 17, 841–858.

Shapiro, A. (1990). On differential stability in stochastic programming. *Mathematical Programming* 47, 107–116.

Shapiro, A. (1991). Asymptotic analysis of stochastic programs. *Annals of Operations Research* 30, 169–186.

Shapiro, A. (1994). Quantitative stability in stochastic programming. *Mathematical Programming* 67, 99–108.

Shapiro, A. (1996). Simulation-based optimization – Concergence analysis and statistical inference. *Communications in Statistics – Stochastic Models* 12, 425–454.

Shapiro, A. (2000). Statistical inference of stochastic optimization problems: *Probabilistic Constrained Optimization*, Uryasev, S.P. (ed.), Kluwer, Dordrecht, pp. 282–307.

Shapiro, A. (2003). *Monte Carlo Sampling Methods*, chapter 6, this volume.

Shapiro, A., T. Homem-de-Mello (2000). On rate of convergence of optimal solutions of Monte Carlo approximations of stochastic programs. *SIAM Journal on Optimization* 11, 70–86.

Talagrand, M. (1994). Sharper bounds for Gaussian and empirical processes. *Annals of Probability* 22, 28–76.

Talagrand, M. (1996). The Glivenko-Cantelli problem, ten years later. *Journal of Theoretical Probability* 9, 371–384.

Topsøe, F. (1967). On the connection between -continuity and -uniformity in weak convergence. *Probability Theory and its Applications* 12, 281–290.

Topsøe, F. (1977). Uniformity in convergence of measures. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* 39, 1–30.

van der Vaart, A.W. (1998). *Asymptotic Statistics*, Cambridge University Press.

van der Vaart, A.W., J.A. Wellner (1996). *Weak Convergence and Empirical Processes*, Springer, New York.

Vogel, S. (1988). Stability results for stochastic programming problems. *Optimization* 19, 269–288.

Vogel, S. (1992). On stability in multiobjective programming–A stochastic approach. *Mathematical Programming* 56, 91–119.

Vogel, S. (1994). A stochastic approach to stability in stochastic programming. *Journal of Computational and Applied Mathematics* 56, 65–96.

Vogel, S. (1995). On stability in stochastic programming–Sufficient conditions for continuous convergence and epi-convergence, Preprint No. M 12/95, Technical University Ilmenau, Faculty of Mathematics and Natural Sciences.

Vogel, S., P. Lachout (2000). On continuous convergence and epi-convergence of random functions – I and II, Research Report No. 1988 and 1989, Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic.

Walkup, D., R.J.-B. Wets (1969a). Lifting projections of convex polyhedra. *Pacific Journal of Mathematics* 28, 465–475.

Walkup, D., R.J.-B. Wets (1969b). A Lipschitzian characterization of convex polyhedra. *Proceedings of the American Mathematical Society* 23, 167–173.

Wang, J. (1985). Distribution sensitivity analysis for stochastic programs with complete recourse. *Mathematical Programming* 31, 286–297.

Wang, J. (1989). Continuity of feasible solution sets of probabilistic constrained programs. *Journal of Optimization Theory and Applications* 63, 79–89.

Wang, J. (1995). Stability of multistage stochastic programming. *Annals of Operations Research* 56, 313–322.

Wets, R.J-B (1974). Stochastic programs with fixed recourse: The equivalent deterministic program. *SIAM Review* 16, 309–339.

Wets, R.J.-B. (1979). A statistical approach to the solution of stochastic programs with (convex) simple recourse, Working Paper, Department of Mathematics, University of Kentucky, Lexington.

Wets, R.J.-B. (1983). Stochastic programming: Solution techniques and approximation schemes, in: Bachem, A., Grötschel, M., Korte, B. (eds.), *Mathematical Programming: the State-of-the-Art,* Springer, Berlin, pp. 566–603.

Wets, R.J.-B. (1989). Stochastic programming, in: G.L. Nemhauser, A.H.G. Rinnooy Kan, M.J. Todd (eds.), *Handbooks in Operations Research and Management Science, Vol. 1, Optimization*, North-Holland, Amsterdam, 573–629.

Wets, R.J.-B. (1991). Constrained estimation: Consistency and asymptotics. *Applied Stochastic Models and Data Analysis* 7, 17–32.

Wets, R.J.-B. (1998). Stochastic programs with chance constraints: Generalized convexity and approximation issues, in: Crouzeix, J.-P., Martínez-Legaz, J.-E., Volle, M. (eds.), *Generalized Convexity, Generalized Monotonicity: Recent Results*, Kluwer, Dordrecht, pp. 61–74.

Zervos, M. (1999). On the epiconvergence of stochastic optimization problems. *Mathematics of Operations Research* 24, 495–508.

Zolotarev, V.M. (1983). Probability metrics. *Theory of Probability and its Applications* 28, 278–302.

A. Ruszczyński and A. Shapiro, Eds., *Handbooks in OR & MS, Vol. 10*

Chapter 9

Stochastic Programming in Transportation and Logistics

Warren B. Powell and Huseyin Topaloglu

Department of Operations Research and Financial Engineering, Princeton University, Princeton, NJ 08544, USA

Abstract

Freight transportation is characterized by highly dynamic information processes: customers call in orders over time to move freight; the movement of freight over long distances is subject to random delays; equipment failures require last minute changes; and decisions are not always executed in the field according to plan. The high-dimensionality of the decisions involved has made transportation a natural application for the techniques of mathematical programming, but the challenge of modeling dynamic information processes has limited their success. In this chapter, we explore the use of concepts from stochastic programming in the context of resource allocation problems that arise in freight transportation. Since transportation problems are often quite large, we focus on the degree to which some techniques exploit the natural structure of these problems. Experimental work in the context of these applications is quite limited, so we highlight the techniques that appear to be the most promising.

1 Introduction

Operational models of problems in transportation and logistics offer a ripe set of applications for stochastic programming since they are typically characterized by highly dynamic information processes. In freight transportation, it is the norm to call a carrier the day before, or sometimes the same day, to request that a shipment be moved. In truckload trucking, last minute phone calls are combined with requests that can be made a few days in advance, putting carriers in the position of committing to move loads without knowing the last minute demands that will be made of them (sometimes by their most important customers). In railroads, requests to move freight might be made a week in the future, but it can take a week to move a freight car to a customer. The effect is the same.

The goal of this chapter is to provide some examples of problems, drawn from the arena of freight transportation, that appear to provide a natural application of stochastic programming. Optimization models in transportation and logistics, as they are applied in practice, are almost always formulated based on deterministic models. Our intent is to show where deterministic models can exhibit fundamental weaknesses, not from the perspective of academic theory, but in terms of practical limitations as perceived by people in industry. At the same time, we want to use the richness of real problems to raise issues that may not have been addressed by the stochastic programming community. We want to highlight what works, what does not, and where there are rich areas for new research.

We do not make any effort to provide a comprehensive treatment of stochastic optimization problems in transportation and logistics. First, we consider only problems in freight transportation (for the uninitiated, "transportation and logistics" refers to the operational problems surrounding the movement of goods). These problems are inherently discrete, giving rise to stochastic, integer programming problems, but we focus on problems where linear programming formulations represent a good starting point. We completely avoid the general area of stochastic vehicle routing or the types of batch processes that often arise in the movement of smaller shipments, and focus instead on problems that can be broadly described as dynamic resource allocation problems.

Our presentation begins in Section 2 with an overview of different classes of applications. This section provides a summary of the different types of uncertainty that arise, and addresses the fundamental question of why stochastic programming is a promising technology for freight transportation. Section 3 provides a general modeling framework that represents a bridge between linear programming formulations and a representation that more explicitly captures the dimensions of transportation applications. In Section 4 we present a case study based on the distribution of freight cars for a railroad. This case study provides us with a problem context where dynamic information processes play an important role. We use this case study in the remainder of the chapter to keep our discussions grounded in the context of a real application.

We approach the stochastic modeling of our freight car problem in two steps. First, we discuss in Section 5 the basic two-stage resource allocation problem. This problem is particularly relevant to the car distribution problem. The characteristics of the car distribution problem nicely illustrate different types of recourse strategies that can arise in practice. Specialized strategies give way to approximations which exploit the underlying network structure. For the most general case (network recourse) we briefly review a broad range of stochastic programming strategies, focusing on their ability to handle the structure of transportation problems.

Section 6 addresses multistage problems. Our approach toward multistage problems is that they can and should be solved as sequences of two-stage

problems. As a result, we solve multistage problems by building on the theory of two-stage problems.

Transportation problems offer far more richness than can be covered in a single chapter. Section 8 provides a hint of the topics that we do not attempt to cover. We close with Section 9 that discusses some of the challenges of actually implementing stochastic models in an operational setting.

2 Applications and issues

It is important to have in mind a set of real problems that arise in transportation and logistics. We begin our discussion of applications by listing some sample problems that arise in practice, and then use these problems: (a) to discuss sources of uncertainty; (b) to raise special modeling problems that arise in transportation applications; and finally (c) to highlight, from a practical perspective, the limitations of deterministic models and how stochastic programming can improve the quality of our models from a practical perspective.

2.1 Some sample problems

Transportation, fundamentally, is the business of moving things so that they are more useful. If there is a resource at a location i, it may be more useful at another location j. Within this simple framework, there is a tremendous variety of problems that pose special modeling and algorithmic issues. Below is a short list of problems that helps to highlight some of the modeling issues that we will have to grapple with.

1) Product distribution—Perhaps one of the oldest and most practical problems is the determination of how much product to ship from a plant to intermediate warehouses before finally shipping to the retailer (or customer). The decision of how much to ship and where must be made before we know the customer demand. There are a number of important variations of this problem, including:

   a) Separability of the distribution process—It is often the case that each customer will be served by a unique warehouse, but substitution among warehouses may be allowed.
   b) Multiple product types with substitution—A company may make multiple product types (for example, different types of salty food snacks) for a market that is willing to purchase different products when one is sold out. For the basic single period distribution problem, substitution between products at different locations is the same as substitution across different types of products, as long as the substitution cost is known (when the cost is a transportation cost,

this is known, whereas when it represents the cost of substituting for different product types, it is usually unknown).

c) Demand backlogging—In multiperiod problems, if demand is not satisfied in one time period, we may assume the demand is lost or backlogged to the next time period. We might add that the same issue arises in the product being managed; highly perishable products vanish if not used at a point in time, whereas nonperishable products stay around.

2) Container management—Often referred to as fleet management in the literature, "containers" represent boxes of various forms that hold freight. These might be trailers, boxcars, or the intermodal containers that are used to move goods across the oceans (and then by truck and rail to inland customers). Containers represent a reusable resource where the act of satisfying a customer demand (moving freight from i to j) also has the effect of changing the state of the system (the container is moved from i and j). The customer demand vanishes from the system, but the container does not. Important problem variations include:

   a) Single commodity problems—These arise when all the containers are the same, or when there are different container types with no substitution between different types of demands. When there is no substitution, the problem decomposes into a series of single commodity problems for each product type.
   b) Multicommodity problems—There may be different container types, and the customers may be willing to substitute between them. For example, they may accept a bigger container, or be willing to move their dry goods in a refrigerated container (although no refrigeration is necessary).
   c) Time windows and demand backlogging—The most common model represents customer demands at a point in time, where they are lost if they are not served at that point in time. In practice, it is usually the case that customer orders can be delayed.
   d) Transshipment and relay points—The simplest models represent a demand as the need to move from i to j, and where the movement is represented as a single decision. More complex operations have to model transportation legs (ocean or rail movements) with relays or transshipment points (ports, rail yards) where the containers move from one mode to the next. A major element of complexity is when capacity constraints are imposed on the transportation legs.

3) Managing complex equipment—The major railroads in North America need to manage fleets of several thousand locomotives. The air mobility command has to move freight and people on a global scale using different types of aircraft. Recently formed companies service a high end

market with personal jet service using jets in which the customers own a fraction. These problems have been modeled in the past using the same framework as container management problems with multiple container types. These complex pieces of equipment require something more. For example, there are four major classes of locomotive, reflecting whether they are high or low "adhesion" (a technology that determines the slippage of the wheels on a rail), and whether they are four axle or six axle units (six axle locomotives are more powerful). On closer inspection, we find that the horsepower rating of a locomotive can be divided into 10 or 12 reasonable divisions. It matters if the locomotive has its home shop in Chicago, Atlanta or southern California. Since locomotives may move from the tracks of one railroad to another, it matters who owns the locomotive. And it matters if the locomotive is due into the shop in $1,2,\ldots,10$ days, or more than 10 days. In short, complex equipment is complex, and does not lend itself easily to a multicommodity formulation. As we show later, this characteristic determines whether the size of the *attribute space* of a resource is small enough to enumerate the entire space, or too large to enumerate.

4) People and crews—Trucks, trains and planes move because people operate them. Not surprisingly, the modeling of the people is not only important, but requires a set of attributes that makes complex equipment look simple. A truck driver, for example, might be characterized by his current location, his home base, his skill level, whether he has experience driving into Mexico or Canada, how many hours he has driven in the last 8 days, how many consecutive hours he has been "on duty" today, and how many hours he has been actively driving during his current duty period. Production software systems have to cover these and many other issues.

These problems are all examples of resource allocation problems where, with few exceptions, a single "resource" serves a single "demand." "Bundling" arises when, for example, you need several locomotives to pull a single train, or two drivers (a sleeper team) to operate a single truck. "Layering" arises when you need an aircraft, a pilot, fuel and special loading equipment to move a load from one airbase to another. In some cases, the resource/task dichotomy breaks down. For example, we may be managing locomotives, crews and boxcars. The boxcar needs to go from A to B. We need the crew to move the train, but the crew needs to get back to its home domicile at C. And the locomotive needs to get to shop at D. We would refer to the locomotives, crew and boxcars as three *resource layers*, since the locomotives, crew and boxcars are all needed to move the train. In fact, for more complex problems, we refer to the objects being managed as *resource layers* (or sometimes, resource classes), where one layer is almost always one that would be referred to as a customer, or job, or task.

2.2 Sources of uncertainty

Uncertainty arises whenever we need to make a decision based on information that is not fully known. We are aware of three scenarios under which this can arise:

1) The information is not yet known, but will become known at some point in the future. This is the standard model of uncertainty.
2) Information is known to someone (or something), but is not known to the decision-maker. We would generally say that this information is *knowable* but for various reasons (most commonly, it is simply too expensive) has not been properly communicated to where the information is needed for a decision.
3) The information will never be known (optimization under incomplete information). For any of a variety of economic or technical reasons, an unknown variable is never measured, even though it would help improve decisions. Since the information is never known, we are not able to develop a probability distribution for it.

Cases (2) and (3) above both represent instances where decisions have to be made without information, but we assume that case (3) represents information that never becomes known explicitly, whereas (2) represents the case where someone knows the information, raising the possibility that the information could be shared (at a cost) or at a minimum, where a probability distribution might be constructed after the fact and shared with others.

Classical uncertainty arises because information arrives over time. It is possible to divide the different types of dynamic information processes into three basic classes: the ''resources'' being managed (including customer demands), the physical processes that govern the evolution of the system over time, and the decisions that are actually implemented to drive the system. This division reflects our modeling framework, presented in Section 3. Since the focus of this volume is on modeling uncertainty, it is useful to give each of these at least a brief discussion.

Resources

Under the heading of ''resources'' we include all the information classes that we are actively managing. More formally, these are ''endogenously controllable information classes which constrain the system,'' a definition that includes not just the trucks, trains and planes that we normally think of as resources, but also the customer orders that these resources are normally serving. Dynamic information processes for resources may include:

a) Information about new (exogenous) arrivals to the system—This normally includes the arrival of customer orders, but may also include the arrivals of the product, equipment or people required to satisfy the customer order. For example, a trucking company is constantly hiring

new drivers (there is a lot of turnover) so the arrival of new drivers to the fleet is a dynamic information process. Similarly, a railroad has to manage boxcars, and the process of boxcars becoming empty turns out to be a highly stochastic process (far more uncertain than the customer orders).

b) Information about resources leaving the system—Drivers may quit, locomotives may be retired from service, product can perish. The challenge of modeling departures is that they depend on the state of the system, whereas exogenous arrivals are normally modeled as being independent of the state of the system.

c) Information about the state of a resource—An aircraft may break down or a driver may call in sick.

An important dimension of the modeling of resources is the concept of *knowability* and *actionability*. It is not uncommon for a customer to call in and book an order in advance. Thus, the order becomes known right now (time t) but actionable when it actually arrives to the system at some point in the future (at time $t' \geq t$). Most stochastic models implicitly assume that a customer demand is not known until it actually arrives. By contrast, most deterministic models assume that we know all orders in advance (or more precisely, that we do not want to make a decision taking into account any order that is not already known). In practice, both extremes arise, as well as the case of prebooking where customers call at least some of their orders in advance.

Processes

Under this category, we include information about parameters that govern the evolution of the system over time. The most important classes include:

a) The time required to complete a decision—In most areas of transportation, travel times are random, and sometimes highly so (although applications vary in the degree to which random travel times actually matter). In air traffic control problems, planes may land at 20 minute intervals. Flights of several hours can easily vary in duration by 10 or 20 minutes, so they have to maintain a short backlog of flights to ensure that there is always an aircraft available to land when the runway has the capacity to handle another arrival. In railroads, it is not unusual for the travel time between two points to take anywhere from 5 to 8 days.

b) The cost of a decision—This is often the least uncertain parameter, but there are a number of reasons why we might not know the cost of a decision until after the fact. Costs which are typically not fully known in advance include tolls, transportation accidents, and processing costs that are not always easy to allocate to a particular activity. Even more uncertain is the revenue that might be received from satisfying a customer which might arise as a result of complex accounting procedures.

c) Parameters that determine the attributes of a resource after a decision—Examples might include the fuel consumption of an aircraft or locomotive (which determines the fuel level), or the maintenance status of the equipment at the end of a trip.

Controls

In real problems, there is a difference between the decisions that we are planning to make, and the decisions that are actually made. The flow of actual decisions is an important exogenous information process. There are several reasons why an actual physical system does not evolve as planned:

1) The decisions made by a model are not as detailed as what is actually needed in operations. The user has to take a *plan* developed by the model and convert it into something implementable.
2) The user has information not available to the model.
3) The user simply prefers to use a different problem solving approach (possibly suboptimal, but this assumes the solution provided by the model is in some way optimal).

When there is a difference between what a model recommends and the decisions that are actually made, we encounter an instance of the user non-compliance problem. This is a source of uncertainty that is often overlooked.

2.3 Special modeling issues in transportation

Transportation problems introduce an array of issues that provide special modeling and algorithmic challenges. These include:

a) Time staging of information—In freight transportation, information arrives over time. This is the heart of any stochastic model.
b) The lagging of information—Often, a customer will call at time t to place an order to be served at time $t' > t$. The same lagging of information may apply to the vehicles used to serve customers. Since we have information about the future, it is tempting to assume that we can make plans about the future, even before new information becomes known.
c) Complex resource attributes—It is often assumed that the number of different types of resources is "not too large." The number of resource types determines the number of constraints. In practice, the attributes of resources can be surprisingly complex, creating problems where the number of constraints can number in the millions. This is a challenge even for deterministic models, but poses special difficulties in the context of stochastic problems.
d) Integrality—Many transportation problems exhibit network structure that makes it much easier to obtain integer or near-integer solutions. This structure can be easily destroyed when uncertainty is introduced.

e) Travel times—The common behavior in transportation problems that it takes time to move from one location to the next is generally a minor issue in deterministic models. In stochastic models, it can introduce major complications. If the travel times are deterministic, the result can be a dramatic growth in the size of the state space. However, it is often the case that travel times not only are stochastic, they are not even measurable when the trip is initiated.
f) Multi-agent control—Large transportation systems might be controlled by different agents who control specific dimensions of the system. The decisions of other agents can appear as random variables to a particular agent.
g) Implementation—What we plan may not be the same as what actually happens. An overlooked source of uncertainty is the difference between planned and executed decisions.

2.4 Why do we need stochastic programming?

There are two types of modeling technologies that are widely used in practice: simulation models, which are used almost entirely for planning purposes where there is a need to understand the behavior of a system that evolves over time, and deterministic optimization models and algorithms, when there is a need for the computer to recommend what action should be taken. Stochastic programming brings the modeling of uncertainty explicitly into the process of making a decision (using an optimization algorithm). But, there is a large community of both academic researchers and consultants who feel that they are being quite productive with the algorithms that they are developing based on deterministic models.

There is a broad perception, in both the academic research community and in engineering practice, that deterministic optimization algorithms are "good enough." In part this can be attributed to both the mathematical maturity that has been required to understand stochastic models, and the lack of practical, problem-solving tools. But equally important, we need to understand the ways in which stochastic models can provide solutions that are not just better, but noticeably better in a way that would attract the attention of industry. An understanding of these issues will also indicate where stochastic models are not necessarily appropriate. A partial list of motivations for stochastic models should include:

1) The newsvendor effect—Providing the right amount of resource to meet demand given the uncertainty in demand and the relative costs of providing too much or too little. A deterministic model will never allocate more than the point forecast, even when there are excess resources. Stochastic models can overallocate or underallocate depending on the overall availability of resources to meet forecasted demands.

2) Robust allocation—We might need the container in city A or city C, but we are not sure, so we send the truck halfway in between to city B where it can wait and respond to the demand at the last minute. A deterministic model will never send capacity to a location that does not need it.
3) The value of advance information—Stochastic models can explicitly model the staging of information over time. A carrier might want to know the value of having customers book orders farther in advance. A proper analysis of this question needs to consider the value of reducing the uncertainty in a forecast.
4) Forecasts of discrete items—Sometimes it is necessary to forecast low volume demands; for example, orders might be 1 with probability 0.20 and 0 with probability 0.80. A point forecast would produce a demand of 0.20, but a routing and scheduling model is unable to assign 0.20 trucks to the order (the algorithm routes a single truck). Integer rounding amounts to little more than Monte Carlo sampling (simple rounding produces biases—it is necessary to round based on a random sample whose expectation is the same).
5) The algorithmic challenge of solving problems over extended planning horizons—Classical optimization algorithms struggle with optimization problems defined over long horizons, typically as a result of degeneracy. Formulations based on a stochastic "view" of the world produces time-staged problems that are much easier to solve. Sequences of two-stage problems are much easier to solve than a single, large integer program.
6) Overoptimizing problems with imperfect data—A deterministic view of the world can produce problems that are larger and more complex than necessary. An appreciation of uncertainty, not only of the future but also of the "here and now" data (which in practice is a major form of uncertainty) produces models that are smaller and more compact.

3 Modeling framework

The first chapter of this handbook provides a basic mathematical framework for multistage stochastic programming problems. The problem with these abstract formulations is spanning the gap between generic mathematical formulations and real problems. In this section, we offer a notational framework that helps to bridge the gap between real-world dynamic resource allocation problems, and the basic framework of math programming in general, and stochastic programming in particular.

We divide our modeling framework between three fundamental dimensions: the resources being managed, the processes that govern the dynamics of the system, and the structure and organization of controls which manage the system. Our presentation is not the most general, but allows us to focus on the

dimensions that are important for modeling the organization and flow of information.

3.1 Resources

To help formalize the discussion, we offer the following definition:

Definition 1. A resource is an endogenously controllable information class that constrains the system.

From a math programming perspective, a resource is anything that shows up as a right hand side of a constraint (no surprise that these are often referred to as "resource constraints"). For transportation, resources include trucks, trains, planes, boxcars, containers, drivers/crews, and special equipment that may be needed to complete a trip. Sometimes, but not always, the "demands" being served also meet this definition. For example, the load of freight that we are moving from one location to the next is both endogenously controllable (we often have to determine when to move the load, and sometimes how it is routed) and it constrains the system.

We describe resources using the following:

$\mathcal{C}^R$ = The set of resource classes (e.g., tractors, trailers, drivers, freight).
$\mathcal{R}_c$ = The set of (discrete) resources in class $c \in \mathcal{C}^R$.
a_r = The attributes of resource $r \in \mathcal{R}_c$, $c \in \mathcal{C}^R$.
$\mathcal{A}_c$ = The space of attributes for resource class $c \in \mathcal{C}^R$, with element $a^c \in \mathcal{A}_c$. We often use $\mathcal{A}$ to represent the attribute space of a generic resource.

The attribute vector is a very flexible device for describing the characteristics of a resource. In truckload trucking, it might be the case that all trucks are the same, in which case the attribute vector consists only of the location of the truck. In rail car distribution, the attribute vector can be the type of car as well as the location. If the resource is a human operator, the vector can grow to include attributes such as the home domicile, days away from home, hours of service, and skill sets.

The definition of the attribute space requires an understanding of how a resource evolves over time, and in particular the flow of information. For example, an air cargo carrier working for the military airlift command might have to move a load of cargo from the eastern United States to southeast Asia. This trip might require midair refueling, as well as stops at several intermediate airbases. Is it necessary to represent the aircraft at each of these intermediate points, or is it enough to assign the aircraft to move a load, and then model its status at the destination? The answer depends on the evolution of information and decisions. For example, if we can completely model all the steps of a trip using the information available when the aircraft first takes off from the origin, then there is no need to model the intermediate points. But we

might wish to model the possibility of a failure in the midair refueling, or the failure of the aircraft itself at any of the intermediate airbases. Both of these represent examples of new information arriving to the system, which requires modeling the status of the aircraft just before the new information arrives. The new information may produce new decisions (we may wish to reroute the aircraft) or a change in the dynamics (the aircraft may be unexpectedly delayed at an airbase).

The need to model our aircraft at intermediate points raises a new and even more complex issue. An aircraft that is fully loaded with freight takes on the characteristics of a *layered* (or composite) resource. That is, we have not only the characteristics of the aircraft, but also the characteristics of the freight on the aircraft. This sort of layering arises frequently in transportation operations. Another example arises in the management of locomotives. A locomotive may be sitting idle at a rail yard, or it may be attached to an inbound train (which is making an intermediate stop). If the locomotive is attached to an inbound train, then we have not only the attributes of the locomotive, but also of the train itself (such as its final destination).

We handle this behavior by defining layered attribute vectors. For example, let:

a^A = The attributes of an aircraft.
a^R = The attributes of a load of freight being moved (known as requirements).
a^C = The attributes of the crew piloting the aircraft.

When an aircraft is loaded and making a set of stops, then the attributes of the composite resource at the intermediate stops would be represented using:

$a^{(A)}$ = The attributes of the aircraft *layer*.
= $a^A \mid a^R \mid a^C$, where a^A, a^R and a^C are the attributes of the primitive aircraft, requirement and crew resources.

A layer is a concatenation of attributes. An aircraft which is currently sitting idle (a primitive resource) would have the attribute $a^{(A)} = a^A \mid a^\phi \mid a^\phi$.

In more complex problems, we may encounter three, four or even five layers. For these problems, we have to define in advance how resources may be combined.

Regardless of our problem class, we let:

$R_{t,a}$ = The number of resources with attribute $a \in \mathcal{A}$ at time t.
R_t = $(R_{t,a})_{a \in \mathcal{A}}$.

One issue that often arises in transportation is the concept of *knowability* and *actionability*. We may *know* of a resource r with attribute a_r at time t

which is not *actionable* until some time $t' > t$. This can arise when a customer calls in an order in advance, or when a plane takes off from airport i at time t but will not arrive at airport j until time t'. Actionability can arise as an "estimated time of arrival," an order pickup time, or the time when a task (such as maintenance) will be finished. Actionability can be viewed as being simply an attribute of a resource, and therefore part of the vector a. But often, the actionable time is sufficiently important that it needs to be represented explicitly. In this case, we write:

$R_{t,at'}$ = Number of resources that we know about with attribute a at time t that will not be actionable until time $t' \geq t$.
$R_{tt'} = (R_{t,at'})_{a \in \mathcal{A}}$.
$R_t = (R_{tt'})_{t' \geq t}$.

Thus, we can continue to use the vector R_t as our general state vector, recognizing that it may be divided into elements $R_{tt'}$.

This discussion illustrates a division in the operations research community on the meaning of a time index. Deterministic models of time-staged processes always use time to refer to when an action will happen ("actionability"). Stochastic models almost always use time to refer to the information content of a variable ("knowability" or, in formal terms, "measurability"). In general problems, it is necessary to use both, but this can sometimes be clumsy. We use the double time index (t, t') when we want to explicitly refer to the information content of a variable ("t"), and when an activity actually takes place ("t'"). Whenever we use a single time index, such as R_t, we will always intend the time index to refer to the information content.

3.2 Processes

A dynamic process evolves because of two types of information processes: exogeneous information processes, that arrive as a series of events which update the state of the system, and endogenous information processes, otherwise known as decisions. Following the conventions described in the first chapter of this volume, we let:

ξ_t = The information arriving in time period t. ξ can represent new information about customer demands, new equipment entering the system, equipment breakdowns, and travel delays.
$\xi = (\xi_t)_{t \in \mathcal{T}}$
= The information process over the model horizon represented by the set of time periods $\mathcal{T}$.

In general, new information arriving from external sources is captured in a *knowledge base* which summarizes all the information known at time t. Following standard convention, we let $\mathcal{F}_t$ be the σ-algebra generated by the vector $(\xi_0, \ldots, \xi_t)$.

The standard representation of information in real problems does not always follow standard assumptions. To illustrate, let:

K_t = Our (data) knowledge base at time t.
U^K = The knowledge updating function which updates K_{t-1} using new information ξ_t.

We would representing our updating process as:

$$K_t \leftarrow U^K(K_{t-1}, \xi_t)$$

Realizing that $\mathcal{F}_{t-1} \subseteq \mathcal{F}_t$, one would expect that $\sigma(K_t)$ (the σ-algebra generated by the random variable K_t) would satisfy $\sigma(K_{t-1}) \subseteq \sigma(K_t)$. This assumes that computer databases do not "forget" information. But this is not always the case. It is not our intent to raise this as a serious issue, but just as a reminder to the reader that standard mathematical assumptions do not always apply to the real world.

For our problems, we can typically divide new information into two classes: the arrivals of new resources (including new customer demands, as well as new equipment or new drivers), and information about model parameters (such as costs and times). This distinction is important in our problem representation, so we define:

$\hat{\rho}_t$ = Updates to model parameters arriving in time period t.
$\hat{R}_{tt'}$ = The vector of new resources arriving in time period t that become actionable at time $t' \geq t$.
$\hat{R}_t$ = $(\hat{R}_{tt'})_{t' \geq t}$.

Thus, we would write $\xi_t = (\hat{\rho}_t, \hat{R}_t)$ with sample realization $\omega_t = \xi_t(\omega) = (\hat{\rho}_t(\omega), \hat{R}_t(\omega))$.

We represent decisions using:

$\mathcal{C}^D$ = The set of decision classes (move empty, move loaded, refuel, maintain the equipment, have a driver go on rest, etc.)
$\mathcal{D}_c$ = The set of discrete decisions in decision class $c \in \mathcal{C}^D$.
$\mathcal{D}$ = $\cup_{c \in \mathcal{C}^D} \mathcal{D}_c$

We use $\mathcal{D}$ to refer to the complete set of decisions. In most transportation applications, it is useful to capture the fact that the set of decisions also depends on the attribute of the resource being acted on. For this purpose we define:

$\mathcal{D}_a$ = The set of decisions that can be used to act on a resource with attribute $a \in \mathcal{A}$.

For the purposes of our presentation, we consider only direct decisions that act on the attributes of a resource (this would exclude, for example, decisions about pricing or what speed to fly an aircraft). For transportation problems, if $d \in \mathcal{D}$ is an instance of a decision, then the impact of the decision is captured through the *modify* function, which is a mapping:

$$M(K_t, a, d) \rightarrow (a', c, \tau) \tag{2.1}$$

where d is a decision acting on a (possibly layered) resource with attribute a at time t producing a resource with attribute a', generating a contribution c and requiring time τ to complete the action. a', c and τ are all functions, which we can represent using the triplet $(a^M(t,a,d), c^M(t,a,d), \tau^M(t,a,d))$ (for notational compactness, we index these functions by time t instead of modeling the explicit dependence on K_t). We call $a^M(t,a,d)$ the *terminal attribute function*. Normally, we represent the costs and times using the vectors $c_{tad} = c^M(t,a,d)$ and $\tau_{tad} = \tau^M(t,a,d)$. We note as an aside that while we will usually model $(a^M(t,a,d), c^M(t,a,d), \tau^M(t,a,d))$ as $\mathcal{F}_t$—measurable, this is certainly not always the case. For example, Section 4 describes an application in rail car distribution. In this application, empty freight cars are moved to customers to move loads of freight. The destination of a load is typically not known until the car is released loaded back to the railroad. The travel time of the movement is not known until the car actually reaches the destination.

The set $\mathcal{D}$ is the set of *types* of decisions we make. The decision vector itself is represented using:

x_{tad} = The number of times that we act on a resource with attribute a using decision d at time t.
$x_t = (x_{tad})_{a \in \mathcal{A}, d \in \mathcal{D}}$.
= The vector of decisions at time t.

Letting c_t similarly represent the vector of contributions at time t provides for a compact representation that matches standard modeling notation. Most transportation costs are linear in the decision variables, and as a result, the total contribution at time t can be written as:

$$\begin{aligned} C_t(x_t) &= \sum_{a \in \mathcal{A}} \sum_{d \in \mathcal{D}} c_{tad} x_{tad} \\ &= c_t x_t. \end{aligned}$$

It is important to realize that our notation for stochastic problems is different in a subtle but important way than the notation conventionally used in deterministic transporation models. For example, it is normal to let x_{ijt} be the flow from location i to location j departing at time t. The index j effectively

presumes a deterministic outcome of the decision (the notation $x_{ijt}(\omega)$ does not fix the problem; we would have to write $x_{i,j(\omega),t}$ which is quite ugly). We might not question the outcome of a decision to send a truck or plane from i to j (frequent fliers will remember at least one occasion when the plane did not arrive at the proper destination as a result of weather problems). But in more complex problems where we are capturing a larger vector of attributes, the terminal attribute function $a^M(t, a, d)$ cannot in general be assumed to be a deterministic function of (t, a, d). The representation of a decision using x_{tad} is important for stochastic problems since the variable is indexed only by information available when the decision is made.

For algebraic purposes, it is useful to define:

$$\begin{aligned}\delta_{t',a'}(t, a, d) &= \text{Change in the system at time } t' \text{ given a decision} \\ &\qquad \text{executed at time } t. \\ &= \begin{cases} 1 & \text{if } M_t(t, a, d) = (a', \cdot, t' - t) \\ 0 & \text{otherwise} \end{cases}\end{aligned}$$

We note that if d represents a decision to couple two resources, then a is the attributes of the resource, d contains the information about the resource being coupled with, and a' is the concatenation of two attribute vectors.

Using this notation, we can now write the dynamics of our resource variable (incorporating the time-lagging of information):

$$R_{t+1,a't'} = R_{t,a't'} + \hat{R}_{t+1,a't'}(\omega) + \sum_{d\in\mathcal{D}} \sum_{a\in\mathcal{A}} \delta_{t',a'}(t, a, d) x_{tad} \quad a' \in \mathcal{A}, \quad t' > t. \tag{2.2}$$

3.3 Controls

It is common in transportation problems to focus on decisions that move resources from one location to the next. While this is the most obvious dimension, it is important to capture other types of decisions.

Our notation for representing decisions offers considerable flexibility. It is a common misconception in the modeling of transportation systems that decisions always represent movements from one location to another. Examples of different classes of decisions other than spatial movements include: cleaning dirty vehicles, repairing or maintaining equipment, sending a driver off-duty, using outside contractors to perform a task, transferring rail cars from one shipper pool to another (this is a form of classification, and does not mean moving from one location to another), buying/selling/leasing equipment, and hiring/firing drivers.

In deterministic problems, decisions are made by solving a particular instance of an optimization problem. In stochastic problems, we have to capture the time staging of decisions and information. We represent the process of making decisions at time t using:

I_t = The set of information available for making a decision.
$X_t^\pi(I_t)$ = The decision function of policy $\pi \in \Pi$ which returns a vector x_t given the information set I_t.

In Section 3.6, we describe different classes of information, and the types of decision functions these produce.

For our problems, the decision function will be some sort of mathematical program, since the decisions typically are vectors, possibly of fairly high dimensionality. Later we provide specific examples of decision functions, but for now, we simply assume that they produce feasible solutions. The most important constraint that must be satisfied is flow conservation:

$$\sum_{d \in \mathcal{D}} x_{tad} = R_{ta} \quad \forall a \in \mathcal{A}.$$

In addition, the flows must be nonnegative and, in many applications (virtually all involving operational problems in transportation) integer.

3.4 Modeling state variables

It is useful at this point to make a brief comment about "state variables," since these take on different meanings in different communities. In our modeling framework, the attribute vector a captures the "state" of a particular resource. $R_t = (R_{ta})_{a \in \mathcal{A}}$ is the "state" of the vector of resources. I_t (which we have not completely defined) is the "information state" of the system. In some subcommunities (notably, people who solve crew scheduling problems using column generation techniques), the management of multiple resources is decomposed into subproblems involving the optimization of a single resource. In this context, someone might talk about a large "state space" but refer to the attribute space of a single resource.

It is very common in the operations research literature (most commonly in the context of dynamic programming and Markov decision processes) to talk about the "state" of the system, where the state variable captures the amount of product being stored or the customer demands that have been backlogged. In this setting, the "state" of the system refers to the resource state variable, R_t. Even recently, discrete dynamic programming models have been proposed using R_t as the state variable. Not surprisingly, the number of possible realizations of R_t (assuming it is discrete) will be huge even for toy problems.

Of course, the real state variable must be what we know or, literally, the state of our knowledge, which we denote by K_t. Other authors refer to this as the *information state*. We let I_t be the information state, but claim that there are potentially four classes of information:

a) Knowledge—This is the data in the vector K_t, capturing the exogenous data that has been provided to the system.
b) Forecasts of exogenous processes—This is information from a forecasting model, representing projections of what might happen in the future. If we are making a decision at time t, this would be a projection of $(\hat{\xi}_{t+1}, \hat{\xi}_{t+2}, \ldots, \hat{\xi}_T)$. We may use a point forecast of future events, or forecast a set of future scenarios which would be represented using the set $\hat{\Omega}_t$ (the set of future events forecasted at time t). If $|\hat{\Omega}_t| = 1$, then we are using a traditional point forecast.
c) Forecasts of the impact of decisions now on the future. In this chapter, this dimension will be captured through the *recourse function* and hence we denote the set of possible recourse functions, estimated at time t (but capturing the impact on the future) by $\mathcal{Q}_t$.
d) Plans—These are projections of decisions to be made in the future, which can be expressed in a variety of ways (it is useful to think of these as *forecasts* of future decisions). A convenient way is to represent them as a vector of decisions $x_t^p = (x_{tt'}^p)_{t' \geq t}$, where $x_{tt'}^p$ is the plan for time t' using the information available at time t. We note that plans are almost always expressed at some level of aggregation. Normally, we use plans as a guide and penalize deviations from a plan.

The last three classes of information are all forms of forecasts. We assume that these are generated from data that is a function of K_t. However, while a forecast is generated from knowledge, they do not represent knowledge itself. All companies seek to improve decision-making by improving the knowledge base K_t, but they also consider the value of including forecasts (many transportation companies do not perform short term operational forecasts, and most research into problems such as dynamic vehicle routing does not use forecasts) or future plans. Companies make explicit decisions to add these classes of information to their decision making process (and adjust the process accordingly).

Using this definition of information, the information state can come in a variety of forms, such as $I_t = (K_t)$, $I_t = (K_t, \hat{\Omega}_t)$, $I_t = (K_t, x_t^p)$ and $I_t = (K_t, Q_t)$. Later we show that different classes of information give rise to the major classes of algorithms known in the operations research community. For the moment, it is necessary only to understand the different ways of representing the "state" of the system. Our notation contrasts with the standard notation S_t for a state variable. The problem is that S_t is not very explicit about what is comprising the state variable. We suggest using S_t when we want to refer to a generic "state," and use a, R_t, K_t or I_t when we want to express explicit dependence on, respectively, the attribute of a single

resource, the resource state vector, the entire knowledge base, or a broader information set.

Using these notions of state variables, it is useful to revisit how we write our cost and decision functions. The representation of costs and decisions using the notation c_{tad} and x_{tad} suggests that both the costs and decisions are a function only of the attribute vector of the resource, although this does not have to be the case. We may write the decision function as $X^{\pi}(R_t)$ if all other types of information are static. The reader may write $X^{\pi}(K_t)$ to express the explicit dependence on the larger knowledge base, but this generality should be reserved for problems where there are parameters which are evolving over time, and whose values affect the forward evolution of the system.

3.5 The optimization problem

Our problem is to find a decision function X^{π} that solves the following expression:

$$F^* = \sup_{\pi \in \Pi} EF^{\pi} \tag{2.3}$$

$$= \sup_{\pi \in \Pi} E\left\{\sum_{t \in \mathcal{T}} C_t(X_t^{\pi}(I_t))\right\}. \tag{2.4}$$

The system has to respect the following equations governing the physical and information dynamics:

Physical dynamics:

$$R_{t+1,a't'}(\omega) = R_{t,a't'}(\omega) + \hat{R}_{t+1,a't'}(\omega) + \sum_{d \in \mathcal{D}} \sum_{a \in \mathcal{A}} \delta_{t',a'}(t, a, d) x_{tad} \quad a' \in \mathcal{A},\ t' > t. \tag{2.5}$$

Informational dynamics:

$$K_{t+1} = U^K(K_t, \xi_{t+1}). \tag{2.6}$$

The decision function X_t^{π} is assumed to produce a feasible decision. For this reason, flow conservation constraints and upper bounds are not included in this formulation.

The optimization problem is one of choosing a function. The structure of the decision function depends on the information available. Within an information class, a decision function is typically characterized by a family of parameters and we have to choose the best value for these parameters.

3.6 A brief taxonomy of problems

Using our modeling framework, we can provide a brief taxonomy of major problem classes that arise in transportation. We divide our taxonomy along the three major dimensions of resources, processes and controls.

Resources

By just using the attribute vector a notation, we can describe six major problem classes in terms of the resources being managed:

1) Basic inventory problems—$a = \{\}$ (no attributes). This is the classical single product inventory problem.
2) Multiproduct inventory problems—$a = \{k\}$ where $k \in \mathcal{K}$ is a product type.
3) Single commodity flow problems—$a = \{i\}$ where $i \in \mathcal{I}$ is a state variable (such as a city or geographical location).
4) Multicommodity flow problems—$a = \{i, k\}$ where $i \in \mathcal{I}$ is a state variable (such as a location) and $k \in \mathcal{K}$ is a commodity class.
5) Heterogeneous resource allocation problem—$a = \{a_1, a_2, \ldots, a_N\}$. In these more complex problems, it is possible to divide the attribute vector into static attributes, a^s which do not change over time, and dynamic attributes, a^d, which do change. Writing $a = \{a^s, a^d\}$, we can think of a^d as a resource state variable, and a^s as a resource type variable.
6) The multilayered resource allocation problem—$a = \{a^1 \mid a^2 \mid \cdots \mid a^L\}$ where a^c is the attributes of resource class c. Here, a is a concatenation of attribute vectors.

Although the sixth class opens the door to multilayered problems, it is useful to divide resource allocations between single layer problems, two-layer problems (which most often involve an *active* resource layer representing people or equipment, and a *passive* layer representing customer requests), and multilayer problems.

We focus on single layer problems in this chapter, which include the first five types of attribute vectors. Of these, the first four are typically characterized by small attribute spaces, where it is possible to enumerate all the elements in $\mathcal{A}$, while heterogeneous resource allocation problems are typically characterized by an attribute space that is too large to enumerate. As we point out later, this creates special problems in the context of stochastic resource allocation problems.

System dynamics

Under the heading of system dynamics, we divide problems along three major dimensions:

1) The time staging of information—The two major problem classes are:
   a) Two-stage problems.
   b) Multistage problems.

2) Travel times (or more general, decision completion times). We define two major classes:
   a) Single-period times—$\tau_{tad} = 1$ for all $a \in \mathcal{A}$, $d \in \mathcal{D}$.
   b) Multiperiod times—$1 \leq \tau_{tad} \leq \tau^{max}$. We assume that $\tau_{tad} \geq 1$ but we can relax this requirement and model problems where $\tau_{tad} = 0$.
3) Measurability of the modify function. We again define two major classes:
   a) The function $M(t, a, d)$ is $\mathcal{F}_t$—measurable. This means that $(a^M(t, a, d), c^M(t, a, d), \tau^M(t, a, d))$ is deterministic given a, d and other parameters that are known at time period t.
   b) The function $M(t, a, d)$ is not $\mathcal{F}_t$-measurable. This is common, although we are not aware of any research addressing this issue.

Controls

We first divide problems into two broad classes based on control structure:

1) Single agent control structure—The entire company is modeled as being controlled by a single agent.
2) Multiagent control structure—We model the division of control between multiple agents.

Starting with the single agent control structure, we can organize problems based on the information available to make a decision. Earlier, we described four classes of information. We can now describe four classes of algorithms built around these information sets:

a) $I_t = (K_t)$—This is our classic myopic algorithm, widely used in simulations. This is also the standard formulation used (both in practice and in the research community) for dynamic vehicle routing problems, and other on-line scheduling problems.
b) $I_t = (K_t, \hat{\Omega}_t)$—If $|\hat{\Omega}_t| = 1$, this is our classical rolling horizon procedure using a point forecast of the future. This represents standard engineering practice for fleet management problems and other dynamic resource allocation problems. If $|\hat{\Omega}_t| > 1$, then we would obtain a scenario-based stochastic programming model. The use of these formulations for multistage problems in transportation and logistics is very limited.
c) $I_t = (K_t, x_t^p)$—Here we are making decisions reflecting what we know now, but using plans to help guide decisions. This information set typically gives rise to proximal point algorithms, where the proximal point term penalizes deviations from plan.
d) $I_t = (K_t, Q_t)$—This information set gives rise to dynamic programming formulations, Benders decomposition and other methods for approximating the future. Typically, the recourse function Q_t is itself a function

of a distributional forecast $\hat{\Omega}_t$, so it is appropriate to write $Q_t(\hat{\Omega}_t)$ to express this dependence.

This breakdown of different types of decision functions, each based on different types of information, nicely distinguishes engineering practice ($I_t = (K_t)$ or $I_t = (K_t, \hat{\Omega}_t)$ with $|\hat{\Omega}| = 1$) from the stochastic programming literature ($I_t = (K_t, \hat{\Omega}_t)$ with $|\hat{\Omega}| > 1$ or $I_t = (K_t, Q_t)$). The use of proximal point algorithms has been studied in the stochastic programming literature, but the use of plans (generated from prior data) to help guide future decisions is often overlooked in the modeling and algorithmic community. If stochastic programming is to gain a foothold in engineering practice (within the transportation and logistics community), it will be necessary to find the problem classes where the more advanced decision sets add value.

Complex problems in transportation, such as railroads, large trucking companies and the air traffic control system, are characterized by multiple decision-making agents. We would represent this structure by defining:

$\mathcal{D}_q$ = The subset of decisions over which agent q has control.
I_{tq} = The information available to agent q at time t.

Then $X^{\pi}_{tq}(I_{tq})$ is the decision function for agent q given information I_{tq} at time t.

Multiagent systems capture the organization of information. By contrast, classical stochastic programming models focus on the flow of information. In transportation, modeling information is important, but we typically have to capture both the organization and flow. We also find that in a multiagent system, we may have to forecast the behavior of another agent (who may work within the same company). This can be an important source of uncertainty in large operations.

4 A case study: freight car distribution

When moving freight by rail (for the purposes of this discussion, we exclude the movement of intermodal freight such as trailers and containers on flatcars), a shipper requests one or more cars, of a particular type, at his dock for a particular day. The request may be for one or two cars, or as many as 100 or more. The railroad identifies specific cars that can be assigned to the request, and issues a ''car movement order'' to get the car to the shipper. The car may be in a nearby yard, requiring only the movement of a ''local'' train to get the car to the shipper. Just as easily, the car may have to move from a much farther location through a sequence of several trains before arriving at the final destination.

Freight cars come in many types, often looking the same to the untrained eye but appearing very different to the shipper. For example, there are 30 types

of open top gondola cars (“gons” in the industry). When a railroad cannot provide the exact type of car from the closest depot on the correct day, it may resort to three types of substitution:

1) Geographic substitution—The railroad may look at different sources of cars and choose a car that is farther away.
2) Temporal substitution—The railroad may provide a car that arrives on a different day.
3) Car type substitution—The railroad may try to satisfy the order using a slightly different car type.

Once the decision has been made to assign a car to a customer request, the railroad begins the process of moving a car to the destination. If the car is far away, this may require movements on several trains, passing through one or more intermediate *classification yards* which handle the sorting process. Travel times are long, and highly variable. It can take up to two or three weeks to move an empty car to a customer, wait for it to load, move it loaded, and then wait for it to unload (known as a car cycle). Travel times typically range between one and five days or more. Travel times between a pair of locations that averages six days can see actual transit times between four and eight days.

From the perspective of car distribution, there are three important classes of dynamic information: the flow of customer requests for capacity, the process of cars becoming empty (either because a shipper has emptied and released the car or because another railroad has returned the car empty), and the travel times for cars moving from one location to another. Customer orders are typically made the week before the car is actually needed, but some orders are made more than a week in advance, and some orders are made at the last minute (especially from large, high priority customers). There is very little advance information about empty cars, and of course, transit times are only known after the movement is completed. Thus, we see information processes where the difference when a resource is knowable and actionable is large (customer orders), small (empty cars), and where the modify function is not $\mathcal{F}_t$-measurable.

It is useful to get a sense of the variability of the data. Fig. 1 is an actual graph of the demand for cars at a regional level, showing actual, predicted, and both 10th and 90th percentiles. This graph ignores the presence of booked orders, and in practice, most orders are known a week into the future. For this reason, customer orders are not the largest source of uncertainty in an operational model. A much more significant source of error arises from the forecast of empty cars. Fig. 2 shows a similar graph for a particular type of freight car at a specific location. We again see a large degree of variability. In this case, there is little advance information.

One of the most difficult sources of uncertainty arises in transit times. In railroads, it is not unusual to see transit times that range between 5 and 10 days. This source of noise is particularly problematic. It means that if we ship 10 cars from i to meet a demand at j, we are not sure when they will arrive.

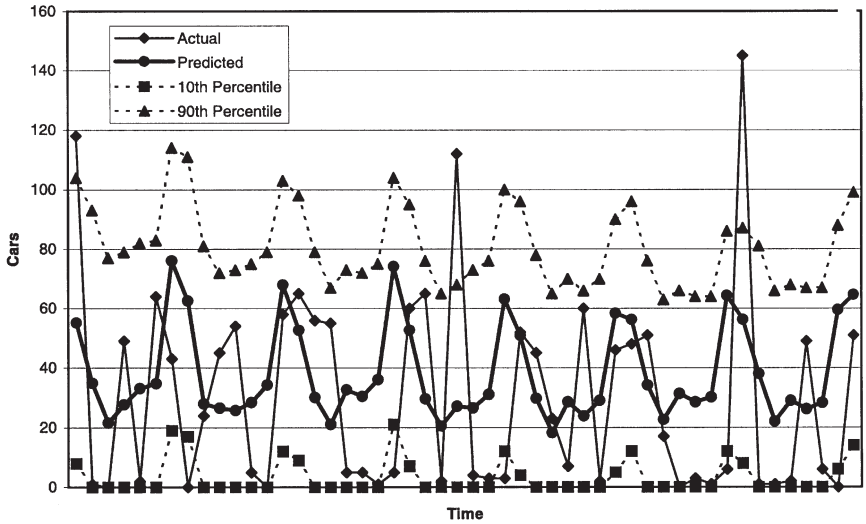

Fig. 1. Actual vs. predicted forecasts of future demands for empty cars, showing the 10th and 90th percentiles.

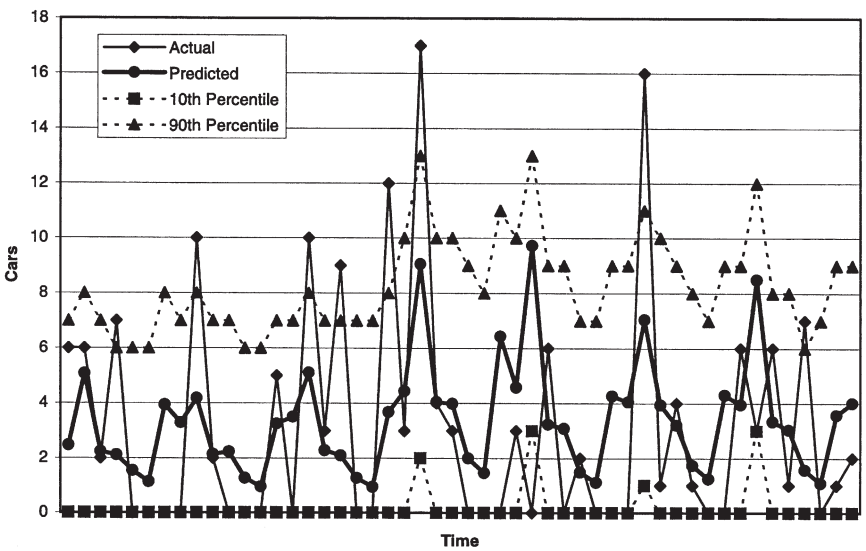

Fig. 2. Actual vs. predicted forecasts of supplies of empty cars, showing the 10th and 90th percentiles.

It has been suggested that we can improve our forecast of empty cars becoming available by using what we know about cars that are currently moving loaded (we know where they are going, so if we could estimate the transit time, we could estimate when they are becoming available). The uncertainty of transit times complicates this analysis.

We are now ready to consider more carefully the decision classes that govern the problem. As a result of the long travel times and high degree of uncertainty, it is not possible to simply wait until orders become known before a car is assigned to satisfy the order. The situation is further complicated by the fact that they cannot always let a car sit until there is an order to assign it to. A car may become available at a location that does not have the capacity to store the car. As a result, the railroad faces four possible classes of decisions when a car becomes empty:

1) Send it directly to a customer who has booked an order. Normally, we assume that this decision is to assign a car to a specific order, but it could be modified to send the car to a customer (where it would be assigned to a specific order after it arrives).
2) Send it to a *regional depot* which only serves customers in the region.
3) Send it to a *classification yard* where cars can be sorted and moved out on different trains. A classification yard at a railroad is a major facility and represents a point where it is easiest to make a decision about a car. From a classification yard, a car may be sent to another classification yard, a regional depot or directly to a customer.
4) Do nothing. This means storing the car at its current location. This is generally not possible if it just became available at a customer, but is possible if it is at a storage depot.

Not every car can be immediately assigned to an order, partly because some orders simply have not been booked yet, and partly because there are times of the year when there are more cars than we need. At the same time, one would expect that we do not always assign a car to a particular order, because not all the available cars are known right now. However, there is a strong bias to find an available car that we know about right now (even if it is a longer distance from the order) than to use a car that might become available later.

5 The two-stage resource allocation problem

We start with the two-stage problem because it is fundamental to multistage problems, and because some important algorithmic issues can be illustrated with minimum complexity. It should not be surprising that we are going to solve multistage problems basically by applying our two-stage logic over and over again. For this reason, it is particularly important that we be able to understand the two-stage problem very well.

We begin our presentation in Section 5.1 with a brief discussion of our notational style. Two-stage problems are relatively simple, and it is common to use notational shortcuts to take advantage of this simplicity. The result, however, is a formulation that is difficult to generalize to harder problems. Section 5.2 summarizes some of the basic notation used specifically for the car

distribution problem. We introduce our first model in Section 5.3 which presents models that are in practice today. We then provide three levels of generalization on this basic model. The first (Section 5.4) introduces uncertainty without any form of substitution, producing the classical "stochastic programming with simple recourse" formulation. The second models the effect of regional depots (Section 5.5), which produces a separable two-stage problem which can be solved using specialized techniques. The last model considers classification yards which requires modeling general substitution (Section 5.6), and brings into play general two-stage stochastic programming, although we take special advantage of the underlying network structure. Finally, Section 5.7 discusses some of the issues that arise for problems with large attribute spaces.

5.1 Notational style

One of the more subtle modeling challenges is the indexing of time. In a two stage problem, this is quite simple. Often, we will let x denote an initial decision, followed by new information (say, ξ), after which there is a second decision (perhaps denoted by y) that is allowed to use the information in the random variable ξ.

This is very simple notation, but does not generalize to multistage problems. Unfortunately, there is not a completely standard notation for indexing activities over time. The problem arises because there are two processes: the information process, and the physical process. Within the information process, there is exogenous information, and the process of making decisions (which can be viewed as endogenously controllable information). In many problems, and especially true of transportation, there is often a lag between the information process (when we know about an activity) and the physical process (when it happens). (We ignore a third process, which is the flow of financial rewards, such as billing a customer for an activity at the end of a month.)

In the operations research literature, it is common to use notation such as x_t to represent the vector of flows occurring (or initiating) in time t. This is virtually always the case in a deterministic model (which ignores completely the time staging of information). In stochastic models, it is more common (although not entirely consistent) to index a variable based on the information content. In our presentation, we uniformly adopt the notation that any variable indexed by time t is able to use the exogenous information up through and including time t (that is, $\xi_0, \xi_1, \ldots, \xi_t$). If x_t is a decision made in time t, then it is also allowed to see the information up through time t. It is often useful to think of ξ_t as information arriving "during time period t" whereas the decision x_t is a function determined at the end of time period t.

We treat $t = 0$ as the starting point in time. The discrete time $t = 1$ refers to the time interval between 0 and 1. As a result, the first set of new information would be ξ_1. If we let S_0 be our initial state variable, we can make an initial

decision using only this information, which would be designated x_0. A decision made using ξ_1 would be designated x_1.

There may be a lag between when the information arrives about an activity and when the activity happens. It is tempting, for example, to let D_t be the demands that arrive in period t, but we would let D_t be the orders that become known in time period t. If a customer calls in an order during time interval t which has to be served during time interval t', then we would denote this variable by $D_{tt'}$. Similarly, we might make a decision in time period t to serve an order in time period t'; such an activity would be indexed by $x_{tt'}$.

A more subtle notational issue arises in the representation of state variables. Here we depart from standard notation in stochastic programming which typically avoids an explicit definition of a state variable (the "state" of the system going into time t is the vector of decisions made in the previous period x_{t-1}). In resource allocation problems, vectors such as x_t can have a very large number of dimensions. These decisions produce future inventories of resources which can be represented using much lower dimensional state variables. In practice, these are much easier to work with.

It is common in multistage problems to let S_t be the state of the system at the beginning of time period t, after which a decision is made, followed by new information. Following our convention, S_t would represent the state after the new information becomes known in period t, but it is ambiguous whether this represents the state of the system before or after a decision has been made. It is most common in the writing of optimality equations to define the state of the system to be all the information needed to make the decision x_t. However, for computational reasons, it is often useful to work in terms of the state of the system immediately after a decision has been made. If we let S_t^+ be the *complete* state variable, giving all the information needed to make a decision, and let S_t be the state of the system immediately after a decision is made, the history of states, information and decisions up through time t would be written:

$$h_t = \{S_0^+, x_0, S_0, \xi_1, S_1^+, x_1, S_1, \xi_2, S_2^+, x_2, S_2, \ \ldots, \xi_t, S_t^+, x_t, S_t, \ldots\}. \tag{5.1}$$

We sometimes refer to S_t as the *incomplete* state variable, because it does not include the information ξ_{t+1} needed to determine the decision x_{t+1}. For reasons that are made clear later (see Section 6.2), we find it more useful to work in terms of the incomplete state variable S_t (and hence use the more cumbersome notation S_t^+ for the complete state variable).

In this section, we are going to focus on two-stage problems, which consist of two sets of decision vectors (the initial decision, and the one after new information becomes known). We do not want to use two different variables (say, x and y) since this does not generalize to multistage problems. It is tempting to want to use x_1 and x_2 for the first and second stage, but we find

that the sequencing in equation (5.1) better communicates the flow of decisions and information. As a result, x_0 is our first stage decision while x_1 is our second stage decision.

5.2 Modeling the car distribution problem

Given the complexity of the problem, the simplicity of the models in engineering practice is amazing. As of this writing, we are aware of two basic classes of models in use in North America: myopic models, which match available cars to orders that have already been booked into the system, and models with deterministic forecasts, which add to the set of known orders additional orders that have been forecasted. We note that the railroad that uses a purely myopic model is also characterized by long distances, and probably has customers which, in response to the long travel times, book farther in advance (by contrast, there is no evidence that even a railroad with long transit times has any more advance information on the availability of empty cars). These models, then, are basically transportation problems, with available cars on the left side of the network and known (and possibly forecasted) orders on the right side.

The freight division of the Swedish National Railroad uses a deterministic time–space network to model the flows of loaded and empty cars and explicitly models the capacities of trains. However, it appears that the train capacity constraints are not very tight, simplifying the problem of forecasting the flows of loaded movements. Also, since the model is a standard, deterministic optimization formulation, a careful model of the dynamics of information has not been presented, nor has this data been analyzed.

The car distribution problem involves moving cars between the locations that handle cars, store cars and serve customers. We represent these using:

$\mathcal{I}^c$ = Set of locations representing customers.
$\mathcal{I}^{rd}$ = Set of locations representing regional depots.
$\mathcal{I}^{cl}$ = Set of locations representing classification yards.

It is common to represent the "state" of a car by its location, but we use our more general attribute vector notation since it allows us to handle issues that arise in practice (and which create special algorithmic challenges for the stochastic programming community):

$\mathcal{A}^c$ = The set of attributes of the cars.
$\mathcal{A}^o$ = The set of attributes of an order, including the number of days into the future on which the order should be served (in our vocabulary, its actionable time).
$R^c_{t,at'}$ = The number of cars with attribute a that we know about at time t that will be available at time t'. The attribute vector includes the location of the car (at time t') as well as its characteristics.

$R^o_{t,at'}$ = The vector of car orders with attribute $a \in \mathcal{A}^o$ that we know about at time t which are needed to be served at time t'.

Following the notational convention in equation (5.1), we let $R^{+,c}_0$ and $R^{+,o}_0$ be the initial vectors of cars and orders at time 0 before any decisions have been made, whereas R^c_0 and R^o_0 are the resource vectors after the initial decision x_0 has been implemented.

It is common to index variables by the location. We use a more general attribute vector a, where one of the elements of an attribute vector a would be the location of a car or order. Rather than indexing the location explicitly, we simply make it one of the attributes.

The decision classes are given by:

$\mathcal{D}^c$ = The decision class to send cars to specific customers, where $\mathcal{D}^c$ consists of the set of customers (each element of $\mathcal{D}^c$ corresponds to a location in $\mathcal{I}^c$).

$\mathcal{D}^o$ = The decision to assign a car to a type of order. Each element of $\mathcal{D}^0$ corresponds to an element of $\mathcal{A}^o$. If $d \in \mathcal{D}^o$ is the decision to assign a type of car (as apposed to a particular car), we let $a_d \in \mathcal{A}^o$ be the attributes of the car type associated with decision d.

$\mathcal{D}^{rd}$ = The decision to send a car to a regional depot (the set $\mathcal{D}^{rd}$ is the set of regional depots—we think of an element of $\mathcal{I}^{rd}$ as a regional depot, while an element of $\mathcal{D}^{rd}$ as a decision to go to a regional depot).

$\mathcal{D}^{cl}$ = The decision to send a car to a classification yard (each element of $\mathcal{D}^{cl}$ is a classification yard).

d^ϕ = The decision to hold the car (''do nothing'').

The different decision classes are illustrated in Fig. 3, where a car can be shipped directly to a customer, a regional depot, or a classification yard.

Our complete set of decisions, then, is $\mathcal{D} = \mathcal{D}^c \cup \mathcal{D}^o \cup \mathcal{D}^{rd} \cup \mathcal{D}^{cl} \cup d^\phi$. We assume that we only act on cars (cars are the only *active* resource class, whereas orders are referred to as a *passive* resource class). We could turn orders into an active resource class if we allowed them to move without a car (this would arise in practice through outsourcing of transportation). Of these, decisions in $\mathcal{D}^o$ are constrained by the number of orders that are actually available. As before, we let x_{tad} be the number of times that we apply decision d to a car with attribute a given what we know at time t.

The contribution function is:

c_{tad} = The contribution from assigning a car with attribute a to an order for cars of type $d \in \mathcal{D}^o$, given what we know at time t. If $d \in \mathcal{D}^o$, then we assume that the contribution is a ''reward'' for satisfying a customer order, minus the costs of getting the car to the order. For all other decision classes, the contributions are the (negative) costs from carrying out the decision.

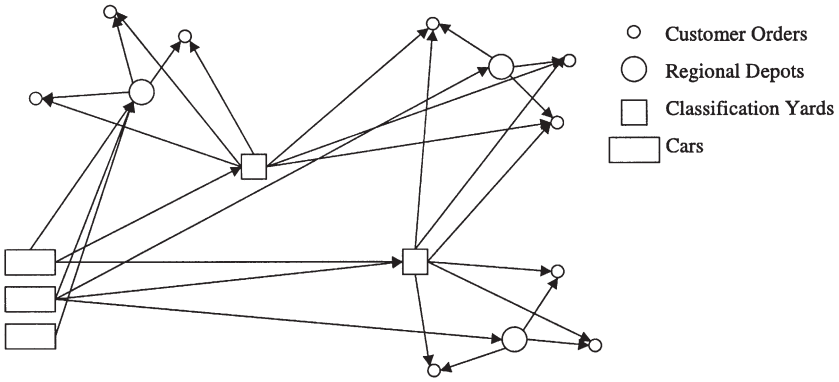

Fig. 3. Car distribution through classification yards.

Since all orders have to be satisfied, it is customary to formulate these models in terms of minimizing costs: the cost of moving a car from its current location to the customer, and the "cost" of assigning a particular type of car to satisfy the order. Since rail costs are extremely complex (what is the marginal cost of moving an additional empty car on a train?), all costs are basically surrogates. The transportation cost could be a time or distance measurement. If we satisfy the customer order with the correct car type, then the car type cost might be zero, with higher costs (basically, penalties) for substituting different car types to satisfy an order. Just the same, we retain our maximization framework because this is more natural as we progress to more general models (where we maximize "profits" rather than minimize costs).

5.3 *Engineering practice—Myopic and deterministic models*

The most basic model used in engineering practice is a myopic model, which means that we only act on the vectors $R^c_{0t'}$ and $R^o_{0t'}$ (we believe that in practice, it is likely that companies even restrict the vector of cars to those that are actionable now, which means R^c_{00}). We only consider decisions based on what we know now (x_{0ad}), and costs that can be computed based on what we know now (c_{0ad}). This produces the following optimization problem:

$$\min_x \sum_{a\in\mathcal{A}} \sum_{d\in\mathcal{D}} c_{0ad} x_{0ad} \tag{5.2}$$

subject to:

$$\sum_{d\in\mathcal{D}} x_{0ad} = R^c_{0a} \quad a \in \mathcal{A} \tag{5.3}$$

$$\sum_{a \in \mathcal{A}} x_{0ad} \leq R^o_{0a_d} \quad d \in \mathcal{D}^o \tag{5.4}$$

$$x_{0ad} \in Z_+. \tag{5.5}$$

Equation (5.4) restricts the total assignment of all car types to a demand type a_d, $d \in \mathcal{D}^o$, by the total known demand for that car type across all actionable times. The model allows a car to be assigned to a demand, even though the car may arrive after the time that the order should have been served. Penalties for late service are assumed to be captured in c_{0ad}.

It is easy to pick this model apart. First, the model will never send a car to a regional depot or classification yard (unless there happens to be a customer order at precisely that location). Second, the model will only send a car to an order that is known. Thus, we would not take a car that otherwise has nothing to do and begin moving to a location which is going to need the car with a high probability. Even worse, the model may move a car to an order which has been booked, when it could have been moved to a much closer location where there probably will be an order (but one has not been booked as yet). If there are more cars than orders, then the model provides almost no guidance as to where cars should be moved in anticipation of future orders.

Amidst these weaknesses are some notable strengths. First, the model is simple to formulate and solve using commercial solvers. Second, the model handles all three types of substitution extremely easily (especially important is substitution across time, something that models often struggle with). But, perhaps the most important feature is that the solution is easy to understand. The most overlooked limitation of more sophisticated models is that their solutions are hard to understand. If the data were perfect, then we would argue that the user should simply trust the model, but the limitations of the data preclude such a casual response.

The first generalization used in practice is to include forecasts of future orders, which we would represent using the vector $R^o_{tt'}$ for $t \in \mathcal{T}^{ph}$, where $\mathcal{T}^{ph}$ is the set of time periods in our planning horizon. The details of the process of forecasting future orders are, of course, not documented. The process of forecasting would generally have to be made at some level of aggregation (daily/weekly, customer level or regional, and the car class). Particularly tricky is handling the time staging of orders. If a forecast is generated for a particular time t' in the future (using, for example, standard time series forecasting techniques applied to a historical dataset showing customer orders by time period), then we would be forecasting the total orders for time t', and then adding in the orders to be satisfied at time t' that are known now. We assume that we have a forecast $R^o_{tt'}$ representing the orders that would be placed at time t to be satisfied at time t'.

We let $R^o_{tt'}$ be a point forecast of future demands for $t \geq 1$, $t' \geq t$, with $R^o_{0t'}$, as before, the orders we know about now. We could also make a forecast of

cars that will become available in the future, but this is still not normally done. As a result, our model using a deterministic forecast is given by:

$$\min_{x} \sum_{a\in\mathcal{A}} \sum_{d\in\mathcal{D}} c_{0ad} x_{0ad} \tag{5.6}$$

subject to:

$$\sum_{d\in\mathcal{D}} x_{0ad} = R^o_{0a} \quad a \in \mathcal{A} \tag{5.7}$$

$$\sum_{a\in\mathcal{A}} x_{0ad} \leq \sum_{t\in\mathcal{T}^{ph}} R^o_{ta_d} \quad d \in \mathcal{D}^o \tag{5.8}$$

$$x_{0ad} \in Z_+ \tag{5.9}$$

Equation (5.8) includes demands that are known now (R^o_{0a}) and all orders that are forecasted to become known within the planning horizon. Note that we are using forecasted orders, but not forecasted cars. One model in industrial practice separately forecasts future cars becoming available, but these forecasts are independent of decisions being made now. To model this process, we would replace equation (5.7) with:

$$\sum_{d\in\mathcal{D}} x_{0ad} = \sum_{t\in\mathcal{T}^{ph}} R^o_{ta} \quad a \in \mathcal{A}$$

It would be possible to use a deterministic, time staged model over a planning horizon, but this would actually be fairly hard to solve, since it would be a moderately large integer multicommodity flow problem with time windows on the loads (it is the time windows that really complicates the formulation).

Models that incorporate forecasted demands have the immediate advantage of providing recommendations for cars which would otherwise not be assigned in a myopic model. The model will send cars to locations which normally book new orders, allowing the railroad to start the process of moving the car, rather than waiting until the last minute. Since we are only using a point forecast, the model will not be able to send cars to a location where they *might* be needed. This can be a problem when we are in a period where there is excess supply. This model can recommend letting cars sit at a location where there is absolutely no chance of them being used, rather than moving them to a location where they might be used.

Our model does not include forecasts of empty cars. The common rationale for leaving these forecasts out is that they are so uncertain (it is not unusual for practitioners to ignore information which cannot be reasonably approximated

by a point forecast). It also ignores many other operational issues such as train capacities (which we have already identified to be highly uncertain), yard capacities (which determines how many cars can be stored at a location) or the value of cars at the end of the horizon (which we could overcome with a multistage model).

There are a number of limitations of these simple models, but we would argue that a serious practical limitation is that the model will never recommend sending a car to a regional depot or classification yard. In one application with which we are familiar, the model will recommend sending a car to a particular customer (perhaps to serve a forecasted order). In the process of routing the car to the customer, the car will have to go through a regional depot. The railroad will then route the car to the classification yard, reoptimizing the assignment of the car to new orders as they become available. This is a highly heuristic way of accounting for uncertainty.

5.4 No substitution—a simple recourse model

Our first effort to incorporate uncertainty is a simple recourse model where we replace the decision class to assign cars to a specific *order* and instead allow us to send cars to a particular *customer* (or equivalently, to a customer location). The difference is that if a customer only places one order, then in the first model we can only send him one car. In our simple recourse model, we may send more cars to the customer location at time t than has been ordered at time t' in the hopes that new orders will come in later. For this case, we define:

$R^o_{t,ct'}$ = The number of orders for customer c that we first learn about at time t that are actionable (must be served) at time t'.
$R^o_{t,c}$ = All the orders for customer c known at time t.
$= (R^o_{t,ct'})_{t' \geq t}$.

Of course, $\mathcal{R}^o_0$ are the orders we know about now, while $(\mathcal{R}^o_t)_{t>0}$ are the forecasted orders for the future. Unlike our first models, we are now going to explicitly model the uncertainty in the forecast of future orders. Looking at equation (5.8), we see that we only need the total forecast of future demands. For this reason, it is simpler to define:

$$\overline{R}^o_{1,c} = \sum_{t \in \mathcal{T}^{ph} \setminus \{0\}} \sum_{t' \in \mathcal{T}^{ph}} R^o_{t,ct'}.$$

$\overline{R}^o_{1c}$ is a random variable representing all "future" demands, which would be derived from a forecasting model. Note that we have aggregated not

only across all orders that would become known in the future (t), but across the dates when the orders would need to be satisfied (t'). Let:

R^c_{0,i_d} = The number of cars (indicated by the superscript c) sent to customer i_d, $d \in \mathcal{D}^c$, where the decision is made at time 0 but the cars can be used at time 1 (the second stage).
$= \sum_{a \in \mathcal{A}} x_{0ad}$.

The decisions x_{0ad} must be made before the orders $(R^o_{ti})_{t>0}$ become known. In our simple recourse model, we assume that a car sent to customer c cannot then, at a later time, be sent to another customer. It is either used to satisfy an order (within our planning horizon) or it sits idle. Let:

c^o_i = The (positive) contribution from satisfying an order for customer $i \in \mathcal{I}^c$.
c^h_i = The contribution (typically negative) from sending a car to customer i and then having it sit.

Now let:

x^o_{1i} = The number of cars assigned to serve an order (after they arrive at customer i).
x^h_{1i} = The number of cars that are held at customer i.

x^o_1 and x^h_1 are random variables defined by:

$$x^o_{1c}\left(R^c_{0,c}, \overline{R}^o_{1c}(\omega)\right) = \min\left\{R^c_{0,c}, \overline{R}^o_{1c}(\omega)\right\}$$
$$x^h_{1c}\left(R^c_{0,c}, \overline{R}^o_{1c}(\omega)\right) = \max\left\{0, R^c_{0,c} - \overline{R}^o_{1c}(\omega)\right\}.$$

We should note that our choices for $x^o_{1c}(R^c_{0,c}, \overline{R}^o_{1c}(\omega))$ and $x^h_{1c}(R^c_{0,c}, \overline{R}^o_{1c}(\omega))$ seem a bit obvious, but they are in fact the result of a trivial optimization problem.

$R^c_{0,c1}$ is a function of the first stage decisions x_0. Given x_0, the expected second stage reward is given by:

$$\begin{aligned}\overline{C}_{0,1}\left(R^c_0(x_0)\right) &= \text{Expected costs using the information available in time period 0 that would be incurred in time period 1.}\\ &= E\left\{\sum_{c \in \mathcal{I}^c}\left(c^o_c x^o_{1c}(R^c_{0c}(x_0), \overline{R}^o_{1c}) + c^h_c x^h_{1c}(R^c_{0c}, \overline{R}^o_{1c})\right)\right\}\\ &= \sum_{c \in \mathcal{I}^c}\left(\overline{C}_{0,c1}(R^c_{0c}(x_0))\right).\end{aligned}$$

The functions $\overline{C}_{0,c1}(R^c_{0c})$ are concave. If the random demands are discrete, then it is also possible to show that $\overline{C}_{0,c1}(R^c_{0c})$ is piecewise linear, concave, with the breakpoints at integer values of R^c_{0c}. Since these functions are computed as expectations of scalar random variables, computing them is quite easy once the distribution of demands is known. Of course, forecasting future demands is in practice fairly tricky, primarily because of the process of customers booking orders in advance.

We can now formulate our problem as follows:

$$\min_{x_0} \left\{ c_0 x_0 + \overline{C}_{0,1}(R^c_0(x_0)) \right\} \tag{5.10}$$

subject to:

$$\sum_{d \in \mathcal{D}} x_{0ad} = R^{+,c}_{0a} \quad a \in \mathcal{A} \tag{5.11}$$

$$\sum_{a \in \mathcal{A}} x_{0ad} \leq R^{+,c}_{0a_d} \quad d \in \mathcal{D}^o \tag{5.12}$$

$$x_{0ad} \in Z_+. \tag{5.13}$$

This is a convex nonlinear programming problem with network constraints. If the demands are discrete, producing piecewise-linear concave reward functions for each shipper, then we can use a standard trick for converting these problems into pure networks, as shown in Fig. 4.

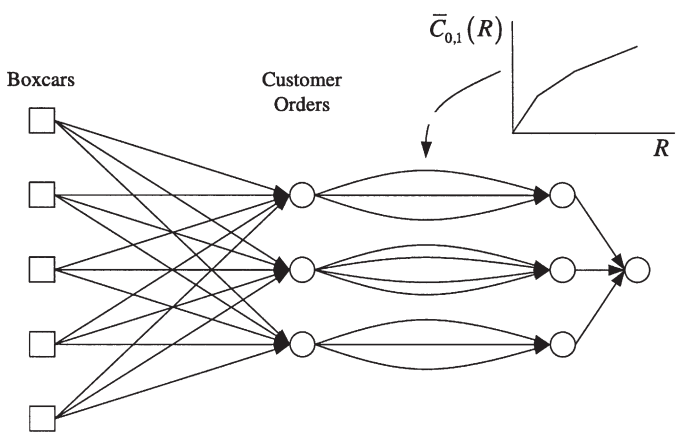

Fig. 4. The simple recourse problem as a pure network.

5.5 Shipping to regional depots—a separable recourse model

The major weakness of the simple recourse model is that it does not capture the ability of the railroad to send cars to a regional depot, and then wait until the last minute to send cars from the depot to the customer. In fact, it is generally not possible to send a car to a customer unless the customer has specifically asked for the car. A more realistic model is to assume that the car has been sent to a local yard (which we refer to as a regional depot) where it is stored waiting for customers.

In this section we present a more general model which captures the ability of a railroad to send cars to a regional depot, after which it can be distributed to customers. We must, however, introduce one key simplification (which we relax later), which is that while we can model general substitution rules between car types and order types in the first stage, we are not going to allow any substitution between car types in the second stage. One way to mitigate this approximation is to aggregate car types into more aggregate categories, and then assume that there is no substitution between major car categories.

We show in this section how we can solve this more general model, producing a solution that requires solving a network with the same structure as that produced for the case of simple recourse (Fig. 4). We begin by assuming that the demands from different shippers are statistically independent, and then present a more general result which uses a technique that we will use for harder problems.

The case of independent demands—an exact result

We begin by setting up some notation that we need for both models. For this work, it is easier to index decisions and contributions by the spatial location. This notation is clearer, although not as general.

For each regional depot there is a set of customers in this region. We represent this using:

$\mathcal{I}_r^c$ = The subset of customers in region $r \in \mathcal{I}^{rd}$. We assume that customers in $\mathcal{I}_r^c$ can only be served using box cars at depot r.

x_{1ri} = The number of cars sent from $r \in \mathcal{I}^{rd}$ to $i \in \mathcal{I}_r^c$ to satisfy customer orders that become known in the second stage (here the destination i plays the role of the decision d in our earlier notation).

c_{1ri} = The contribution from sending cars from $r \in \mathcal{I}^{rd}$ to $i \in \mathcal{I}_r^c$ to satisfy customer orders that become known in the second stage.

R_{1c}^o = Random variable giving the number of orders for customer c in the second stage.

$R_{0r}^c(x_0)$ = Total number of cars sent to region r as a result of decisions made in the first period.

Both x_1 and R_1^o are random variables. For a given realization of the second stage orders, we would find ourselves solving:

$$Q\left(R_{0r}^c, R_{1r}^o(\omega)\right) = \max_{x_1} \sum_{r \in \mathcal{I}^{rd}} \sum_{i \in \mathcal{I}_r^c} c_{1ri} x_{1ri}(\omega) \tag{5.14}$$

subject to:

$$\sum_{i \in \mathcal{I}_r^c} x_{1ri}(\omega) + x_{1rd^\phi}(\omega) = R_{0r}^c \quad \forall r \in \mathcal{I}^{rd} \tag{5.15}$$

$$x_{1ri}(\omega) \leq R_{1i}^o(\omega) \quad \forall i \in \mathcal{I}_r^c, r \in \mathcal{I}^{rd} \tag{5.16}$$

$$x_{1ri}(\omega) \geq 0 \quad \forall i \in \mathcal{I}_r^c, r \in \mathcal{I}^{rd} \tag{5.17}$$

where, as a reminder, d^ϕ is the "do nothing" decision. Problem (5.14)–(5.17) decomposes by regional depot, where the problem for each regional depot is easily solved as a sort. For a given region $r \in \mathcal{I}_r^c$, assume that

$$c_{r1}^o \geq c_{r2}^o \geq \ldots \geq c_{r|\mathcal{I}_r^c|}^o \geq c_r^h$$

where $|\mathcal{I}_r^c|$ is the number of customers in region r. We have ordered the customers so that customer 1 is the most attractive, 2 is the second most attractive, and we have assumed that satisfying any customer is better than doing nothing (this assumption is easy to relax). Clearly, we would like to assign as much capacity as possible to the most valuable customers. We want to find the expectation of $E[Q_r(R_{0r}^c, R_{1r}^o)] = Q_r(R_{0r}^c)$. We are in particular interested in the slopes $Q_R(R_{0r}^c + 1) - Q_r(R_{0r}^c)$, since these form the coefficients on the arcs which give the marginal value of each additional unit of flow. We solve this using the following simple observation. Let $s = R_{0r}^c$, and let $\mathcal{E}(s, i)$ be the event that results in the sth unit of flow being assigned to the ith—most valuable customer. Define:

$$\begin{aligned}\overline{R}_1^o(J) &= \sum_{j=1}^{J} R_{1,j}^o \\ &= \text{Cumulative number of orders made by the top } J \text{ customers.}\end{aligned}$$

The probability of the event $\mathcal{E}(s, J)$, then, is given by:

$$\begin{aligned}Prob[\mathcal{E}(s, J)] &= Prob[(\overline{R}_1^o(J-1) < s) \cap (\overline{R}_1^o(J) \geq s)] \\ &= Prob[\overline{R}_1^o(J-1) < s] + Prob[\overline{R}_1^o(J) \geq s] \\ &\quad - Prob[(\overline{R}_1^o(J-1) < s) \cup (\overline{R}_1^o(J) \geq s)]\end{aligned} \tag{5.18}$$

The events $(R_1^o(J-1) < s)$ and $(\overline{R}_1^o(J) \geq s)$ are collectively exhaustive, so the last probability in equation (5.18) is equal to one. This allows us to reduce (5.18) to:

$$\begin{aligned} Prob\,[\mathcal{E}(s, J)] &= Prob\,[\overline{R}_1^o(J-1) < s] - (1 - Prob\,[\overline{R}_1^o(J) \geq s]) \\ &= Prob[\overline{R}_1^o(J-1) < s] - Prob[\overline{R}_1^o(J) < s]. \end{aligned}$$

Thus, the probability that the sth unit of flow is assigned to the Jth option is simply the difference between two cumulative distributions. These are easy to compute if the demands across customers are independent. Now let $v_r(s)$ be the expected value of the sth unit of flow in depot r, given by:

$$v_r(s) = \sum_{i \in \mathcal{I}_r} c_i^o Prob[\mathcal{E}(s, i)] + c_r^h \left(1 - \sum_{i \in \mathcal{I}_r} Prob[\mathcal{E}(s, i)]\right).$$

The values $v_r(s)$ give the expected marginal value of each additional unit of flow sent into a regional depot.

Using the marginal values $v_r(s)$, our first stage problem is again a pure network very similar to the one used for simple recourse, but now with the property that the decision to send flow to a regional depot is considered explicitly. Our model will now send cars either directly to customers (to serve orders that have already been booked) or to regional depots for later assignment to orders that become known in the future.

Earlier, we considered the problem where we would send cars directly to the customer before knowing the customer demand. We would then incur an overage or underage penalty after learning the outcome. This strategy is referred to as simple recourse. In this section, we send a car to a regional depot; then, after we learn the demand, we decide which customers to allocate cars to. Since we are assigning cars from a single node over several links, this strategy has been referred to as *nodal recourse*.

Our analysis has been simplified in part by the assumption that the demands are independent (making it possible to find the partial cumulative distributions) and to an even greater degree by the assumption that each customer can be served by a single regional depot. We first generalize our analysis to relax the assumption of independent demands, where we use a technique that will also allow us to relax the assumption that each customer is served by a single regional depot.

The general case—Monte Carlo methods

We have seen that in both the simple recourse case and the regional depot (nodal recourse) case, the problem reduces to finding piecewise linear, concave functions characterizing the value of cars at a location. Now we are going to

introduce another technique for estimating these concave functions based on Monte Carlo sampling, which does not require making any independence assumptions between the demands of different customers.

Our second stage problem consists of finding:

$$Q(R_0^c) = EQ(R_0^c, R_1^o). \tag{5.19}$$

Our strategy is to solve this iteratively. At each iteration, we would choose an outcome ω. For this outcome, the conditional second stage function is given by:

$$Q(R_0^c, R_1^o(\omega)) = \max_{x_1(\omega)} \sum_{r \in \mathcal{I}^{rd}} \sum_{i \in \mathcal{I}_r^c} c_{1ri} x_{1ri}(\omega) \tag{5.20}$$

subject to:

$$\sum_{i \in \mathcal{I}_r^c} x_{1ri}(\omega) + x_{1rd^\phi}(\omega) = R_{0r}^c \quad \forall r \in \mathcal{I}^{rd} \tag{5.21}$$

$$x_{1ri}(\omega) \leq R_{1i}^o(\omega) \quad \forall i \in \mathcal{I}_r^c, r \in \mathcal{I}^{rd} \tag{5.22}$$

$$x_{1ri}(\omega) \geq 0 \quad \forall r \in \mathcal{I}^{rd}, i \in \mathcal{I}_r^c. \tag{5.23}$$

Problem (5.20)–(5.23) is pretty easy to solve for a sample realization. Let $\hat{q}_{1r}(\omega)$ be the dual variable for constraint (5.21), reflecting the marginal value of another car. We would like to use this sample gradient information to build an approximation of $Q(R_0^c)$. The simplest strategy, of course, is to build a linear approximation of the form:

$$\hat{Q}(R_0^c) = \hat{q} \cdot R_0^c \tag{5.24}$$

but these are notoriously unstable. Although techniques are available to help these techniques (proximal point strategies, auxiliary functions), we are going to try to build a nonlinear function similar to the exact functions that we have seen so far. The simplest that we have seen starts with a piecewise linear function and then "tilts" it using stochastic subgradients. For example, we could start with any concave function such as:

$$\hat{Q}^0(R) = \rho_0(1 - e^{-\rho_1 R})$$
$$\hat{Q}^0(R) = \ln(R + 1)$$
$$\hat{Q}^0(R) = -\rho_0(R - \rho_1)^2$$

where R is a scalar. As an alternative, we would initialize the function by assuming independence between the demands. Continuous functions can be converted to piecewise linear functions by extrapolating the function between integer values of R. Let $\tilde{q}^n = q(\omega^n)$ be a stochastic subgradient of Q (given by the dual variable of equaion (5.21)), and let R^n be the resource vector at the nth iteration. We can then update our approximation $\hat{Q}$ using the following updating equation:

$$\hat{Q}^{n+1}(R) = \hat{Q}^n(R) + \alpha^n(\tilde{q}^n - \nabla\hat{Q}^n(R^n)) \cdot R. \tag{5.25}$$

This strategy, dubbed the "SHAPE" algorithm, is provably convergent when the function $Q(R)$ (and its approximations $\hat{Q}^n(R)$) are continuously differentiable, but in transportation, we are typically managing discrete resources, and we are interested in integer solutions.

When we are using piecewise linear functions, we can get an even better estimate by using left and right gradients of $Q(R_0^c, R_1^o(\omega))$ rather than a simple subgradient. Let $\tilde{q}^{n+}$ and $\tilde{q}^{n-}$ be the right and left gradients, respectively, of $Q(R_0^c, R_1^o(\omega))$. Then we can perform a two-sided update using:

$$\tilde{Q}^{n+1}(R) = \begin{cases} \tilde{Q}^n(R) + \alpha^n(\tilde{q}^{n+} - \nabla\tilde{Q}^n(R^n)) \cdot R & R \geq R^n \\ \tilde{Q}^n(R) + \alpha^n(\tilde{q}^{n-} - \nabla\tilde{Q}^n(R^n)) \cdot R & R < R^n \end{cases} \tag{5.26}$$

There is another class of strategies that we refer to broadly as structured adaptive functional estimators (or "SAFE" algorithms). In our problem, we are trying to estimate piecewise linear, concave functions which can be represented by a sequence of slopes that are decreasing monotonically. At each iteration, we obtain stochastic gradients that allow us to update estimates of these slopes, but it is important to maintain the concavity of our function or, equivalently, the monotonicity of the slopes. We briefly review two strategies for performing this estimation. The first is referred to as a *leveling* technique since violations of concavity are fixed by leveling the estimates of the slopes (see below). The second is called a separable, projective approximation routine (SPAR), since we maintain monotonicity in the slopes by performing a projection of the updated function onto the space of concave functions.

Both approaches begin by representing the piecewise linear function $Q(R)$ by its slopes as follows. Let:

$$q_r = Q(r+1) - Q(r)$$

be the right derivative of $Q(R)$ at $R = r$. We can the write:

$$Q(R) = Q(0) + \sum_{r=0}^{R-1} q_r.$$

Let $\hat{q}_r^n$ be an estimate of q_r at iteration n. As before, let $\tilde{q}^n$ be a stochastic gradient of Q at iteration n, and assume they have the property that $E[\tilde{q}_r^n] = q_r$. Assume that at iteration n we sample $r = R^n(\omega)$. We could estimate the slopes using the simple updating equations:

$$\hat{q}_r^{n+1} = \begin{cases} (1-\alpha^n)\hat{q}^n + \alpha^n \tilde{q}^n & \text{if } r = R^n(\omega) \\ q_r^n & \text{otherwise} \end{cases} \tag{5.27}$$

If we assume assume that we are going to sample all the slopes infinitely often, then it is not hard to show that $\lim_{n\to\infty} \hat{q}_r^n = q_r$. But this updating scheme would not work in practice since it does not maintain the concavity of the function $\hat{Q}^n(R)$. We know from the concavity of $Q(R)$ that $q_0 \geq q_1 \geq \ldots \geq q_r$. It is apparent that equation (5.27) would not maintain this relationship between the slopes. Within an algorithm, this forces us to solve nonconcave optimization problems, which is quite hard. We note that concavity is automatically maintained in equation (5.25) since we are updating a concave approximation with a linear updating term. Concavity is also maintained in equation (5.26), since we are guaranteed that $\tilde{q}^{n-} \geq \tilde{q}^{n+}$.

The first of our two approaches maintains concavity (monotonicity in the slopes) by using a technique that we call *leveling*. Here, all we are doing is identifying a violation of concavity after a basic update (as in equation (5.27)), and then adjusting the neighbors of the updated slope so that concavity is maintained. As before, let $R^n(\omega)$ be the point that we sample in iteration n. The updating equations are given by (see Fig. 5):

$$\hat{q}_r^{n+1} = \begin{cases} \alpha^n \tilde{q}^n(\omega) + (1-\alpha^n)\hat{q}_r^n & \text{if } R^n(\omega) = r \\ \alpha^n \tilde{q}^n(\omega) + (1-\alpha^n)\hat{q}_i^n & \text{if } R^n(\omega) = i < r \text{ and } \alpha^n \tilde{q}^n(\omega) + (1-\alpha^n)\hat{q}_i^n < \hat{q}_r^n \\ \alpha^n \tilde{q}^n(\omega) + (1-\alpha^n)\hat{q}_i^n & \text{if } R^n(\omega) = i > r \text{ and } \alpha^n \tilde{q}^n(\omega) + (1-\alpha^n)\hat{q}_i^n > \hat{q}_r^n \\ \hat{q}_r^n & \text{otherwise} \end{cases} \tag{5.28}$$

The second method starts with the estimate of the slopes given by equation (5.27) and then performs a projection onto the space of functions whose slopes are monotonically decreasing. We start by letting the left hand side of equation (5.27) be denoted by the vector $\overline{q}^{n+1}$ which clearly may violate concavity. We can now think of $\mathcal{Q}$ as the space of concave functions, and let

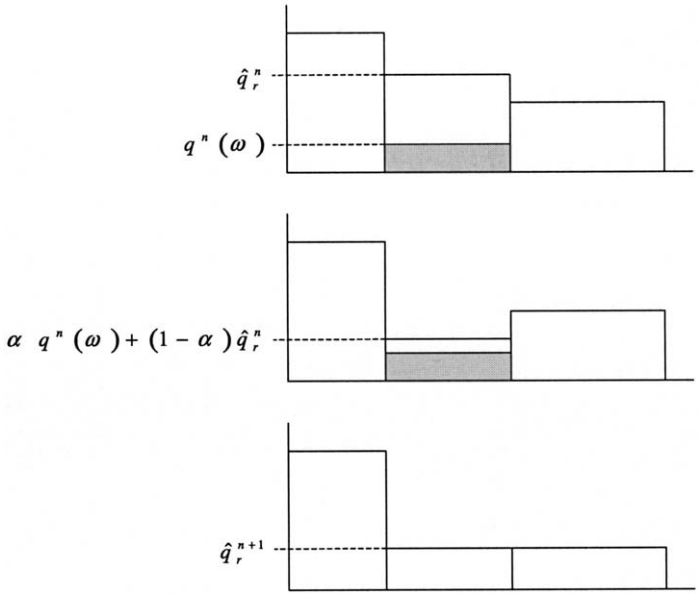

Fig. 5. Maintaining concavity by the levelling method.

$\Pi_{\mathcal{Q}}$ be the nearest point projection onto the space $\mathcal{Q}$. This allows us to represent the process of converting the vector $\overline{q}^{n+1}$ as the projection:

$$\hat{q}^{n+1} = \Pi_{\mathcal{Q}}(\overline{q}^{n+1}) \tag{5.29}$$

The projection $\Pi_{\mathcal{Q}}$ is the solution to the quadratic programming problem:

$$\min_{q} \|q - \overline{q}^n\|^2 \tag{5.30}$$

subject to:

$$q_{r+1} - q_r \leq 0 \tag{5.31}$$

Solving this projection problem is very easy. Assume that after the basic update, we have an instance where $\overline{q}_{r-1}^n < \overline{q}_r^n$. Let $\bar{r} = \arg\min_{r' < r}\{\overline{q}_{r'}^n < \overline{q}_r^n\}$ be the smallest index such that $\overline{q}^n < \overline{q}_r^n$. Now find the average over all these elements:

$$\overline{q}_{[\bar{r},\, r]}^n = \frac{1}{r - \bar{r} + 1} \sum_{r'=\bar{r}}^{r} \overline{q}_{r'}^n$$

Finally, we let

$$\hat{q}_{r'}^{n+1} = \begin{cases} \overline{q}_{[\bar{r},r]}^{n} & \text{if } \bar{r} \geq r' \geq r \\ \overline{q}_{r}^{n} & \text{otherwise} \end{cases}$$

Both the leveling method and the projection method produce convergent algorithms from two perspectives. First, if all the slopes are sampled infinitely often, then we obtain that $\lim_{n\to\infty} \hat{q}_r^n = q_r$ for all r a.s. But, we are not going to sample all the slopes infinitely often. What we want to do is to use the approximation $\hat{Q}^n$ as an approximation of the second stage to determine the solution to the first stage. Thus, our algorithm is going to proceed by solving:

$$x_0^n = \arg\max_{x_0} c_0 x_0 + \hat{Q}^n(R_0(x_0)) \tag{5.32}$$

subject to our first stage constraints:

$$\sum_{d\in\mathcal{D}} x_{0ad} = R_{0,a}^{+,c} \quad a \in \mathcal{A} \tag{5.33}$$

$$\sum_{a\in\mathcal{A}} x_{0ad}\delta_{1,a'}(0, a, d) = R_{0,a'}^{c} \quad a' \in \mathcal{A} \tag{5.34}$$

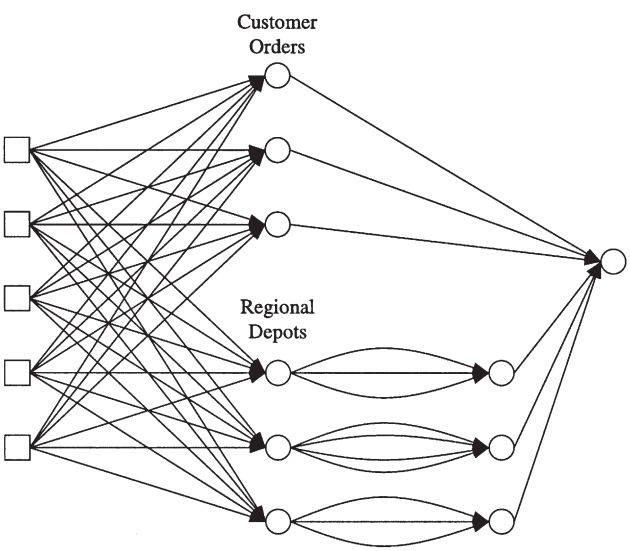

Fig. 6. The separable recourse problem as a pure network.

$$\sum_{a \in \mathcal{A}} x_{0ad} \leq R^{+,c}_{0,a_d} \quad d \in \mathcal{D}^o \tag{5.35}$$

$$x_{0ad} \in Z_+ \tag{5.36}$$

The problem (5.32)–(5.36) is a pure network shown in Fig. 6. Once we obtain x^n_0, we find a sample realization ω^n and solve the optimization problem in (5.20) again. The duals from this problem are used to update the value function, and the process repeats itself.

Our algorithm, then, does not sample the entire domain for R_0, but rather only those that are produced by solving our first stage approximation. Fortunately we can show that this algorithm will visit the optimal solution infinitely often. A side benefit is that we are solving sequences of pure networks which readily yield integer solutions. Integrality can be a major headache in transportation applications, but we have now designed an algorithm which always produces integer solutions. Of central importance in this regard is the fact that our algorithm never performs smoothing on the decision variables, as would be required if we used stochastic linearization methods.

At this point it may seem that we are simply solving very special cases. In fact, as we soon show, we actually have all the machinery we need to solve very general instances of this problem.

5.6 Shipping to classification yards—a network recourse model

The next level of generalization is the challenge of working with what we call "classification yards." For the purpose of this presentation, we are going to assume that we can send cars from classification yards to any customers in the network (for the moment, we are not going to allow ourselves to send cars to regional depots, since this would take us past our basic two-stage model). For the moment, we are going to continue to assume that once cars reach a classification yard that there is no substitution between car types: if a customer order is for car type k, then we must provide car type k. But we are going to assume that a single customer can be served from more than one depot.

This problem is known to the stochastic programming community as a two-stage stochastic program with network recourse. The reason is that, unlike our previous models, the second stage is now a general network problem (as opposed to the much simpler problems posed by simple or nodal recourse). Solving a network problem in the second stage is almost as difficult as solving a general linear program, which means that we should consider algorithms designed for general two-stage stochastic linear programs.

The research community has developed a number of algorithmic strategies over the years. The question of whether an algorithmic strategy works has to be answered in three levels: (1) Does the algorithm appear to work in theory? Does it capture the mathematical properties of the problem? (2) Does it

produce reasonable numerical results in laboratory experiments? For example, using datasets reflecting specific classes of problems, we would like to know if it converges quickly, producing stable, high quality solutions. (3) Does it work in practice, producing recommendations that are acceptable to those who have to implement them?

As of the writing of this chapter, there is a strong handle on the theory, but numerical testing is extremely limited (given the broad diversity of problems). For example, showing that an algorithm works well on car distribution problems for one railroad will not always convince another railroad that it will work on their network! Container management problems (in trucking, rail and intermodal applications) come in a variety of sizes and characteristics. The dynamics of short-haul regional truckload carriers are completely different from those of long-haul national carriers. Experiments in pure transportation applications do not tell us whether it would work in other resource allocation settings such as supply chain management and distribution problems. And we are not even talking about applications outside of transportation and logistics. In short, each subcommunity (and these can be very specialized) needs to see numerical work to demonstrate effectiveness on its own problem class.

Given the very limited amount of laboratory testing of the different algorithmic strategies (even within general transportation problems), our discussion focuses on the qualities of different algorithms and their potential strengths and weaknesses for our problem. We cannot definitively state what will and will not work for our problem class, but we can discuss the qualities of different approaches. In particular, we are interested in the degree to which a technique allows us to exploit the underlying structure of the transportation problem. Many transportation problems require integer solutions, and also exhibit near-network structure. Algorithms which allow us to exploit this network structure are more likely to yield integer solutions from LP relaxations, or at least provide tight LP relaxations.

Scenario methods

Perhaps the best known algorithmic strategy in stochastic programming is *scenario programming*, popular because of its conceptual simplicity, generality, and use of general-purpose optimization algorithms. But, its effectiveness for transportation applications is doubtful.

Let $\hat{\Omega}$ be a (not too large) sample of outcomes (future car orders, future car supplies, as well as travel times). Further let $\hat{p}(\omega)$ be the probability of outcome $\omega \in \hat{\Omega}$. We can approximate our original problem using the method of scenarios:

$$\max_{x_0, x_1} c_0 x_0 + \sum_{\omega \in \hat{\Omega}} \hat{p}(\omega) c_1 x_1(\omega) \tag{5.37}$$

subject to:

First-stage constraints:

$$\sum_{d\in\mathcal{D}} x_{0ad} = R^c_{0a} \quad a \in \mathcal{A} \tag{5.38}$$

$$\sum_{a\in\mathcal{A}} x_{0ad} = R^c_{0a_d} \quad d \in \mathcal{D}^o \tag{5.39}$$

$$\sum_{a\in\mathcal{A}}\sum_{d\in\mathcal{D}} x_{0ad}\delta_{1,a'}(0,a,d) - R^c_{0a'1} = 0 \quad a' \in \mathcal{A} \tag{5.40}$$

$$x_{0ad} \in Z_+ \quad a \in \mathcal{A}, d \in \mathcal{D}. \tag{5.41}$$

Second stage constraints:

$$\sum_{d\in\mathcal{D}^c_a} x_{1ad}(\omega) + x_{1ad^\phi}(\omega) = R^c_{0,a} \quad \forall a \in \mathcal{A}, \quad \forall \omega \in \hat{\Omega} \tag{5.42}$$

$$\sum_{a\in\mathcal{A}} x_{1ad}(\omega) \le R^o_{1c_d}(\omega) \quad d \in \mathcal{D}^c, \quad \forall \omega \in \hat{\Omega} \tag{5.43}$$

$$x_{1ad}(\omega) \ge 0 \quad \forall a \in \mathcal{A}, d \in \mathcal{D}^c_a, \forall \omega \in \hat{\Omega}. \tag{5.44}$$

We note that our decision class $\mathcal{D}$ only allows us to assign cars to a known order, or to reposition cars to a general or regional depot. As a result, the second stage problem is primarily one of assigning cars from the general and regional depots to orders that became known in the second stage.

Scenario methods have been very popular in financial applications, but we feel that there are specific characteristics of financial applications that are not shared in transportation applications, and vice versa. Financial applications are characterized by very complex stochastic processes with high levels of interdependence reflecting the dependence of random outcomes on a relatively smaller number of common factors. It is easier, then, to reasonably approximate the future with a smaller number of scenarios. Also, financial applications typically do not have integer variables.

Transportation applications, on the other hand, are characterized by a large number of relatively independent random variables. The optimization problems, which are typically integer, are often so large that deterministic problems are hard (although they often have embedded network structures). The formulation in (5.37)–(5.44) has the effect of taking a computationally intractable problem and blowing it up into a problem that is many times larger. Furthermore, many transportation problems exhibit a natural network structure that is destroyed by the representation of the second stage problem.

Benders decomposition

Benders decomposition is an appealing algorithm that replaces the very large problems posed in scenario optimization with sequences of relatively small problems of the form (which we state here as a minimization problem as is standard practice in the literature):

$$\min_{x_0} c_0 x_0 + z \tag{5.45}$$

subject to the first stage constraints (5.38)–(5.41) which we represent compactly using:

$$A_0 x_0 = R_0 \tag{5.46}$$

$$x_0 \geq 0 \tag{5.47}$$

and the constraints:

$$z - \beta_i x_0 \geq \alpha_i, \quad \forall \quad i = 1, \ldots, n \tag{5.48}$$

where β_i and α_i are generated by solving the dual of the second stage problem, which for compactness we can write as:

$$\min_{x_1} c_1 x_1(\omega)$$

subject to:

$$A_1 x_1(\omega) = R_1(\omega) + B_0 x_0$$
$$x_1(\omega) \geq 0.$$

Different algorithms have been proposed for generating cuts. The first algorithm of this class is the so-called "L-shaped" decomposition, which works on a finite set of outcomes (which cannot be too large, since we have to solve a linear program for each outcome). This concept was generalized by the stochastic decomposition algorithm which generates cuts from a potentially infinite sample space. A sketch of the algorithm is given in Fig. 7. This algorithm converges almost surely to the optimal solution, but the rate of convergence on practical applications remains an open question.

The CUPPS algorithm (outlined in Fig. 8) requires a finite sample space which can be quite large (for example, thousands or tens of thousands of scenarios). The critical step in the stochastic decomposition is equation (5.50) which requires smoothing on the coefficients of the cuts. The critical step in the CUPPS algorithm is equation (5.51) which requires a simple

Step 1. Solve the following *master problem*:

$$x_0^n = \arg\min\{c_0 x_0 + z : \ A_0 x_0 = R_0,\ z - \beta_t^n x \geq \alpha_t^n,\ t = 1, \ \ldots, n-1,\ x \geq 0\}$$

Step 2. Sample $\omega^n \in \Omega$ and solve the following *subproblem*:

$$\min\{c_1 x_1 : \ A_1 x_1 = R_1(\omega^n) + B_0 x_0^n,\ x_1 \geq 0\}$$

to obtain the optimal dual solution:

$$v(x_0^n, \omega^n) = \arg\min_v\{(R_1(\omega^n) + B_0 x_0^n)\ v : \ A_1^T v \leq c_1\}$$

Augment the set of dual vertices by:

$$\mathcal{V}^n = \mathcal{V}^{n-1} \bigcup \{v(x_0^n, \omega^n)\}$$

Step 3. Set:

$$v_t^n = \arg\max\{(R_1(\omega^t) + B_0 x_0^n)v : \ v \in \mathcal{V}^n\} \text{ for all } t = 1, \ \ldots, n$$

Step 4. Construct the coefficients of the nth cut to be added to the master problem by:

$$\alpha_n^n + \beta_n^n x_0 \equiv \frac{1}{n} \sum_{k=1}^{n} (R_1(\omega^k) + B_0 x_0) v_k^n \tag{5.49}$$

Step 5. Update the previously generated cuts by:

$$\alpha_k^n = \frac{n-1}{n}\ \alpha_k^{n-1},\ \beta_k^n = \frac{n-1}{n}\ \beta_k^{n-1},\ k = 1, \ \ldots, n-1 \tag{5.50}$$

Fig. 7. Sketch of the stochastic decomposition algorithm.

arithmetic calculation over the entire sample space. Since equation (5.51) is quite simple, it is not hard to execute even for tens of thousands of scenarios, but it prevents the algorithm from ever being applied rigorously to complete sample spaces (for realistic problems) which can be of the order 10^{10} or even 10^{100}. From a practical perspective, it is not clear if this is useful.

We need to keep in mind that Benders decomposition is probably limited (in transportation applications) to the types of resource allocation problems that we have been considering (since these can be reasonably approximated as continuous linear programs). However, there are unanswered experimental questions even for this special problem class. First, there is the usual issue of rate of convergence. Real car distribution problems may have over 100 regional depots and thousands of customers (for our model, it is the number of regional depots that really impacts the second stage problem). If there are 50 car types (a conservative estimate) and 100 depots, then (realizing that we do not have all car types at all locations) we can still anticipate upwards of a

Step 1. Solve the following *master problem*:

$$x_0^n = \arg\min\{cx + z : \ A_0 x_0 = R_0,\ z - \beta_k^n x \geq \alpha_k^n,\ k = 1, \ \ldots, n-1,\ x \geq 0\}$$

Step 2. Sample $\omega^n \in \Omega$ and solve the following dual *subproblem*:

$$v(x^n, \omega^n) = \arg\min\{(R^o(\omega^n) + B_0 x_0^n)v : \ A_1^T v \leq c_1\}$$

Augment the set of dual vertices by:

$$V^n = V^{n-1} \bigcup \{v(x^n, \omega^n)\}$$

Step 3. Set:

$$v^n(\omega) = \arg\max\{(R^o(\omega) + B_0 x_0^n)v : v \in V^n\} \quad \text{for all } \omega \in \Omega \qquad (5.51)$$

Step 4. Construct the coefficients of the nth cut to be added to the master problem by:

$$\alpha_n^n + \beta_n^n x_0 \equiv \sum_{\omega \in \Omega} p(\omega)(R_1(\omega) + B_0 x_0)v^n(\omega)$$

Fig. 8. Sketch of the CUPPS algorithm.

thousand resource states for the second stage problem. How quickly does Benders decomposition converge for problems of this size? The problem is that a single cut may not improve our approximation of the value of cars in a particular location.

A second issue with Benders decomposition is that real applications require integer solutions. When the flows are relatively large, solutions are easily rounded to obtain reasonable approximations to the discrete version of the problem. In actual applications, there can be many instances where flows (to small locations, or of unusual car types) are quite sparse and fractional solutions become more problematic. In these cases, fractional solutions can involve flows that are less than one, and simple rounding may produce infeasible solutions. At a minimum, dealing with fractional solutions can be a nuisance.

Despite these questions, Benders decomposition is a promising technique that needs to be tested in the context of specific applications.

Stochastic linearization techniques

A powerful set of techniques is based on linear approximations of the recourse function. These fall into two groups. The first uses pure linear approximations, but performs smoothing on the first stage variables to stabilize the solution. The second group introduces some sort of nonlinear stabilization term.

The pure linearization strategy solves sequences of problems of the form:

$$\hat{x}_0^n = \arg\min_{x_0} c_0 x_0 + \overline{q}^{n-1} \cdot R_1(x_0). \tag{5.52}$$

To calculate $\overline{q}^n$, let $\tilde{q}^n$ as before be our stochastic gradient obtained as the dual variable of the resource constraint from the second stage using a Monte Carlo realization ω^n. We then smooth these gradients to obtain:

$$\overline{q}^n = (1-\alpha^n)\overline{q}^{n-1} + \alpha^n \tilde{q}^n. \tag{5.53}$$

Having obtained the solution $\hat{x}_1^n$, we then smooth this as well:

$$\overline{x}_0^n = (1-\beta^n)\overline{x}_0^{n-1} + \beta^n \hat{x}_0^n. \tag{5.54}$$

For both equations (5.53) and (5.54), we would normally require the usual conditions on the stepsizes for stochastic problems, namely that $\sum_{n=0}^{\infty} \alpha^n = \infty$ and $\sum_{n=0}^{\infty} (\alpha^n)^2 < \infty$, although we note in passing that this is not normally satisfied by the stepsize rules used in practice.

Stochastic linearization techniques are clearly the simplest to use, but are unlikely to work in practice simply because of the lack of stability. Stability is imposed on the solution primarily through the use of declining stepsizes, but this is an artificial form of stability. Furthermore, the smoothing on the first stage variables performed in equation (5.54) is a serious practical problem because it completely destroys integrality. Rounding is hard for large problems because it can be difficult ensuring conservation of flow in the presence of substitutable resources.

Despite these weaknesses, techniques based on linear approximations are attractive for transportation applications since they retain the structure of the original problem. If the first stage is a network problem, then adding a linear adjustment term retains this property. Linear approximations are also easy to compute and store. We may be able to overcome the instability of pure linearization techniques by employing any of a variety of nonlinear stabilization terms. One notable example is proximal point algorithms, which solve sequences of problems of the form:

$$x_0^n = \arg\min_{x_0} c_0 x_0 + \overline{q}^{n-1} \cdot R_1(x_0) + \theta\psi(x_0, \overline{x}_0^{n-1})$$

where $\psi(x, \overline{x}^{n-1})$ is a distance metric such as $\psi(x, \overline{x}^{n-1}) = \|x - \overline{x}^{n-1}\|^2$. $\overline{x}^n$ is computed using:

$$\overline{x}^n = (1-\beta^n)\overline{x}^{n-1} + \beta^n x^n.$$

Note that at the last iteration, the final solution is x^n, not $\overline{x}^n$. $\overline{x}^n$ is used only to stabilize the solution.

In engineering practice, we can be creative in the construction of the distance metric $\psi(x, \overline{x}^{n-1})$. If it is separable in x, then we may find ourselves solving a separable, convex optimization problem. If we are particularly interested in integer solutions, we can construct a piecewise linear function defined on discrete values of x (even if $\overline{x}^{n-1}$ is fractional).

A second type of nonlinear stabilization strategy is the SHAPE algorithm first presented in Section 5.5.2. This is a type of auxiliary function algorithm where we start with an artificial auxiliary function $\hat{Q}^0(R)$, and then update it using stochastic gradients as demonstrated in equation (5.25). An attractive feature of this algorithm is that the auxiliary function can (and should) be chosen to retain the structure of the first stage problem as much as possible. For our car distribution problem (and similar applications), piecewise linear, separable approximations are particularly attractive.

Nonlinear functional approximations

Our last class of strategies tries to explicitly approximate the recourse function, without any guarantee of convergence to the exact function. We do not include algorithms such as the one-sided SHAPE algorithm (equation (5.25)) in this group because there is no explicit attempt to approximate the recourse function. We also do not include Benders decomposition simply because we feel that this strategy is in a class by itself. But we do include both the two-sided SHAPE algorithm (equation (5.26)) and the structured adaptive functional estimators.

We first introduced these algorithms in the context of a recourse function $Q(R)$ which could be written as a separable function of the resource vector R. Now consider what would happen if we apply the exact same algorithms to a nonseparable problem. We still produce a separable approximation of Q, and we still solve sequences of networks that are identical to Fig. 5. The important difference is that we are solving sequences of separable approximations of otherwise nonseparable functions. For continuously differentiable problems, this can be an optimal strategy (in the limit). For nondifferentiable problems (as we are considering) the result is an algorithm that is very near optimal with a much faster rate of convergence than has been achieved using Benders decomposition.

Extension to substitution across car types

In the previous section, we retained our assumption that there was no substitution between car types in the second stage. However, the substitution between classification yards (spatial substitution) produced a problem with network recourse (we had to solve a transshipment problem in the second stage). Now consider what happens if we allow substitution between car types, which produces the second stage network illustrated in Fig. 9. We quickly see that while this expands the set of options for a car, it is still a network, and is

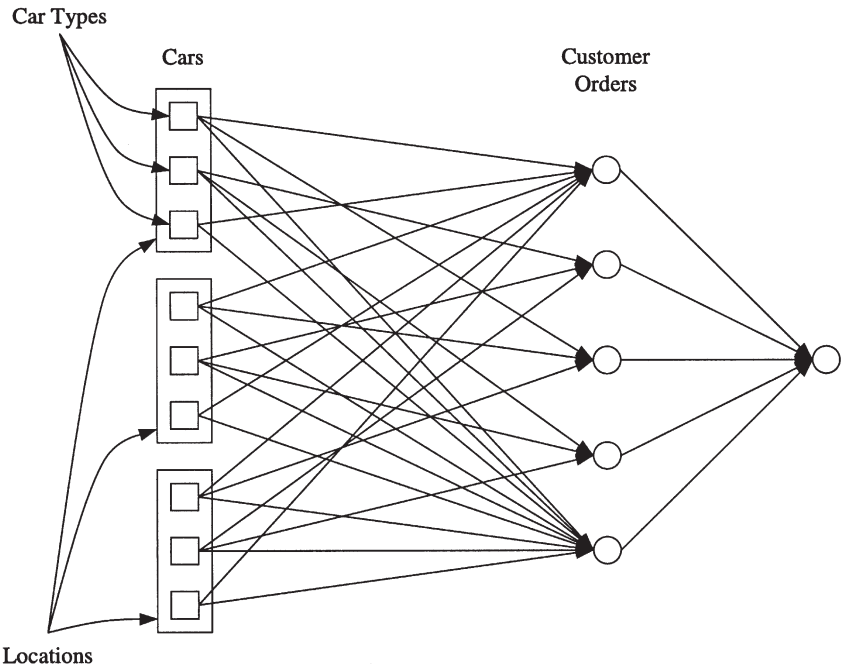

Fig. 9. Illustration of second stage substitution between car types and locations.

mathematically equivalent to the problem which only allows spatial substitution. In addition, we are also implicitly allowing temporal substitution. In the second stage, we will forecast demands that are actionable on different days, as well as cars that will become available on different days. However, we are allowing general assignments of cars to demands that may require cars to be held to meet the demand, or require demands to wait until the car arrives.

We see, then, that we can use the same algorithmic strategy for handling substitution between car types as we did for geographic substitution. But this avoids the more practical question: will it work? It is much more convincing to argue that spatial problems will produce approximately separable recourse functions than would arise in the case with other forms of substitution. For example, it is quite likely that the cost of substituting other box cars in the same car group is quite small. In fact, this is the reason that a railroad might reasonably ignore car subgroups and just model car groups.

For two-stage problems, there is a strong reason to believe that separable approximations will work well, even when the recourse function is not even approximately separable. The logic behind this argument is that for n sufficiently large, $\hat{Q}^n$ will stabilize, and therefore so will x_0^n. As x_0^n (approximately) approaches a limit point, $R_1^n(x_0^n)$ will approach a limit point, allowing us to produce an accurate (but not perfect) piecewise linear approximation

$\hat{Q}_j^n(R_j)$ of the jth dimension of $Q(R)$, at the point $R = R_1^n(x_0^n)$. Thus, a separable approximation only needs to be a good local approximation to provide good results.

Summary

The choice of the best algorithm for two-stage resource allocation problems remains an open question. Two-stage problems are an important foundational problem, but in transportation applications, the usual goal is to solve multistage problems. However, it is important to study the performance of an algorithm in a two-stage setting first, and it is difficult to believe that an algorithm that does not work well for two-stage problems would turn out to work well in a multistage application. But the converse may not be true; an algorithm that does work well for a two-stage problem may not work well in a multistage setting. It is possible for an algorithm (such as the separable, nonlinear functional approximations) to exploit the limiting behavior of the first stage decisions, a property that we lose immediately in the context of multistage problems.

Our belief is that while scenario methods are unlikely to prove attractive in practice, the other three major classes of techniques (Benders decomposition, stochastic linearization with nonlinear stabilization strategies, and nonlinear functional approximations) all deserve serious study. We believe that nonlinear functional approximations are going to work the best for two-stage problems because: (a) they attempt to explicitly approximate the recourse function; and (b) they exploit the structure of problems that arise in transportation. However, we have not yet addressed some of the more difficult dimensions of transportation applications (some of which are touched on below). Nonlinear approximations can work the best in laboratory experiments, but are much more difficult to use than linear approximations. Pure linear approximations are too unstable, but linear approximations with nonlinear stabilization terms (proximal point algorithms or auxiliary functions) may offer an attractive alternative. All of these approximations are separable, and it is simply not clear how well these will work in multistage applications.

5.7 Extension to large attribute spaces

Up to now, we have considered a problem where there are $|\mathcal{K}|$ car types spread among $|\mathcal{I}|$ locations. Using our earlier notation, the attribute vector of a car would be represented by $a = (k, i)$. Thus, our attribute space $\mathcal{A}$ has $|\mathcal{A}| = |\mathcal{K}| \times |\mathcal{I}|$ elements. A large railroad might have several hundred regional depots and 50 to 100 car types, with a total of several thousand combinations. Large, but not hard to enumerate on modern computers.

Now consider what a real car distribution problem looks like. While there are, in fact, 50 to 100 real car types, these cars are allocated among

several dozen *pools* which represent groups of cars controlled by individual shippers. The allocation of cars to pools is negotiated periodically between the shipper and the railroad, and cars can be moved from one pool to another, reflecting the evolving needs of the shippers. Cars are also characterized by the railroad that owns them (when they are not in a pool). Box cars sometimes carry dirty freight that limits their ability to carry cargo (such as food) that requires a clean car. For this reason, cars carry the attribute of the commodity type that they last carried. Finally, some cars may have equipment problems that require maintenance. These problems may range from minor issues that do not affect the use of the car to more serious equipment problems.

When we use this full vector of attributes, we now find that there are not several thousand possible attributes, but several million. Now we cannot even generate the entire attribute space. This creates an interesting problem. In our optimization model, we may wish to consider acting on a car with what is now a multiattribute vector a with decision d, producing a car with attribute vector a'. Since we are not able to enumerate the space $\mathcal{A}$, we may not have an approximation for $\hat{Q}^n_{a'}(R_{a'})$. As a result, we have to devise a strategy for approximating $\hat{Q}^n_{a'}(R_{a'})$.

We are not able to address this issue in depth, but it is important to understand some of the problems that arise. First, it is easy to see that we should try to make sure that our initial approximation $\mathrm{d}\hat{Q}^0_{a'}(R_{a'})/\mathrm{d}R_{a'}|_{R_{a'}=0}$ is an optimistic estimate of $\partial Q(R)/\partial R_{a'}|_{R=0}$. If we did not do this, then a low estimate might result in us choosing not to make the decision d that produces a resource with attribute a'. Since we never visit that resource state, we never improve our approximation.

In practice, the use of optimistic estimates of the value of a resource may not work. We have found that initial approximations that are guaranteed to be optimistic are actually too optimistic. Consider choosing between decisions d' and d''. Assume that decision d' produces a resource with attribute a' while decision d'' produces a resource with attribute a''. Further assume that we have generated the attribute a' (and therefore have an approximation $\hat{Q}^n_{a'}(R_{a'})$), but we have never generated the attribute a''. If we use an optimistic approximation for $\hat{Q}^n_{a''}(R_{a''})$, then we would choose d'' just because we have never tried it before. The result is the steady exploration of every possible decision, and a virtual enumeration of the attribute space $\mathcal{A}$.

A more practical approach is to assume that we have access to an aggregation function $G(a) \mapsto \hat{a}$ where $\hat{a} \in \hat{\mathcal{A}}$ is an aggregation of the original attribute space. We assume that $\hat{\mathcal{A}}$ is not too large and can be enumerated. We further assume that $\hat{Q}^n_{\hat{a}}(R_{\hat{a}})$ is a "good" approximation of $\hat{Q}^n_a(R_a)$ when $G(a) = \hat{a}$. We then make sure that we repeatedly sample gradients and update $\hat{Q}^n_{\hat{a}}(R_{\hat{a}})$ for all $\hat{a} \in \hat{\mathcal{A}}$.

We are not aware of any formal convergence theory for this approach (or for any other algorithm for this problem class). But real problems in

transportation are characterized by a much richer vector of attributes than is normally considered by the academic community.

6 Multistage resource allocation problems

We now turn to the challenge of solving multistage problems. In multistage problems, we have to capture the sequential decision-making process as information (and decisions) evolve over time. For transportation problems, we encounter the reusability of resources; once a vehicle moves a load, it is available to be used again.

We could motivate our multistage problem using our rail car distribution example, but it is useful to bring other applications into the picture. Examples include:

1) Fleet management for truckload trucking—In truckload trucking, a truck moves an entire load of freight from origin to destination. Uncertainty plays a major role in the long haul truckload market, where loads can take between 1 and 4 days to deliver. Customers sometimes request trucks the same day the order is made, but more often call in one or 2 days in advance. In sharp contrast with the rail industry, the truckload carrier does not have to accept every load, and this is one of the major challenges. Emerging electronic market places, where loads are posted on web sites, open the possibility of taking loads that do not have to move for several days. This is a classic problem of decision-making under uncertainty.
2) Driver management for long-haul less-than-truckload motor carriers—LTL carriers face the problem of timing the movement of loads over the network, requiring the careful management of drivers.
3) Management of jets in the fractional ownership industry—In this business, high net worth individuals and business executives will own a fraction of a jet. This gives them access to the entire fleet of jets. They may call the company with as little as 8 h notice and request that a jet move them from a local airport to any other airport. After the move, the fleet operator will move the jet to another location.
4) Routing and scheduling transport aircraft for the air mobility command—The AMC works like a large trucking company, moving freight for the military using large transport aircraft. They are typically used in support of emergency situations where requests for freight movement arise dynamically.

Compared to our rail car distribution problem, these applications are characterized by relatively shorter travel times and less flexibility to satisfy customer orders at times other than when they were requested.

Multistage problems are, of course, much harder than two-stage problems, but we are going to approach them by building on the tools we have

already introduced. Our strategy for solving multistage problems is to solve them as sequences of two-stage problems. Students of dynamic programming will see strong similarities in our modeling approach. But multistage problems do introduce a fresh set of modeling and algorithmic issues that simply do not arise in two-stage problems.

We start in Section 6.1 with a formulation of the problem. Then, Section 6.2 outlines the general algorithmic strategy. Section 6.3 describes the implementation in the context of single commodity flow problems. Section 6.4 outlines the challenges that arise when we solve multicommodity problems in a multistage setting. Finally, Section 6.5 describes the complications that are introduced when we model the property that it takes more than one time period to go from one location to another.

6.1 Formulation

We present the basic multistage problem with somewhat more generality than we have used previously. We first define the exogenous information arriving to our system:

$\hat{R}_t$ = Vector of new arrivals in period t, where $\hat{R}_t = (\hat{R}^o_t, \hat{R}^c_t)$.
ξ_t = Complete vector of new information arriving in period t, including both $\hat{R}_t$ as well as other information about system parameters (travel times, costs, and parameters governing the physics of the problem).

For our purposes, we are only interested in the resource state R_t, and the only information process we are modeling at the moment is the arrival of new resources, $\hat{R}_t$. Using this notation, our history of states, information and decisions (given earlier in equation (5.1)) would look like:

$$h_t = \{R^+_0, x_0, R_0, \hat{R}_1, R^+_1, x_1, R_1, \hat{R}_2, R^+_2, x_2, R_2, \ldots, \hat{R}_t, R^+_t, x_t, R_t, \ldots\} \tag{6.1}$$

There are three perspectives of the state of our system:

R_t = Vector of resources available in period t after decisions x_t have been made.
K_t = What is known at time t after the new information ξ_t has been incorporated. K_t includes R_t plus what we know about parameters that govern the dynamics of the system.
I_t = Set of information available at time t for making a decision. I_t includes K_t, but it might also include forecasts of future activities (activities which are not "known" now, but are the result of a forecasting exercise).

Our process is controlled by the decisions we make:

x_t = Decisions which are made after new information in period t has become known.
$X_t^\pi(I_t)$ = The decision function of policy π.

Our decisions are chosen to maximize the expected total contribution over a planning horizon. Our contribution function is expressed as:

$c_t(x_t, K_t)$ = The contribution generated in period t given decision x_t, and what is known, K_t.

When resources are allocated in time t, they have to satisfy flow conservation equations of the form:

$$\sum_{d\in\mathcal{D}} x_{tad} = R_{t-1,at} + \hat{R}_{t,at}$$

where we assume that we can only act on resources that are actionable now (R_{tt}). The physical dynamics of the system are given by:

$$R_{t,a't'} = R_{t-1,a't'} + \sum_{a\in\mathcal{A}} \sum_{d\in\mathcal{D}_a} \delta_{t',a'}(t, a, d)x_{tad} \quad \forall a' \in \mathcal{A}, \quad t' \geq t+1. \tag{6.2}$$

It is often useful to express flow conservation and system dynamics in matrix form, which we do using:

$$A_t x_t = R_{t-1} + \hat{R}_t \tag{6.3}$$

$$R_t - B_t x_t = R_{t-1}. \tag{6.4}$$

For reasons that are made clear in Section 6.2, we have written our equations directly in terms of the incomplete resource vector R_t. The complete resource vector is simply $R_t^+ = R_{t-1} + \hat{R}_t$.

The informational dynamics can be written generally as:

$$K_{t+1} = U^K(K_t, \xi_{t+1}) \tag{6.5}$$

which is how we would represent the process of updating demand forecasting equations, parameter estimation equations, and the storage of other types of information.

The basic statement of the multistage problem is now given by:

$$\max_{\pi\in\Pi} E\left\{\sum_{t\in\mathcal{T}} c_t(X^\pi(I_t), K_t)\right\}$$

subject to flow conservation (equation (6.3)), resource dynamics (equation (6.4)) and informational dynamics (equation (6.5)). There may also be upper bounds on flows representing physical constraints.

The challenge, now, is choosing a function $X^\pi(I_t)$. Popular choices include myopic policies ($I_t = K_t$), or rolling horizon procedures, where $I_t = (K_t, \hat{\Omega}_t)$ where $\hat{\Omega}_t$ represents a forecast of future events made with the information known at time t. If $|\Omega_t| = 1$ then we are using a point estimate of the future and we obtain classical deterministic methods for handling the future. In the next section, we discuss how adaptive dynamic programming methods can be used.

6.2 *Our algorithmic strategy*

Our strategy for solving multistage problems is based on techniques from approximate dynamic programming. Since this approach is not familiar to the stochastic programming community, some background presentation is useful. Recall from Section 6.1 (and in particular equation (6.1)) that we can measure the state of our system before or after we make a decision. It is common in the dynamic programming and control community to write the optimality equations using the state before we make a decision, producing optimality equations of the form:

$$Q_t^+(R_t^+) = \arg\max_{x_t} c_t(X^\pi(I_t), R_t^+) + E\{Q_{t+1}^+(R_{t+1}^+) \mid R_t^+\}. \tag{6.6}$$

Classical dynamic programming techniques are computationally intractable for this problem class. Solving equation (6.6) using classical discrete dynamic programming techniques encounters three "curses of dimensionality": the state space, the outcome space and the action space. Each of these variables are vectors (and potentially vectors of high dimensionality). Computing the value functions using a backward recursion requires computing equation (6.6) for each possible value of the state variable. Computing the expectation requires summing over all the outcomes in the outcome space. Finally, since the expected value function may not have any special structure, solving the optimization problem requires evaluating all possible values of the decision vector x_t.

We overcome this problem using the following strategy. First, recognizing that we do not know Q_{t+1}^+, we replace it with an appropriate approximation that for the moment we denote by $\hat{Q}_{t+1}^+(R_{t+1}^+)$. Next, we recognize that we cannot compute the expectation in equation (6.6). The common strategy is to

replace the expectation with an approximation based on a sample taking from the outcome space:

$$Q_t^+(R_t^+) = \arg\max_{x_t} c_t(X^\pi(I_t),\, R_t^+) + \sum_{\hat{\omega}\in\hat{\Omega}} p(\hat{\omega}) Q_{t+1}^+(R_{t+1}^+(\hat{\omega})). \tag{6.7}$$

Equation (6.7) can be exceptionally difficult to solve for the types of high dimensional, discrete resource allocation problems that arise in transportation. It is particularly inelegant when we realize that it is often the case that the myopic problem (maximizing $c_t x_t$) is a pure network, which means the introduction of the approximate value function is making a trivial integer program quite difficult (the LP relaxation is not a good approximation).

Equation (6.7) is quite easy to solve if we use a linear approximation for $\hat{Q}_{t+1}$. In this case:

$$\begin{aligned}\hat{Q}_{t+1}^+(R_{t+1}^+) &= \hat{q}_{t+1}^+ \cdot R_{t+1}^+ \\ &= \hat{q}_{t+1}^+ \cdot (R_t + A_t x_t + \hat{R}_{t+1}).\end{aligned} \tag{6.8}$$

Taking conditional expectations of both sides of equation (6.8) gives:

$$\begin{aligned}E\{\hat{Q}_{t+1}^+(R_{t+1}^+) \mid R_t^+\} &= E\{\hat{q}_{t+1}^+ \cdot (R_t + A_t x_t + \hat{R}_{t+1}) \mid R_t^+\} \\ &= \hat{q}_{t+1}^1 R_t + \hat{q}_{t+1}^+ A_t x_t + E\{\hat{q}_{t+1}^+ \hat{R}_{t+1} \mid R_t^+\}.\end{aligned} \tag{6.9}$$

The only term on the right side of equation (6.9) involving the expectation is not a function of x_t, so it is only a constant term and can be dropped. The resulting optimization problem is identical to the original myopic optimization problem with a linear adjustment term which would never destroy any nice structural properties of the original problem (for example, network structure).

So, a linear approximation allows us to avoid the problem of taking multidimensional expectations. But what if we use a nonlinear approximation? Now the presence of the expectation presents a serious computational complication. We can circumvent the problem by formulating our optimality recursion around the incomplete state variable R_t. This gives us optimality equations of the form:

$$Q_{t-1}(R_{t-1}) = E\{\arg\max_{x_t} c_t x_t + Q_t(R_t(x_t)) | R_{t-1}\}. \tag{6.10}$$

Again, we propose to replace the recourse function $Q_t(R_t)$ with an approximation $\hat{Q}_t(R_t)$. Note the shift in the time index, which reflects the information content of each variable. We still face the problem of computing (at least approximately) the expectation. However, we are not really interested in computing the expectation. Instead, we need an action x_t which depends on

both R_{t-1} and the new arrivals $\hat{R}_t$. For this reason, we would make a decision contingent on a single sample realization, allowing us to write our decision function using:

$$X_t^{\pi}(R_{t-1}, \hat{R}_t(\omega)) = \arg\max_{x_t} c_t x_t + \hat{Q}_t(R_t(x_t, \omega)) \tag{6.11}$$

subject to:

$$\sum_{d \in \mathcal{D}} x_{tad} = R_{ta}^c + \hat{R}_{ta}^c(\omega) \quad a \in \mathcal{A} \tag{6.12}$$

$$\sum_{a \in \mathcal{A}} x_{tad} \leq R_{ta_d}^o \hat{R}_{ta_d}^o(\omega) \quad d \in \mathcal{D}^o \tag{6.13}$$

$$x_{tad} \in Z_+ \tag{6.14}$$

X_t^{π} is an approximate decision function which can be viewed as a class of policy (in the language of Markov decision theory), where the policy is determined by the choice of $\hat{Q}_t$. Because we have formulated the recourse function in terms of the incomplete state variable, there is no need to directly approximate the expectation (this is being done indirectly through the estimation of $\hat{Q}_t$). We now face the challenge of determining $\hat{Q}_t$. Fortunately, we only have to use the techniques that we described earlier for the two stage problems. Linear approximations remain the simplest to use, but current experimental evidence suggests that piecewise linear, separable approximations are both relatively easy to solve and also provide much higher quality solutions.

Our overall algorithmic strategy is shown in Fig. 10. We refer to this as the "single-pass" version of the algorithm. We initialize $\hat{Q}_t$ for all t. We then simulate forward in time, using dual variables to update the approximation of $\hat{Q}_t$, sampling new information as the algorithm progresses.

The single pass version of the algorithm is the easiest to implement, but it may not work the best. The problem is that it takes a full forward iteration to pass information back one time period. An alternative is to use a two-pass version of the algorithm, where there is a forward pass making decisions, and a backward pass updating dual variables. This version is described in Fig. 11.

These algorithms do not describe specifically how to update the functional approximations $\hat{Q}^n$. For the single-pass version of the algorithm, this updating process is identical to that used for the two-stage problem. We simply use the dual variables for the flow conservation constraint to update $\hat{Q}^n$ just as we did in Section 5.5.2. As we have seen, there are a number of ways to update the value function, so we represent this in general using the notation:

$$\hat{Q}_t^n \leftarrow U^Q(\hat{Q}_t^{n-1}, \overline{q}_t^n, R_t^n).$$

Step 0. Initialization: Initialize $\hat{Q}_t^0$, $t \in \mathcal{T}$. Set $n = 0$.
Step 1. Do while $n \leq N$: Choose $\omega^n \in \Omega$
Step 2. Do for $t = 0, 1, \ldots, T-1$:
  Step 2a. Solve equation (6.11) to obtain $x_t^n = X_t^\pi(R_t^n, \hat{Q}_{t+1}^{n-1})$ and the duals $\hat{q}_t^n$ of the resource constraint equation (6.12).
  Step 2b. Update the resource state: R_{t+1}^n.
  Step 2c. Update the value function approximations using $\hat{Q}_t^n$.
Step 3. Return the policy X_t^π and $\hat{Q}^N$.

Fig. 10. Single pass version of the adaptive dynamic programming algorithm.

Step 0. Initialize $\hat{Q}_t^0$, $t \in T$. Set $n = 0$.
Step 1. Do while $n \leq N$: Choose $\omega^n \in \Omega$
Step 2. Do for $t = 0, 1, \ldots, T-1$:
  Step 2a. Solve (6.11) to obtain $x_t^n = X_t^\pi(R_t^n, \hat{Q}_{t+1}^{n-1})$ and the duals $\hat{Q}_t^n$ of the resource constraint (6.12).
  Step 2b. Update the resource state: R_{t+1}^n.
Step 3. Do for $t = T-1, T-2, \ldots, 1, 0$:
  Step 3a. Compute marginal value of a resource, $\overline{q}_t^n$, using $\hat{Q}_{t+1}^n$ and the optimal basis from the forward pass.
  Step 3b. Update the value function approximations, $\hat{Q}_t^n \leftarrow U^Q(\hat{Q}_t^{n-1}, \overline{q}_t^n, R_t^n)$.
Step 4. Return policy X_t^π and $\hat{Q}^N$.

Fig. 11. Double pass version of the adaptive dynamic programming algorithm.

Updating the value function in the two-pass version is a bit more involved, but the payoff is an updating process that possesses one particularly nice theoretical property. In addition, it appears to work better in practical applications. For the pedagogical purposes of this chapter, we are going to outline the basic idea graphically. Recall that we are solving sequences of pure networks at each time t. At each point in time, we obtain not only an optimal solution but also an optimal basis. Fig. 12 shows the sequence of optimal bases over three time periods. Recall that the network structure of our one-period problems consists of links representing decisions in time period t plus links that represent our piecewise linear value functions. We are interested only in the portion of the basis that consists of links in time period t (with coefficients from the vector c_t), and not in the links which represent the approximation $\hat{Q}$. We note that as a result of our network structure, each basis path from a resource node consists of one or more links in time period t, finally ending in a node in a future time period.

After building the basis paths in the forward simulation, we now have a set of paths extending through the entire horizon. We then compute the cost of a path from a resource node t for attribute vector a until the end of the horizon. Let $\overline{q}_{ta}^n$ be the cost of the path from resource node a at time t until the end

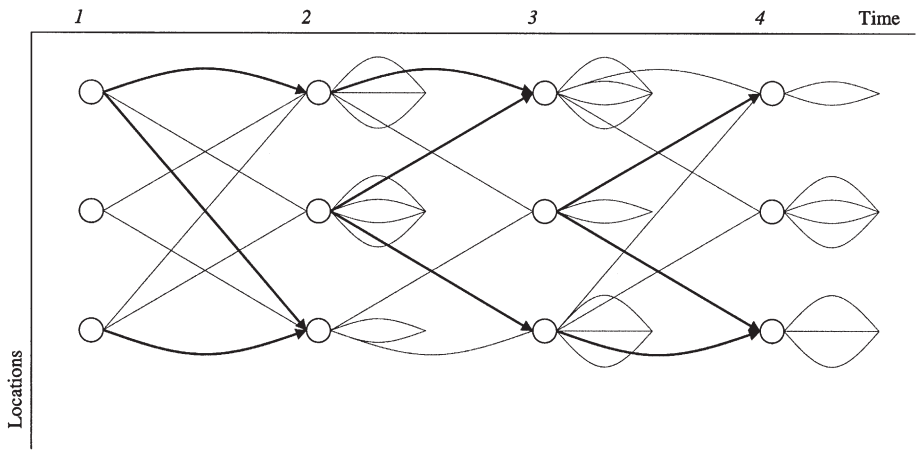

Fig. 12. Optimal network basis from the forward pass.

of the horizon along this basis path. These path costs have a very nice property. Let:

$$F_t^{\pi}(R_t, \omega^n) = \sum_{t'=t}^{T} c_t X_t^{\pi}(I_t(\omega^n)) \tag{6.15}$$

be the costs of a policy π (determined by the functional approximation $\hat{Q}^n$) for outcome ω^n in iteration n, starting in time period t. Then we have:

Theorem 2. *Let $\overline{q}_t^n = (\overline{q}_{ta}^n)_{a \in \mathcal{A}}$ be the vector of path costs from time t to the end of the horizon, given outcome ω^n and functional approximations $\{\hat{Q}_t^n\}_{t \in \mathcal{T}}$ computed from a backward pass. Then $\overline{q}_t$ satisfies:*

$$F_t^{\pi}(R_t, \omega^n) - F_t^{\pi}(R_t', \omega^n) \leq \overline{q}_t^n \cdot (R_t - R_t')$$

Furthermore, if the basis paths in each period t are flow augmenting paths into the supersink, then $\overline{q}_t^n$ is a right gradient of $F_t^{\pi}(R_t, \omega^n)$.

This is a very nice result. These paths are not too hard to compute and provide accurate estimates of the future value of a resource. It turns out that the ability to compute right derivatives is very important. If we just use the dual variables, we overestimate the value of a resource in the future, producing unwanted empty moves.

With our ability to compute valid stochastic gradients of the cost function, we are ready to apply all the tricks we learned for two-stage problems for computing the approximate value functions, $\hat{Q}_t^n$. The forward pass/backward

pass logic is easy to execute and computationally tractable. The Monte Carlo sampling logic avoids problems with expectations. This almost looks too good to be true.

Multistage problems, however, have a few more surprises for us. We begin by discussing the problem of computing expectations, even when we use approximations of the recourse function. We next focus on the issue of problem structure. Keep in mind that transportation problems can be very large, and we are still interested in integer solutions. The first challenge that arises in multistage problems is that resources are reusable (box cars do not simply vanish after we use them). This introduces structural problems that did not occur with two-stage problems, which we review in the context of both single and multicommodity flow problems. We then briefly discuss one of the more annoying, but unavoidable, features of transportation problems: multiperiod travel times.

6.3 Single commodity problems

The difference between the two-stage problem and the one-period problem in a multistage application is that the assignment of a resource to a task produces a resource in the future. In two-stage problems, once the car was assigned to an order in the second stage, it vanished from the system. In multistage problems, we find that we have to solve the type of network depicted in Fig. 13, which is a pure network. This can be solved with commercial LP solvers or specialized network algorithms. Assuming that the functions $\hat{Q}^n_{t+1}(R)$ is piecewise linear, with breakpoints defined for integer values of R, our pure network has integer data (upper bounds and resources) and as a result simplex-based algorithms will return integer solutions.

Pure networks have another attractive property. Linear programming codes will return a dual variable $\tilde{q}_{ta}$ for the resource constraint equation (6.12). It is far more desirable to obtain explicit right and, if possible, left

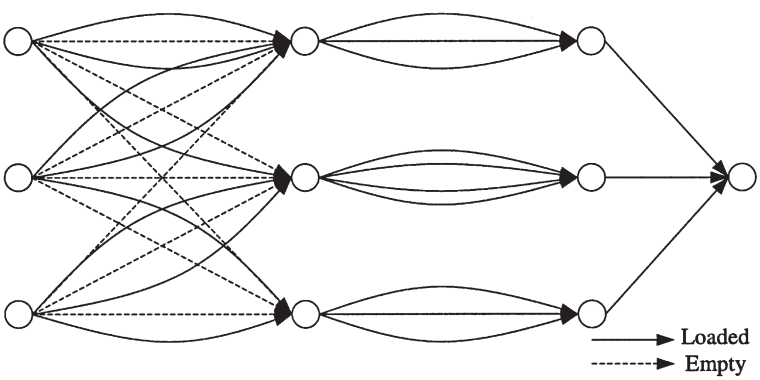

Fig. 13. Network structure of a one-period single commodity problem.

gradients, which we can denote $\tilde{q}_t^+$ and $\tilde{q}_t^-$. With pure networks, left and right gradients can be found by solving flow augmenting path problems into and out of (respectively) the supersink. A right gradient gives us a precise estimate of a particular slope of the recourse function. The computation of explicit left and right gradients is very important in problems with large attribute spaces, where the values of R_{ta} can be relatively small. If $\mathcal{A}$ is small relative to the number of resources being managed (implying that the values of R_{ta} are much greater than one), then the issue is unlikely to be noticeable.

The pure network structure actually arises in a larger class of problems. Assume that we are modeling resources with attribute a, and recall from Section 3 that $a^M(t, a, d)$ is the *terminal attribute function*, giving the attributes of a resource with attribute a after decision d has been applied to it. Now define:

Definition 3. A resource allocation problem has the Markov property if $a^M(t, a, d) = a^M(t, a', d), \forall a, a' \in \mathcal{A}$.

The Markov property for resource allocation problems implies that the attributes of a resource after it has been acted on is purely a function of the decision. Classical single commodity flow problems exhibit this property because a truck in location i which is then assigned to move a load from i to j, is now a truck at location j. If all the trucks are the same, then the attribute of the truck (its location) is purely a function of the attribute of the decision which was to move a load to j. If there were different types of trucks that could serve the load (a multicommodity problem), then the attribute of the truck after moving the load continues to have the attribute of the truck before the load (a characteristic that has nothing to do with the decision).

It is apparent that classical multicommodity problems which allow substitution of different commodity types for the same task will never have the Markov property, but multiattribute problems (where the attribute vector a consists purely of dynamic attributes), can possess this property. Consider a chemical trailer that can move basic argon gas or purified argon gas. Only "clean" trailers can move purified argon gas. A clean trailer can move basic gas, but then it is no longer clean. There is no substitution for purified argon gas, and any truck moving basic argon gas is no longer clean (although it can be put through a cleansing process). Thus, the attributes of the truck after a trip are determined completely by the characteristics of the trip.

6.4 Multicommodity problems

When we encountered single commodity problems, we found that the single period problems were pure networks if the approximation of the recourse function were linear, or piecewise linear, separable. Now we consider what happens when we try to solve multicommodity problems. Recall that we let $\mathcal{I}$ be a set of locations and $\mathcal{K}$ be the set of commodities. We follow the standard notation of multicommodity flow problems and let R_{ti}^k be the number of

resources of type k in location i, and let x^k_{tid} be the number of resources of type k in location i that we act on with decision d at time t.

Below, we show that multistage, multicommodity problems are especially easy to solve if we use linear value function approximations, whereas the use of separable nonlinear (piecewise-linear) approximations introduces complications which, while tractable, have to be addressed.

The case of linear approximations

When we replace the value function $Q_t(R_t)$ with an approximation $\hat{Q}_t(R_t)$, we obtain the decision function:

$$X^\pi_t(I_t) = \arg\max_x \sum_{i\in\mathcal{I}} \sum_{d\in\mathcal{D}} c_{tid}x_{tid} + E\{\hat{Q}_t(R_t(x_t)) \mid R_{t-1}\} \tag{6.16}$$

If we use a linear approximation for $\hat{Q}$, the equation (6.16) reduces to:

$$X^\pi_t(I_t) = \arg\max_x \sum_{i\in\mathcal{I}} \sum_{d\in\mathcal{D}} c_{tid}x_{tid} + \sum_{t'>t} \sum_{j\in\mathcal{I}} \hat{q}_{t,jt'}(R_{t,jt'}(x_t, \omega_t)) \tag{6.17}$$

$$= \arg\max_x \sum_{i\in\mathcal{I}} \sum_{d\in\mathcal{D}} c_{tid}x_{tid} + \sum_{t'>t} \sum_{j\in\mathcal{I}} \hat{q}_{t,jt'}R_{t,jt'}(x_t, \omega_t) \tag{6.18}$$

where:

$$R_{t,jt'} = \sum_{i\in\mathcal{I}} \sum_{d\in\mathcal{D}_i} \delta_{t',j}(t, i, d)x_{tid} \tag{6.19}$$

Substituting equation (6.19) into (6.18) gives:

$$X^\pi(I_t) = \max_x \sum_{i\in\mathcal{I}} \sum_{d\in\mathcal{D}} c_{tid}x_{tid} + \sum_{t'>t} \sum_{j\in\mathcal{I}} \hat{q}_{t,jt'} \sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{D}} \delta_{t',j}(t, i, d)x_{tid} \tag{6.20}$$

$$= \max_x \sum_{i\in\mathcal{I}} \sum_{d\in\mathcal{D}} c_{tid}x_{tid} + \sum_{i\in\mathcal{I}} \sum_{d\in\mathcal{D}} \left(\sum_{t'>t} \sum_{j\in\mathcal{I}} \delta_{t',j}(t, i, d)\hat{q}_{t,jt'}x_{tid} \right) \tag{6.21}$$

It is easy to see that:

$$\sum_{t'>t} \sum_{j\in\mathcal{I}} \delta_{t',j}(t, i, d)\hat{q}_{t,jt'}x_{tid} = \hat{q}_{t,i^M_{tid},t+\tau^M_{tid}}x_{tid} \tag{6.22}$$

where i^M_{tid} is our terminal attribute function (using location indices instead of attribute vectors) and τ^M_{tid} is the time required to complete the decision. This allows us to reduce equation (6.22) to:

$$X^\pi(I_t) = \max_x \sum_{i\in\mathcal{I}} \sum_{d\in\mathcal{D}} \left(c_{tid} + \hat{q}_{t+1,i^M_{tid},t+\tau^M_{tid}}\right) x_{tid}. \tag{6.23}$$

Equation (6.23) demonstrates that a linear approximation of the recourse function is the same as adding a price to the cost of each assignment, with the same structure as the original one-period problem. So, solving multicommodity flow problems with linear approximations is no harder than solving single commodity problems.

The case of nonlinear approximations

The situation is somewhat different when we are using nonlinear functional approximations. Our one-period problem now takes on the network structure shown in Fig. 14 which is itself a multicommodity flow problem, if we are using nonlinear functional approximations. We note, however, that these are not especially large multicommodity flow problems, since they are for a single time period. Perhaps the most significant practical problem that might arise is the loss of integrality. In fact, we have found that a commercial LP package solving a continuous relaxation of the problem returns integer solutions 99.9% of the time (the rare instance of a fractional solution is quickly solved with branch and bound or simple rounding heuristics).

Perhaps the more practical challenge of multicommodity flow problems is that we do lose the ability to find left and right gradients using flow

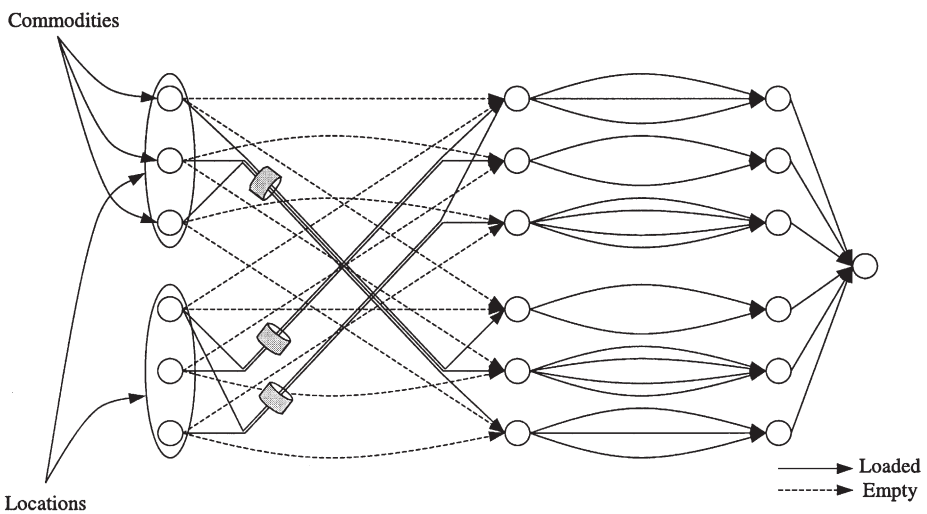

Fig. 14. Network structure of a one-period multicommodity problem.

augmenting path calculations. As we pointed out before, this is not necessary for all problem classes, but we have worked on problems with large attribute spaces where dual variables from the LP solver are simply not good enough. In this case, we are resorting to performing numerical derivatives (incrementing the right hand side and solving the problem again). Since the optimal basis is often optimal for the perturbed problem, this procedure can be quite fast.

6.5 *The problem of travel times*

One of the most annoying characteristics of transportation is the property that it takes time to complete a decision. Using the vocabulary of discrete time models, we refer to these problems as having "multiperiod" travel times. More generally, we would refer to these as "multiperiod transfer times," since there are many activities in transportation that do not actually involve moving from one location to another (drivers have to go on rest, trailers have to be cleaned, locomotives have to be maintained). But, the concept of traveling between locations is easy to visualize.

The implication of multiperiod travel times is that after acting on a resource at time t, the resource is committed to an activity in time period $t+1$ and we cannot act on it. At the same time, we cannot ignore it, because it will eventually complete its movement, and we have to take this into account when we make decisions in time period $t+1$. Fig. 15 illustrates the issue in the context of fleet management. Assume that we will need containers at location **c** at time $t=6$, and we can move them there from either **a**, which is five time periods away, or from **d**, which is only two time periods away. Because **a** is

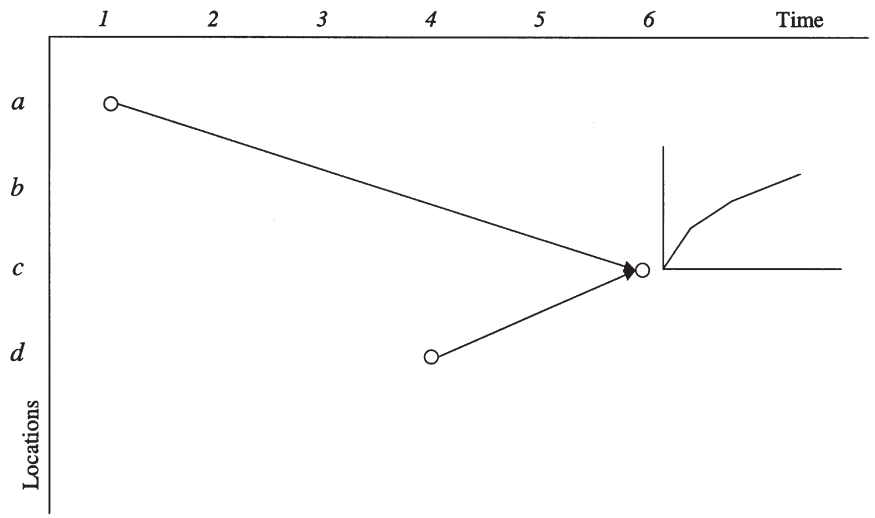

Fig. 15. Multiperiod travel times imply that decisions to move vehicles at different points in time can produce available capacity at the same time in the future.

farther away, we will first look at the problem of the shortage of containers at **c** for time $t=6$ at time $t=1$. At this point, we would not have considered moving containers from location **d**, since this decision would not be made until $t=4$. Seeing the shortage, we might move containers the longer distance from **a** to **c**, rather than waiting until time $t=4$ and moving them from the closer location at **d**.

The modeling of multiperiod travel times is the same as the modeling of lagged information processes. Earlier in the chapter, we introduced the notation $R_{tt'}$ which gives the vector of resources that we know about at time t which become actionable at time t'. The difference between t' and t is the information lag. Lags can arise when customers call in orders in advance. In the case of our rail car distribution problem, it can arise when a shipper tells the carrier that a freight car will become empty in 3 days. Information lags also arise whenever the travel time from one location to another is more than one time period.

The problem of multiperiod travel times is unique to multistage stochastic models, and furthermore it is unique to the usage of nonlinear functional approximations. With a nonlinear function, location **a** "sees" the steeper part of the slope of the function at **c**, since we have not yet made the decision to move cars from **d**. The result is something that we call the "long haul bias," which arises in any application where resources can be committed in the future.

The standard solution for problems of this type is to use an augmented state variable. Assume that a container started moving from i to j, departing at time t, on a trip that requires time τ_{ij}. Now let the variable s be the remaining time in the trip, so at time $t+1$, we would have $s=\tau_{ij}-1$. Using our multiattribute notation, the remaining trip time s simply becomes a new dimension of the attribute vector a. Given this representation, we can solve multiperiod problems using the same techniques that we have described up to now.

This particular solution creates practical problems in transportation applications. Problems in trucking and rail, for example, can have trip times that range from 30 min to 3 days (for movements from the midwest to the west coast). We can often work with time steps of 2 or 4 h, producing trips that are often 10–20 time periods in length. The result is an attribute space $\mathcal{A}$ that is now approximately 10 times bigger (a dramatic increase in the size of the problem). Since a single movement is now broken into 10 or 20 time steps, pure forward pass algorithms become completely unworkable, although we can partly overcome the slow backward communication of duals using shortcuts that take advantage of the properties of the problem.

The augmented state representation has to be viewed as a brute force solution to the problem of multiperiod travel times which can actually hide nice structural properties. For example, it is not readily apparent using this representation that the problem of multiperiod travel times vanishes when we

use linear functional approximations. When $\hat{Q}_{C,t+6}$ is a linear function, both locations **a** and **d** "see" the same slope. If they both send containers in response to an attractive slope, then the sample gradient of the function at location **c** will be reduced, and the location will become less attractive. Over sufficient iterations, the model should discover the slope that attracts capacity from closer locations but not farther ones.

It is beyond the scope of our presentation to fully describe the solution to the "multiperiod travel time" problem when using nonlinear functional approximations, but we note that it involves replacing the single functional approximation $\hat{Q}_{t'}$ with a family of functions $\hat{Q}_{tt'}$ which are used to describe the impact of decisions made at time t on future points in time $t' > t$. This approach produces solutions that are comparable in quality to those obtained using nonlinear approximations with single-period travel times, but as of this writing, the theory behind this approach is immature.

We again see that linear approximations avoid complexities that arise in the context of nonlinear approximations. But, the jury is still out regarding the approach that will produce the best results in the laboratory, and implementable results in the field.

7 Some experimental results

There is surprisingly little work comparing different stochastic programming approaches. This is especially true of multistage problems, but it is even true of the much more mature two-stage problem. Furthermore, the work on two-stage problems has not been done explicitly in the context of resource allocation problems (they are simply characterized as two-stage linear programs) which makes it difficult to generate a library of datasets which focus on the specific dimensions of resource allocation problems (such as number of locations, number of commodity types, repositioning costs in the second stage, and so on). As a result, we do not have a standard library of test datasets for either two-stage or multistage resource allocation problems.

This chapter has described the use of a relatively new class of approximation strategies that are especially well suited to resource allocation problems. These approximations focus on using linear or separable, nonlinear approximations of the recourse function. Many resource allocation problems require integer solutions. Separable, nonlinear functions can be constructed as piecewise linear approximations which produces first stage problems that are either pure networks, or integer multicommodity flow problems with very tight LP relaxations. In this section, we provide some preliminary comparisons between these approximation strategies and a variant of Benders decomposition called the CUPPS algorithm.

Section 7.1 describes some experiments that were performed in the context of two-stage problems. Section 7.2 then presents some results for multistage problems.

7.1 Experimental results for two-stage problems

Virtually all problems in transportation and logistics involve multistage problems, but as our discussion has demonstrated, multistage problems are typically solved as sequences of two-stage problems. As a result, we have to begin with an understanding of how well we can solve two-stage problems.

We undertook a series of preliminary experiments focusing on questions regarding rate of convergence, scalability and solution quality. We used a randomly generated dataset (which allowed us control over its characteristics), and compared SPAR (which uses separable, piecewise linear approximations) and CUPPS (based on Benders decomposition). Our evaluation strategy consisted of running each algorithm for a fixed number of training iterations, and then comparing solution quality over a series of testing iterations.

Our datasets were created very simply. N locations were uniformly generated over a 100×100 square. Initial supplies of resources were randomly generated and spread around these locations. The resources and demands were generated in such a way that ensured that the expected number of resources equaled the expected number of demands (our work has shown that this is when the problems are the most difficult, the most interesting and the most realistic). Demands (in the form of loads of freight to be moved) were randomly generated with both an origin and a destination, where the contribution from covering a demand was set at \$1.5 per mile, where the length of the load is given by the distance from the origin to the destination. Transportation costs in the first stage are fixed at \$1 per mile, while transportation costs in the second stage are fixed at \$2 per mile. This provides an incentive to reposition in the first stage before we know what the demands are.

Our first experiments studied the rate of convergence of each algorithm on a dataset with 30 locations. Fig. 16 shows the objective function for SPAR and CUPPS averaged over 50 testing iterations, as a function of the number of training iterations. With 950 training iterations, the methods are virtually identical. However, as the number of training iterations diminishes, SPAR seems to perform better, suggesting that it has a faster rate of convergence.

This conclusion is supported in Fig. 17 which compares SPAR and CUPPS as a function of the number of locations. For each run, we used 200 training iterations, and the algorithms were run on problems with 20, 30, 40 and 90 locations. The results show that SPAR and CUPPS work similarly for smaller problems, but that SPAR works better (with 200 training iterations) for larger problems. This suggests that the SPAR-class algorithms exhibit a faster rate of convergence, especially as the problem size grows.

We finally looked at the results for every observation within the test sample to get a sense of the distribution of the difference between SPAR and CUPPS (see Fig. 18). To our surprise, we found that SPAR and CUPPS provide almost identical results for every outcome for smaller datasets. For larger datasets, SPAR outperformed CUPPS on *every* outcome.

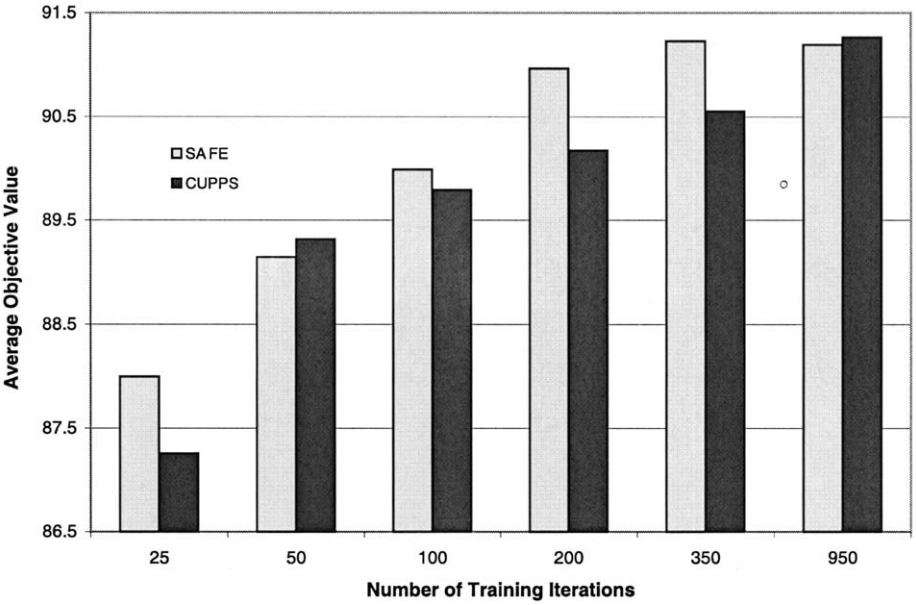

Fig. 16. The effect of the number of training iterations on SPAR and CUPPS, illustrating the faster convergence for SPAR.

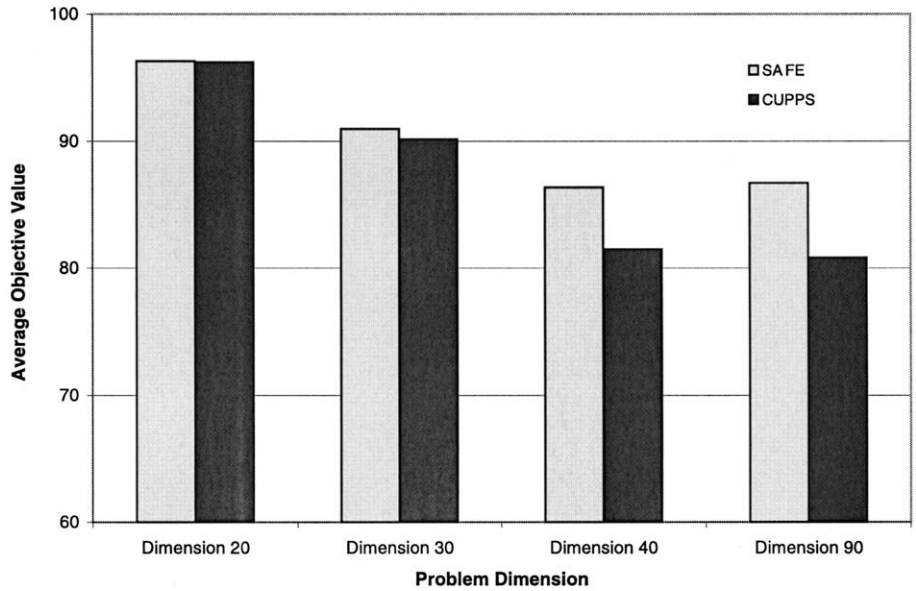

Fig. 17. SPAR vs. CUPPS for two-stage problems, illustrating better results for SPAR when the problem size grows.

7.2 Experimental results for multistage problems

The solution of multistage problems for our applications consist of solving sequences of two-stage problems. The question now is, how well does this work? Should we expect that an optimal or near-optimal algorithm for two-stage problems will work similarly on multistage problems?

The biggest difference between two-stage and multistage problems for our problem class is that the sequences of two-stage problems that we solve in a multistage setting have random initial starting states. When we solve a two-stage problem, the initial state is (normally) deterministic. This means that the optimal solution to the first stage is deterministic, which means that our approximation for the second stage has to be accurate only in the vicinity of the optimal solution of the first stage. In the case of multistage problems, the initial resource state at some time t in the future depends on previous decisions, while the approximation of the recourse function for this problem is fixed, and must perform well over a range of initial states. As a result, the demands on the accuracy of the recourse function for the second stage are much higher.

A major difficulty that arises in the evaluation of approximations for multistage problems is identifying a good benchmark. Optimal solutions are simply not obtainable, and tight bounds are not readily available. For this reason, we use two strategies. First, it is useful to see how well a stochastic algorithm works on a deterministic problem. This need arises since it is typically the case that a company will want to test how well the algorithm works by running it on past history (which is deterministic). Deterministic formulations tend to be the benchmark, and if a stochastic algorithm does not work well on a deterministic problem, it raises the question of how it can be a good method for a stochastic problem.

The second strategy we use is to compare against deterministic rolling horizon procedures using stochastic data. Again, this is the most common strategy used in engineering practice for solving stochastic problems.

We first ran the SPAR algorithm on a deterministic, single commodity problem which can be formulated as a pure network. A significant assumption is that the time at which a load had to be moved was fixed (so-called tight time windows). Table 1 reports these results for problems with 20, 40 and

Table 1
Percentage of integer optimal value obtained using SAFE for second set of deterministic experiments with single-period time windows (network problems)

Locations	Simulation horizon		
	15	30	60
20	100.00%	100.00%	100.00%
40	100.00%	99.99%	100.00%
80	99.99%	100.00%	99.99%

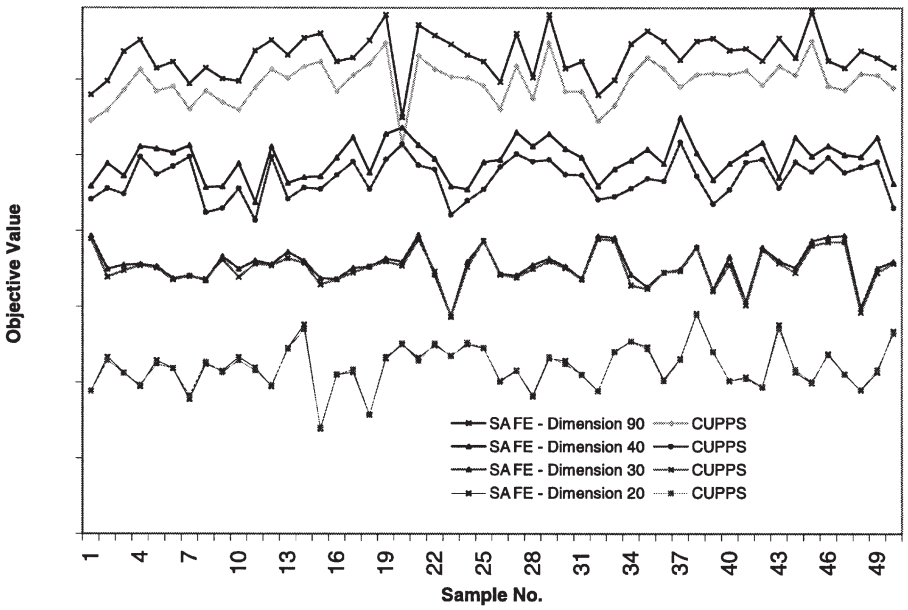

Fig. 18. SPAR outperforms CUPPS as the problem size grows for every outcome in the testing dataset.

80 locations. These results indicate that the algorithm is effectively returning optimal solutions (Fig. 18).

We then ran the algorithm on four stochastic datasets and compared against a rolling horizon procedure (RHP). The RHP used a point forecast of the future to make decisions for the current time period. Tests were run with different planning horizons to ensure that we were using the best possible planning horizon for the RHP. The results are shown in Fig. 19 which shows that the SPAR algorithm is producing results that are significantly better than a deterministic RHP.

Further research is needed before we understand the true value of a stochastic formulation. Our rolling horizon experiments were performed assuming that there was no advance information. The value of a stochastic model also depends on the economics of making the wrong decision (the recourse).

8 A list of extensions

This chapter has introduced a basic problem class, discussing along the way practical algorithmic strategies, but steadily introducing issues that address the richness of transportation problems. Our notational framework, which at first may seem clumsy to researchers accustomed to even simpler notation, is

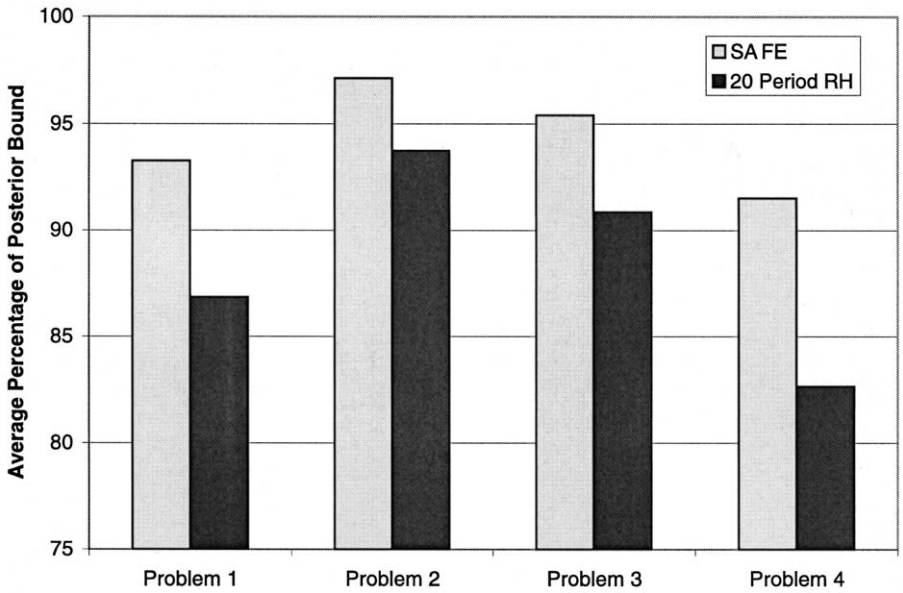

Fig. 19. Comparison of SPAR approximations to rolling horizon procedures for 20 location datasets and different substitution rules.

designed to provide a natural bridge between the classical notation of mathematical and stochastic programming, while providing for some of the issues that arise in practice. Our progression of problems, from two-stage to multistage, from single commodity with no substitution through multi-commodity and heterogeneous resource allocation problems, was designed to bring the reader through a list of issues that add to the realism of the model being considered.

We have brought the reader to the edge of realistic problems that rise in practice, but real problems of the type that arise in freight transportation have even more surprises to delight and frustrate the research community. In this section, we provide a hint of problems that remain to be addressed. All of the issues listed below represent forms of inaccurate information.

a) Random travel times—Not only are there multiperiod travel times, it is normally the case that the travel time is random. A travel time may be uncertain even when a decision is made.
b) Advance information—Customers place orders in advance. Railroads might let you know that they will give you empty cars in 3 days. A driver might tell you that he will start in a week. We need to model the presence of information that is known, but not actionable.
c) Demand backlogging and the "now vs. later" problem—If we do not serve the customer request now, we may be able to serve it later (at a price). We can assign a driver to a distant load now, or wait for a closer

load to be called in later. We need to identify when we should make a decision now (to serve a request) or wait until later.

d) Multiple, reusable layers—Perhaps the most challenging problem is the presence of multiple resource layers. (Recall that customer demands can also represent a kind of resource layer). The simplest example arises with backlogging demands: if we do not serve a demand now, it is still in the system in the next time period. In freight transportation, we may have to manage drivers, tractors and trailers, or locomotives, boxcars and crews. It is possible to model some freight transportation operations with four or five layers, and most of them are reusable.

e) User noncompliance—An overlooked dimension of most models is that what the model is recommending is not what is being implemented. So the costs and contributions that we are adding up in the computer are not the costs and contributions we are getting in the field. The difference between what a model recommends and what is implemented in the field is a source of randomness that draws into question the value of so-called optimal solutions that ignore this source of randomness.

f) Multiagent control—Large problems are typically broken into smaller problems which are managed by different people. Decisions made by one person need to anticipate the impact on another. But it is impossible to predict what someone else will do with certainty. This is not the same as the user compliance problem, but it is another instance of solving a problem where the randomness is in predicting what someone will do.

g) Data errors—We all know about data problems and recognize that we have to fix them, but we overlook the fact that the presence of data errors is again a source of noise. If data errors are going to require humans to override model recommendations, then so-called ''optimal'' solutions are no longer optimal (even within a single stage).

h) Incomplete information—Random variables arise when we have information that is not known now, but can be measured later. There are problems where information is unknown but can never be measured directly, which means we could never estimate a probability distribution. But the missing information is captured in historical databases of past decisions.

9 Implementing stochastic programming models in the real world

We close our discussion of stochastic problems by raising some of the issues that arise when we try to implement stochastic programming models in practice. In the beginning of this chapter, we made the argument that explicit models of uncertainty produce more realistic behaviors, often exactly the behaviors that humans will mimic (perhaps with less precision) but deterministic models will overlook.

Humans have an innate ability to deal with imperfect information and noise. We allow more time to make an appointment. We provide slack time in airline schedules to allow for possible delays. We own extra trucks and locomotives to account for breakdowns, congestion and spikes in demand. Trucking companies have large rooms of people planning operations who spend 90% of their time collecting and verifying data, and only 10% actually making decisions. We would expect that planners would quickly embrace models which capture the uncertainty that they spend so much time dealing with.

Surprisingly, this is not the case. Stochastic models arise because of the need to forecast an activity in the future, a forecast that is made with uncertainty. And yet in the business world, the word "forecast" is synonymous with the concept of a "point forecast." When we ask for a forecast of the number of box cars needed in a region next week, we do not want to be told "somewhere between 70 and 100." When we ask for a forecast, we expect an answer such as "85."

At the heart of any stochastic programming application is a *distributional forecast*, an explicit recognition that a range of outcomes is possible. As much as people in operations have learned to deal with uncertainty, they uniformly have difficulty working with models which capture this uncertainty. Often this is because they are looking either for a specific recommendation of what to do right now (which even a stochastic model should do) or a *plan* of what should be done in the future. It can be said that a plan is a (point) forecast of a decision. Even for events in the future, people want to be told what is going to happen (even if they realize a plan has uncertainty, they are still looking for a specific estimate of what is going to happen).

Consider the practical difficulty of testing a stochastic model. In a deterministic model, we expect the model to move exactly the number of freight cars that are needed. If a deterministic model moved 60 cars to satisfy orders for 50, we might reasonably conclude that there was some sort of bug in the software. Yet this might be exactly the right answer for a stochastic model. But how do we know that we are getting optimal behavior, or simply the result of a programming error?

For this reason, we usually first test a stochastic programming model by solving a deterministic problem, and comparing the solution against a standard solver for deterministic problems. It seems obvious that an algorithm that can solve a stochastic problem should also be able to solve a deterministic problem, but this can be harder than it looks. Figure 20 shows a recourse function for deterministic and stochastic models. As a rule, stochastic problems are smoother and better behaved. As a result, linear approximations can work quite well. However, these same techniques will not work as well on the sharply angled function produced when we replace our range of possible outcomes with a point forecast. While we can argue that we should not have to test our stochastic algorithm on a deterministic model, this is a powerful debugging and testing tool, and we have found that it is necessary

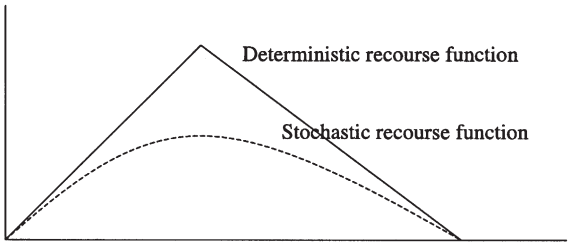

Fig. 20. Recourse functions for deterministic and stochastic functions.

to develop techniques that work well on deterministic as well as stochastic problems.

10 Bibliographic notes

The techniques in this chapter have their roots in stochastic approximation methods (Robbins and Monro (1951), Blum (1954), Dvoretzky (1956), Gladyshev (1965)), stochastic gradient methods Ermoliev (1988), general stochastic linear programming (Birge and Louveaux (1997), Infanger (1994), Kall and Wallace (1994), Higle and Sen (1991)) and dynamic programming (both classical methods reviewed in Puterman (1994), and approximate methods, such as those covered in Bertsekas and Tsitsiklis (1996)). For reviews of these topics, the reader is referred to the introductory chapters in this volume.

Applications in transportation and logistics represented some of the earliest motivations for stochastic programming. Dantzig (1955) used fleet management (in an airline setting) as an early motivation for stochastic programming. Ermoliev et al. (1976) formulated the planning of empty shipping containers as a stochastic program.

There has been a rich history of research in fleet management in the context of the "car distribution problem" of railroads. Most of this work consists of deterministic linear programming models (Feeney (1957), Leddon and Wrathall (1967), Gorenstein et al. (1971), Misra (1972), White (1972), Herren (1973, 1977), White and Bomberault (1969), Haghani (1989), Mendiratta (1981) and Mendiratta and Turnquist (1982)). Dejax and Crainic (1987) provide a thorough review of the research in fleet management at the time, covering both rail and intermodal container applications. Crainic et al. (1993) provide a general stochastic, dynamic model for container distribution.

Jordan and Turnquist (1983) provide a stochastic formulation of the empty car distribution problem. In their model, a car could be assigned to at most one demand, and cars could not be repositioned more than once. This structure allowed the problem to be formulated as a nonlinear programming problem. Powell (1986) extended this methodological approach, using the trucking industry as the context, to multistage problems with reusable resources.

This formulation involved forming deterministic decision variables which specified the percentage of the supply of trucks at a node that would be moved loaded or empty from one location to another. This line of research, however, was unable to properly model the recourse strategy where a truck might be assigned to choose from a set of loads going out of a location to various destinations.

Powell (1987) solved this problem in a way that produced a pure network formulation for the first stage subproblem (see Powell (1988) for an overview of different ways of formulating the problem). A major strength of the technique was that it produced nonlinear functions of the value of vehicles in the future. The ideas behind this research produced a series of articles that approximated the recourse function by using the structure of the underlying recourse function (Powell and Frantzeskakis (1992), Frantzeskakis and Powell (1990), Powell and Cheung (1994a,b) and Cheung and Powell (1996)). These papers introduced concepts such as nodal recourse (the recourse is to allocate flow over different links out of a node), tree recourse (the recourse is to optimize flows over a tree) and other restricted recourse strategies aimed at approximating more complex problems. Although the results were promising, this line of research required approximating the future in a way that prevented the techniques from being applied to the most general (and realistic) problems.

Powell (1998) (see also Carvalho and Powell (2000)) took a different tact and solved the problem as a sequence of network problems, using linear approximations of the value of the future. This approach was computationally quite easy, and scaled to much harder problems. Powell et al. (2002b) showed how the technique could be applied to the heterogeneous resource allocation problem, which is a more general problem than the multicommodity flow problem (the resource attribute space is much larger). The use of linear approximations to represent the future, however, produced instabilities that were solved by the use of control variables that limited the amount of flow that could be moved from one location to another. This worked very well for single commodity problems, but did not scale well to multicommodity or heterogeneous resource allocation problems.

Godfrey and Powell (2001) introduce an adaptive sampling technique, dubbed the CAVE algorithm, that produces nonlinear approximations of a recourse function using stochastic gradients. The method is easy to apply and produces piecewise linear, separable approximations of a recourse function. Furthermore, all it requires is the dual variable from a second stage problem, and does not require that the problem have any particular structure. Experimental work suggested that the algorithm might be optimal for two-stage problems, but at a minimum it produced results that were extremely close to optimal for both deterministic and specially structured stochastic problems. Godfrey and Powell (2002a) apply it to stochastic, multistage problems and demonstrate very good results relative to rolling horizon procedures. Godfrey and Powell (2002b) investigated the case of resource allocation where the travel time from one location to another can be multiple periods.

A naive application of the CAVE algorithm produced extremely poor results, and a variation is suggested that provides results that are comparable to the single period travel time case.

Topaloglu and Powell (2002) applies similar techniques to stochastic multicommodity flow problems. This problem class introduces the additional complexity that multicommodity problems, combined with nonlinear approximations of the future, produce sequences of (usually integer) multicommodity flow problems. The method appears to work well on both deterministic and multistage stochastic integer multicommodity flow problems.

The SPAR algorithm is part of a broader class of algorithms that use stochastic gradients but maintain structural properties such as concavity. This strategy was first suggested in Godfrey and Powell (2001) as the CAVE algorithm, but the SPAR algorithm, introduced by Powell et al. (2002a), offers several provable convergence results and appears to also work better experimentally (see Topaloglu and Powell, 2002 for the use of these techniques for multistage problems). Auxiliary function methods (Culioli and Cohen, 1990; Cheung and Powell, 2000) maintain concavity by starting with a concave function and using stochastic gradients to update the function (effectively tilting it).

References

Bertsekas, D., J. Tsitsiklis (1996). *Neuro-Dynamic Programming*, Athena Scientific, Belmont, MA.

Birge, J., F. Louveaux (1997). *Introduction to Stochastic Programming*, Springer-Verlag, New York.

Blum, J. (1954). Multidimensional stochastic approximation methods. *Annals of Mathematical Statistics* 25, 737–744.

Carvalho, T.A., W.B. Powell (2000). A multiplier adjustment method for dynamic resource allocation problems. *Transportation Science* 34, 150–164.

Cheung, R., W.B. Powell (1996). An algorithm for multistage dynamic networks with random arc capacities, with an application to dynamic fleet management. *Operations Research* 44(6), 951–963.

Cheung, R.K.-M., W.B. Powell (2000). SHAPE: A stochastic hybrid approximation procedure for two-stage stochastic programs. *Operations Research* 48(1), 73–79.

Crainic, T., M. Gendreau, P. Dejax (1993). Dynamic stochastic models for the allocation of empty containers. *Operations Research* 41, 102–126.

Culioli, J.-C., G. Cohen (1990). Decomposition/coordination algorithms in stochastic optimization. *SIAM Journal of Control and Optimization* 28, 1372–1403.

Dantzig, G. (1955). Linear programming under uncertainty. *Management Science* 1, 197–206.

Dejax, P., T. Crainic (1987). A review of empty flows and fleet management models in freight transportation. *Transportation Science* 21, 227–247.

Dvoretzky, A. (1956). On stochastic approximation, in: J. Neyman (ed.), *Proc. Berkeley Sym. on Math. Stat. and Prob.*, University of California Press, Berkeley, pp. 39–55.

Ermoliev, Y. (1988). Stochastic quasigradient methods, in: Y. Ermoliev, Wets, R. (eds.), *Numerical Techniques for Stochastic Optimization*, Springer-Verlag, Berlin.

Ermoliev, Y., T. Krivets, V. Petukhov (1976). Planning of shipping empty seaborne containers. *Cybernetics* 12, 664.

Feeney, G. (1957), Controlling the distribution of empty cars, in: *Proc. 10th National Meeting, Operations Research Society of America.*

Frantzeskakis, L., W.B. Powell (1990). A successive linear approximation procedure for stochastic dynamic vehicle allocation problems. *Transportation Science* 24(1), 40–57.

Gladyshev, E.G. (1965). On stochastic approximation. *Theory of Prob. and its Appl.* 10, 275–278.

Godfrey, G., W.B. Powell (2002). An adaptive, dynamic programming algorithm for stochastic resource allocation problems I: Single period travel times. *Transportation Science* 36(1), 21–39.

Godfrey, G., W.B. Powell (2002). An adaptive, dynamic programming algorithm for stochastic resource allocation problems II: Multi-period travel times. *Transportation Science* 36(1), 40–54.

Godfrey, G.A., W.B. Powell (2001). An adaptive, distribution-free approximation for the newsvendor problem with censored demands, with applications to inventory and distribution problems. *Management Science* 47(8), 1101–1112.

Gorenstein, S., S. Poley, W. White (1971). On the scheduling of the railroad freight operations, Technical report 320-2999, ibm philadelphia scientific center, IBM.

Haghani, A. (1989). Formulation and solution of a combined train routing and makeup, and empty car distribution model. *Transportation Research* 23B(6), 433–452.

Herren, H. (1973). The distribution of empty wagons by means of computer: An analytical model for the Swiss Federal Railways (SSB). *Rail International* 4(1), 1005–1010.

Herren, H. (1977). Computer controlled empty wagon distribution on the SSB. *Rail International* 8(1), 25–32.

Higle, J.L. and S. Sen (1991), Stochastic Decomposition: an algorithm for two stage linear programs with recourse, *Math. of Operations Research* 16, 215–240.

Infanger, G. (1994). *Planning under Uncertainty: Solving Large-scale Stochastic Linear Programs*, The Scientific Press Series, Boyd & Fraser: New York.

Jordan, W., M. Turnquist (1983). A stochastic dynamic network model for railroad car distribution. *Transportation Science* 17, 123–145.

Kall, P., S. Wallace (1994). *Stochastic Programming*, John Wiley and Sons, New York.

Leddon, C., E. Wrathall (1967). Scheduling empty freight car fleets on the louisville and nashville railroad, in: *Second International Symposium on the Use of Cybernetics on the Railways*, *October*, Montreal, Canada, pp. 1–6.

Mendiratta, V. (1981). A dynamic optimization model of the empty car distribution process, Ph.D. Dissertation, Department of Civil Engineering, Northwestern University.

Mendiratta, V., M. Turnquist (1982). A model for the management of empty freight cars. *Trans. Res. Rec.* 838, 50–55.

Misra, S. (1972). Linear programming of empty wagon disposition. *Rail International* 3, 151–158.

Powell, W.B. (1986). A stochastic model of the dynamic vehicle allocation problem. *Transportation Science* 20, 117–129.

Powell, W.B. (1987). An operational planning model for the dynamic vehicle allocation problem with uncertain demands. *Transportation Research* 21B, 217–232.

Powell, W.B. (1988). A comparative review of alternative algorithms for the dynamic vehicle allocation problem, in: B. Golden, A. Assad (eds.), *Vehicle Routing: Methods and Studies*, North Holland: Amsterdam, pp. 249–292.

Powell, W.B., A., C.T. (1998). Dynamic control of logistics queuing network for large-scale fleet management, Transportation Science 32(2), 90–109.

Powell, W.B., R. Cheung (1994). A network recourse decomposition method for dynamic networks with random arc capacities. *Networks* 24, 369–384.

Powell, W.B., R. Cheung (1994). Stochastic programs over trees with random arc capacities. *Networks* 24, 161–175.

Powell, W.B., L. Frantzeskakis (1992). Restricted recourse strategies for stochastic, dynamic networks. *Transportation Science* 28(1), 3–23.

Powell, W.B., A. Ruszczynski, H. Topaloglu (2002a). Learning algorithms for separable approximations of stochastic optimization problems, Technical report, Princeton University, Department of Operations Research and Financial Engineering.

Powell, W.B., J.A. Shapiro, H.P. Simão (2002). An adaptive dynamic programming algorithm for the heterogeneous resource allocation problem. *Transportation Science* 36(2), 231–249.

Puterman, M.L. (1994). *Markov Decision Processes*, John Wiley and Sons, Inc, New York.
Robbins, H., S. Monro (1951). A stochastic approximation method. *Annals of Math Stat*. 22, 400–407.
Sutton, R., A. Barto (1998). *Reinforcement Learning*, The MIT Press, Cambridge, Massachusetts.
Topaloglu, H., W.B. Powell (2002). Dynamic programming approximations for stochastic, time-staged integer multicommodity flow problems, Technical report, Princeton University, Department of Operations Research and Financial Engineering.
White, W. (1972). Dynamic transshipment networks: An algorithm and its application to the distribution of empty containers. *Networks* 2(3), 211–236.
White, W., A. Bomberault (1969). A network algorithm for empty freight car allocation. *IBM Systems Journal* 8(2), 147–171.

A. Ruszczyński and A. Shapiro, Eds., *Handbooks in OR & MS, Vol. 10*

Chapter 10

Stochastic Programming Models in Energy

Stein W. Wallace

Molde University College, Servicebox 8, NO-6405 Molde, Norway

Stein-Erik Fleten

Department of Industrial Economics and Technology Management, Alfred Getz v. 1, Norwegian University of Science and Technology, NO-7491 Trondheim, Norway

Abstract

We give the reader a tour of good energy optimization models that explicitly deal with uncertainty. The uncertainty usually stems from unpredictability of demand and/or prices of energy, or from resource availability and prices. Since most energy investments or operations involve irreversible decisions, a stochastic programming approach is meaningful. Many of the models deal with electricity investments and operations, but some oil and gas applications are also presented. We consider both traditional cost minimization models and newer models that reflect industry deregulation processes. The oldest research precedes the development of linear programming, and most models within the market paradigm have not yet found their final form.

Key words: Stochastic programming, energy, regulated markets, deregulation, uncertainty, electricity, natural gas, oil.

1 Introduction

The purpose of this chapter is to discuss the use of stochastic programming in energy models. This is not a well defined topic. Let us therefore start by outlining what this chapter is and what it is not. First, this is not an annotated bibliography. Its purpose is to help the reader see where stochastic programming can be used, and to point to relevant existing literature. We do not attempt to be complete in our references, only to help the reader find good starting points. We shall discuss both existing models and the potential for new arenas.

Then, what shall we understand by the reference to energy models in *stochastic programming*? Generally, stochastic programming refers to a problem class, and not to the choice of solution procedures. Many of the models in this class can be solved both with tools from mathematical programming and as stochastic dynamic programs (SDPs). This book is about stochastic programs solved with tools from mathematical programming. However, the view we have taken in this chapter is that we cannot include or exclude interesting models solely on the basis of what solution method the authors have chosen. Hence, if an existing model represents a stochastic dynamic decision problem which can be formulated as a stochastic program, we include it irrespective of whether it is solved with methodology from mathematical programming or set and solved as an SDP.

Furthermore, to have made the point, this chapter is not about operations research and energy. This ought not to affect our models too much, as we are of the opinion that most real decision are made under uncertainty, but it will affect our referencing to the literature.

As part of the preparation for this chapter we had the privilege of reading a text, which for our field, is very old. Massé (1946) authored two volumes on hydro scheduling. The books are based on work performed before and during World War II. Of course, he does not discuss stochastic programming as such—the term was not invented at the time—but he discusses models and methodology that would fit the premises of this chapter. It is very interesting to see how he walks his readers through some very deep arguments about why deterministic model are not good enough. He points to the fact that looking at a deterministic future is far too optimistic, and that flexibility will be disregarded.

His major point is that hydro scheduling is about releasing water such that the immediate financial gain equals the expected future value of water. The expected future value of water is presented as a function of reservoir level, present inflow (to the extent that there is memory in that process), and time (to represent seasonality). He gives optimality conditions for this case. In fact, he has a long discussion to the effect that all uses of natural resources is a tradeoff between use now and use in a stochastic future. To illustrate the use of statistics about the future, he makes the reference that if you wish to check the probability that you are alive tomorrow, you look at your present health, if you wish to know if you are alive in thirty years, you resort to statistics.

Another fact, dear to all stochastic programmers, is his pointing out that while deterministic multiperiod optimization yields decisions for all periods, a stochastic approach only yields policies or strategies.

A further major issue in the books is the objective function of the optimization. Should we maximize expected profit or expected utility (which he denotes psychological expectation in contrast to mathematical expectation)? He is concerned about some of the well known paradoxes when using expected profit, and he always refers to Borel for these examples. He is also very much concerned about risk, and strongly believes that risk will always be with us.

(He clearly had not thought of hedging in the financial markets.) He comes very close to defining an efficient frontier with expected profit on one axis and the probability of shortage of water on the other. His premise here is that the owner of a reservoir has agreements with some supplier, and that any reasonable agreement will be such that in extremely dry years, the contract cannot be fulfilled.

When decision problems are formulated and solved as deterministic problems, odd and special situations are often automatically excluded from consideration as only the expected values—the normal cases—are considered. Massé has the view that this can be dangerous, as what may appear to be a detail at the time of analysis, may later turn out to have had a major effect on the development: ''Le nez de Cléopâtre, s'il eût été plus court, toute la face de la terre aurait été changée.''

This chapter has five more sections. Section 2 is on regulated electricity markets. This is clearly the best developed area for use of stochastic programming in energy. Section 3 discusses the much newer area of de/re-regulated electricity markets. We close with two shorter sections on oil and natural gas, and a conclusion.

2 Electricity in regulated markets

This section discusses models for electricity production, thermal and hydro-based, in a setting of a regulated utility. Transmission planning and operations will not be considered. The utility can be either a single producer, or several producers that are perfectly coordinated by choice or by law. Much of the newer literature on electricity production is set in a framework of de(re)regulation and competition. We wait with this subject until the next session.

2.1 Overview of models

Many different models are used in power systems planning. A possible classification divides the different models according to the planning horizon. Long term planning models deal with investments and typically have a 15–20 year horizon. Medium term planning is done over a 1–3 year range, and deal, for example, with reservoir management. Short term planning typically deals with problems with horizons of one week or shorter, such as unit commitment and economic dispatch.

The perspective taken in these models is that of a social planner or an ideal public utility. The industry has traditionally been heavily regulated with considerable central planning. The reason for regulation is that the industry is prune to market failure; use of the transmission grid causes changes in its capacity in other parts of the network (externalities). If demand and supply is not matched at each instant, the whole or large parts of the system breaks

down; reliability is a public good. Local utilities have typically had a monopoly within their area, preventing competition.

The demand in these models is mostly portrayed as price-inelastic; a deterministic or stochastic demand is to be met at minimum cost. A standard textbook on electricity operations is Wood and Wollenberg (1996). Some of these models may still have some relevance in a market setting, because the deregulated utility often still has the obligation to serve local demand. Similarly, the utility may have committed to a particular load schedule in the spot market.

2.2 *Long term planning*

By long term planning we will normally mean planning large investments, be that building thermal units, or constructing hydro reservoirs and turbines. The starting point for such an analysis is always a projection of future load (demand). Let us first briefly see how the load is normally presented in such a setting. The starting point will a load curve for each individual day of the year—possibly split into groups of days with similar patterns. These curved will possess the common pattern of having peaks in the morning and afternoon, reflecting our way of life. The first step is to sort these curves according to decreasing loads, to achieve daily load duration curves. These curves are then added together to form a curve like the one in Fig. 1. To the left is the hour with the highest load during the whole year, and to the right the one with the lowest. For example, h is the number of hours in a year for which the load is at least L (MW).

Normally, this curve is not smooth, but step-wise, as the unit along the horizontal axis is at least one hour. To simplify the presentation, let us assume that there are only two steps, the base load period and the peak load period. What we are about to present is inspired by Murphy et al. (1982), and the first result we are to illustrate comes from that paper. Consider the simplified picture in Fig. 2

The total energy consumption in base load is given by t_2L_1, and in peak load $t_1(L_2 - L_1)$. Assume we have two technologies, for example two different

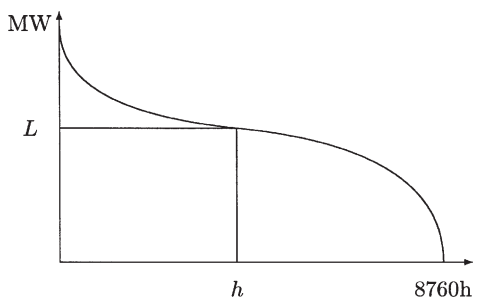

Fig. 1. Load duration curve for one year (8760 h).

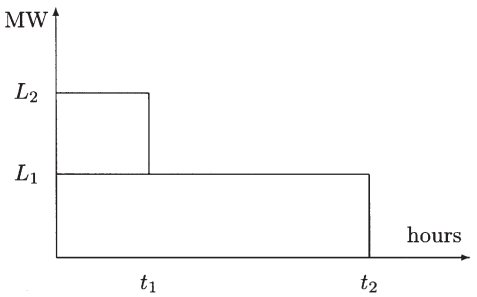

Fig. 2. Simplified load duration curve for illustrative purposes.

types of thermal units, and let us simply denote them '1' and '2'. Let f_1 be the annualized fixed costs per unit of production capacity for technology 1, and let c_1 be the operating costs per MWh, also for technology 1. For technology 2, f_2 and c_2 are similarly defined. Let x_i for $i = 1, 2$ be the installed capacity of technology i. Furthermore, let y_{ib} be the production of base load for technology i and correspondingly y_{ip} for peak load. The (deterministic) problem now becomes:

$$\min \sum_{i=1}^{2} f_i x_i + \sum_{i=1}^{2} c_i(t_2 y_{ib} + t_1 y_{ip}) \tag{2.1}$$

subject to

$$\begin{aligned} y_{1b} + y_{1p} &\leq x_1 \\ y_{2b} + y_{2p} &\leq x_2 \\ t_2(y_{1b} + y_{2b}) &= d_b (= t_2 L_1) \\ t_1(y_{1p} + y_{2p}) &= d_p (= t_1(L_2 - L_1)) \end{aligned}$$

and non-negativity (2.2)

where d_b is the load in the base period and d_p the additional demand in the peak period. The first two constraint say that production cannot exceed installed capacity of the two technologies, whereas the last two constraints express that base load and additional peak load must be satisfied.

This model helps us find the optimal investment to meet a known future demand. But, of course, future demand is not known. Hence, as a first step of making the model more realistic, let us assume that there are several possible future load duration curves. Let curve k occur with probability p_k. A straightforward recourse model now becomes.

$$\min \sum_{i=1}^{2} f_i x_i + \sum_{k=1}^{K} p_k \sum_{i=1}^{2} c_i^k (t_2 y_{ib}^k + t_1 y_{ip}^k) \tag{2.3}$$

subject to, for all k

$$\begin{aligned} y^k_{1b} + y^k_{1p} &\le x_1 \\ y^k_{2b} + y^k_{2p} &\le x_2 \\ t^k_2(y^k_{1b} + y^k_{2b}) &= d^k_b (= t^k_2 L^k_1) \\ t^k_1(y^k_{1p} + y^k_{2p}) &= d^k_p (= t^k_1 (L^k_2 - L^k_1)) \\ \text{and non-negativity} & \end{aligned} \tag{2.4}$$

where superscript k always refers to load duration curve (scenario) k. Note that the first stage variables x_i do not have a superscript k. Also note that for given values of the first-stage decisions x_i, this problem falls apart and becomes standard transportation network flow problems, one for each load duration curve, see Wallace (1986) for details.

What would we expect to get from (2.3) and (2.4) if we compare its solution to the case where (2.1) and (2.2) are solved with the expected load duration curve, that is, the case where we instead of scenarios use a 'typical' or average year and arrive at a deterministic model? Or even more importantly, how do we expect the solution to the stochastic optimization problem to differ from individual scenario problems? Different technologies for energy production will vary in some aspects. Some will have long lead times, high investments costs and low operating costs, others will have shorter lead times, lower investment costs, but at the price of higher operating costs. In a given scenario, the future demand will be known with certainty. That will tend to produce the use of a technology which perfectly matches the load. This will typically be a technology with high investment costs and low operating costs. The capacity of the unit will perfectly match the needs reflected in the load duration curve (scenario). Over-investments will never take place, and shortages will never occur. Smaller units which are more flexible, but have higher operating costs, will tend to lose out, as their qualities will not be reflected in the model. Why pay for flexibility you do not need? An example of these aspects is shown by Smeers (1990), where he discusses the relationship between coal (expensive but flexible) and nuclear (inexpensive but also inflexible).

An interesting question raised in Murphy et al. (1982) is if there is a version of the deterministic problem, such that if we solve that problem, we obtain the solution to the stochastic problem. From a general point of view, this is something we normally do not find for stochastic programs, but in this case there is a result.

Take the individual load duration curves, and multiply the duration of each block by p_k. In our simplified example, multiply all numbers on the horizontal axis by p_k to obtain $t^k_1 p_k$ and $t^k_2 p_k$. Notice that a year is 8760 h, thus $\sum t^k_2 p_k = 8760$. Create a new load duration curve from these new curves to

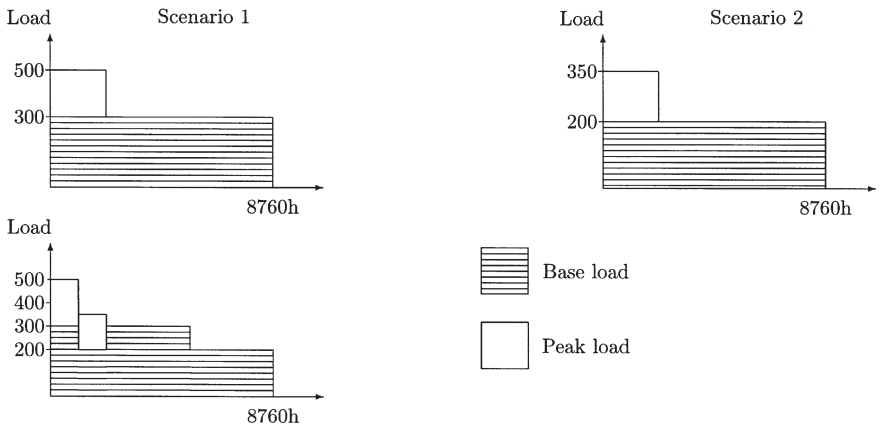

Fig. 3. Horizontal aggregation of two load duration curves.

arrive at a new aggregated curve. This is illustrated in Fig. 3. In the *horizontal* summation we have kept the patterns for peak and base load in the scenarios. This is only to make it easier to see where the columns come from. In reality, in the summation, there are four different load levels, to be treated with four parameters t_i and four load levels L_i. Generally, the number of steps in the sum equals the sum of the number of steps in the individual load duration curves.

The main result here is that if we carry out a deterministic investment analysis using this horizontally aggregated load duration curve, we obtain the same solution as if we had solved the recourse problem above. But the problem is simpler as the number of constraints saying that we cannot use capacity not installed will decrease from k times the number of technologies to just the number of technologies.

But the result is dependent on some assumptions, in particular that the operating costs do not vary with output level, implying that the c_i^k are fixed irrespective of technology i being used for base or peak load (as it is in (2.3) and (2.4)), and irrespective of scenario, so $c_i^k \equiv c_i$. But even so, this is a strong and interesting result for the two-stage case.

Sherali et al. (1982, 1984) discuss the model in greater detail, with an emphasis on cost sharing for the fixed charges f_i. This brings the problem into the realm of peak load pricing, that is, the cost of capacity is always related to those (users or scenarios) that have the maximal use of the capacity.

The load duration curve can also be approximated in a different way. An example is given in Fig. 4. In this case the resulting problem is quadratic. For a discussion, see the very early paper of Louveaux (1980).

Discrete decisions

The basic recourse model above can of course be extended in many directions. First, in many cases the variables x_i can take on only a finite

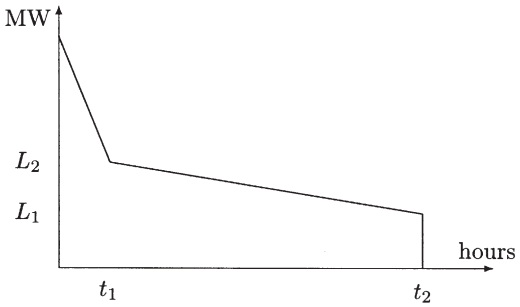

Fig. 4. Discretization of load duration curve resulting in a quadratic objective function.

number of values, which brings us into (stochastic) mixed integer programming. Bienstock and Shapiro (1988) integrates capacity expansion, supply contracts and operations planning in an electricity utility via a two-stage stochastic program with integer variables in both stages.

Multi-stage and lead times

In the same way, there are obvious possibilities in setting up a multi-stage problem, where load duration curves are revealed over time, and investments made stage by stage. Such a setup will show the importance of lead times in construction. The option to wait will favor the technologies with short lead times. This effect is not easy to capture in single- or two-stage models. Hence, this dynamic effect comes in addition to the effect discussed earlier where a stochastic model will be able to favor flexible technologies (that are never optimal in deterministic worlds). A good discussion of this problem can be found in Gardner and Rogers (1999). Flexibility is also discussed by Gardner (1996), where there is a focus on the relationship between flexibility in the financial sense, and the effects of emission control of acid-gas. The paper shows that when acid-gas emission control is added, some of the more flexible technologies lose in the competition. This is particularly true for gas combustion turbines. For further discussions of emission policies, see Manne and Richels (1991, 1995), Manne and Olsen (1996) and Birge and Rosa (1996).

Other contributions in the multi-stage setting are Manne and Richels (1978), Gorenstin et al. (1993), Dantzig and Infanger (1993), Escudero et al. (1995) and Pereira et al. (1995).

Shortage—or lost load

It is traditional in monopoly-based production planning to take it for granted that all demand must be satisfied. Qiu and Girgis (1993) take a different view, and say that since, ultimately, end users must always pay for the investments, they may be better off with a slight probability of an outage. They therefore set up a capacity investment problem where outages are priced rather than forbidden. Taking into account that scenarios (possible load duration curves) will always be somewhat subjectively chosen, and that

outages to some extent correspond to worst-case analysis, it may be very good, modeling-wise, to allow for outages at a high cost. There will always be a slight chance that something even worse than the worst scenario of the model could occur, and hence, that an outage could occur even if the model claimed otherwise.

A further discussion of long-term planning can be found in Hobbs (1995).

2.3 *Medium-term planning*

Hydro-thermal scheduling

An important problem in the medium-term scale is hydro-thermal scheduling. For hydro reservoirs, the problem is essentially to strike a balance between immediate and future (opportunity) costs of using the water. Stochastic optimization models for this problem are in daily use in hydro-dominated systems.

The following section presents the production scheduling problem. There are T time periods, or *stages*, as illustrated in Fig. 5.

Periods are time intervals between stages, which are discrete points in time. The first period is deterministic. To simplify exposition, the problem is formulated for a producer with only one reservoir.

The producer is operating an ongoing business with an infinite future. We would like to avoid end effects, which are distortions in the model decisions due to the fact that the model has a finite horizon, whereas the real business problem has an infinite horizon. For example, if in the model the value of the reservoir at the end of the model horizon is too low, say equal to zero, then the end effect would be that too much water is sold in the last stage. There are several alternatives for this problem. One is choosing the date of stage T such that it makes sense to constrain the reservoir to be either empty or full at that date, i.e., in the spring before snowmelt, or in the fall before winter sets in. Another alternative requires estimating the end-of-horizon value of water in the reservoirs from a more aggregate model with a longer time span. Third, one can choose the time horizon of the model to be long enough to make the first stage decisions unaffected by the choice of horizon reservoir value function. See Grinold (1980) for an approach to dealing with end effects in general energy models. Grinold's observation is that the dual variables at the end of the horizon should be in a steady state, and suggests introducing dual constraints (i.e., primal variables) to ensure this. Another way of achieving steady state dual variables in the reservoir management problem is described by Lindqvist (1962). It involves starting with a guess for the dual variables, which in this case equals the incremental values of stored water in the

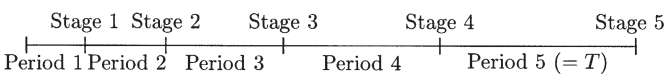

Fig. 5. Time scale example.

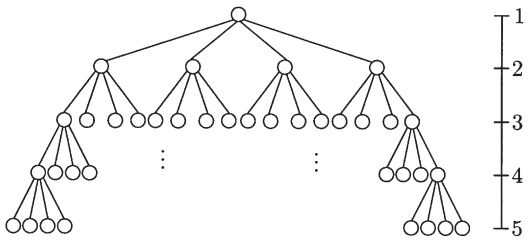

Fig. 6. Event tree and time scale example for $T = 5$. The nodes represent decisions, while the arcs represent realizations of the uncertain variables.

(equivalent) reservoir, and iterating over the last year of the planning horizon replacing the guesses with the water values obtained for one year before the end of the horizon.

The stochastic variables are inflow, and demand δ. Scenarios are possible histories of realizations of the stochastic variables up to the end of the horizon. The event tree in Fig. 6 shows how the uncertainty unfolds over time. A scenario in the event tree is a path from the root node to a leaf node. Each node n represents a *decision point*, or equivalently a *state*, corresponding to a realization of the random variables up to the stage of state n, denoted $t(n)$. The root state is $n = 1$, and scenarios are uniquely identified by states at the last stage, belonging to the set $\mathcal{S}$. The set of all states is denoted $\mathcal{N}$. The states have unconditional probabilities P_n, satisfying $\forall\, t \sum_{n|t(n)=t} P_n = 1$. Every state except the root has a parent state $a(n)$. Let stage t decisions (for period t) be made *after* learning the realization of the stochastic variables for that period.

The inflow process is multidimensional and has strong seasonal components. The main bulk of inflow to reservoirs in North America and northern Europe comes during spring, whereas in winter the precipitation accumulates as snow. Forecasting the inflows and capturing the structure of the processes and their degree of predictability is of vital importance to hydro scheduling models. This issue is discussed by Tejada-Guibert et al. (1995).

The decision variables are reservoir discharge, v_n, spill, r_n, and reservoir level l_n. Each variable in the problem is indexed by the state to which it belongs. Power generation is generally a nonlinear function of the height of the water in the reservoir and the discharge, and could be non-convex. In our exposition we disregard head variation effects, and assume that generation is proportional to flow through the station, ρv_n, where ρ is the constant hydro-plant efficiency. In practice however, head variation effects can be significant, particularly when balancing reservoirs with different characteristics. If a downstream reservoir has little storage capacity, then keeping its head up in order to maximize efficiency may lead to increased risk of spilling water if inflow increases too rapidly. This area needs further research.

Let $V_L(l_n)$ be the value of the reservoir at the end of the horizon as a function of the reservoir level. This function must be specified to avoid end effects. If a long term scheduling model is available, V_L may be extracted

from this model, e.g., in the form of incremental value of stored water in reservoirs.

It is assumed that there is no direct variable cost of hydro production. Thermal generation is represented by p_{in}, for energy generated by unit $i \in \mathcal{I}$ in state n. This incurs variable cost $FC_i(p_{in})$, a convex function of p_{in}. It consists mainly of fuel costs, and is usually modeled as linear, piecewise linear or quadratic. Nonconvex cases are plausible, for example in the case of a unit that can be fed multiple types of fuels. With a typical minimum time resolution of one week it is reasonable not to include startup costs for thermal generation. Let γ be a discount interest rate, N_t is the number of years (in fractions of years) to period t,

$$\min_{p,v} \sum_{n\in\mathcal{N}} \left[P_n(1+\gamma)^{-N_{t(n)}} \sum_{i\in\mathcal{I}} FC_i(p_{in}) \right] - \sum_{s\in\mathcal{S}} P_s(1+\gamma)^{-N_T} V_L(l_s) \quad (2.5)$$

The hydro reservoir balance is

$$l_n - l_{a(n)} + v_n + r_n = v_n, \quad (2.6)$$

The demand constraint reads

$$\rho v_n + \sum_{i\in\mathcal{I}} p_{in} \geq \delta_n \quad (2.7)$$

Time-dependent upper and lower limits on release and reservoir level are imposed using

$$\underline{v}_{t(n)} \leq v_n \leq \overline{v}_{t(n)}, \quad (2.8)$$

$$\underline{l}_{t(n)} \leq l_n \leq \overline{l}_{t(n)}, \quad (2.9)$$

$$r_n \geq 0, \quad (2.10)$$

for $n \in \mathcal{N}$ and with initial reservoir level given.

Load curtailment is sometimes modeled as an extra thermal unit having a marginal cost equal to an estimate of the marginal cost of unserved energy.

In many systems the transmission system limits the opportunities for hydro-thermal scheduling. There will be cases when transferring more electric energy from one node of the system to another will not be possible. Electricity flow in transmission networks is governed by Kirchoff's laws and is limited by line capacities. These physical phenomena must be taken into account when including transmission constraints in the scheduling problem. Accurate mathematical representations of these features typically involve nonlinear and nonconvex equations with phase angles, voltages and power flows.

For hydro-thermal scheduling, the network constraints are usually linearized, however, into linearized (DC) power flow (Wood and Wollenberg, 1996). This problem has been modeled by Gorenstin et al. (1992).

Some schedulers feel more comfortable using deterministic models for this problem. An important question in this context is what is the value of stochastic optimization? Starting with Massé (1946), many researchers argue that the stochastic aspects of the problem are important, and their neglect should result in some loss. Deterministic solutions will underestimate the true costs and the risk of spilling water, and deterministic models will not see any value in waiting with releasing water in order to learn more about future demand and/or inflow. The degree of cost underestimation depends on the problem, e.g., Tejada-Guibert et al. (1995) show that it depends, for a given system, on the demand and the severity of penalties on shortages. Philbrick and Kitanidis (1999) show that the performance of deterministic solutions is particularly poor for reservoir systems with limited storage capacity.

The typical horizon for hydro scheduling is a few months to a few years. A typical length of the first time step ranges from one week to a month. The hydro scheduling model gives signals to hydro unit commitment via marginal values of stored water in the reservoirs and/or via total generation during the first week.

2.4 Short term planning

Hydro unit commitment

In the hydro unit commitment problem the scheduler must determine what turbine units to run in each time step (hourly or shorter) the next day or week, and at which output level the running units should generate. Generating stations may have several turbines each and may be coupled by their location along the same river system. Turbines incur startup costs and the generation of each station varies nonlinearly with the volume and with the net head[1] of the hydro discharge. This problem can be formulated as a large mixed-integer nonlinear programming model with an objective of minimizing cost subject to meeting a given demand. The cost is measured in terms of the volume of water used or as the opportunity cost associated with that volume, i.e., using the marginal values of stored water in each reservoir coming from medium-term generation planning models. Stochasticity in such models may reside in load, inflow, unit availability or cost. Of these, load is considered most important, since it is temperature dependent, and temperature cannot be predicted with a precision better than a few degrees even just a few hours in advance. Other factors such as hydro inflow are accurately predictable on such short time scales. The stochastic hydro unit commitment problem has been studied by Philpott et al. (2000).

[1] The net head is the difference between the height of the water immediately before and immediately after the power station.

If the net head varies significantly on short term with upper reservoir storage level, which is usually the case for small upper reservoirs and/or small maximum head, there may be significant gains from letting reservoir levels cycle between high and low levels during the course of the day or week. Such problems call for global optimization techniques, such as in Feltenmark and Lindberg (1997). Other complicating issues are time delays from flows leaving a station to arriving downstream, a delay that may depend on the flow rate. Further, the generation function is not always concave in discharge, making the standard approach of replacing it with a piecewise linear function problematic. Discharge ramping and reservoir level constraints due to navigational, environmental and recreational requirements add to the difficulty. To avoid end effects, horizon unit states must be constrained or valued. In the case of time delays, in-transition flows at the horizon must be dealt with similarly.

Thermal unit commitment

In contrast to hydro unit commitment, where there are power stations situated along rivers, the problem here is characterized by higher startup costs and restrictions preventing thermal stress. When starting up a coal fired unit there is a time delay before the unit is available for generation. The task is to find a cost-minimal schedule, and a production level, for each generating unit over time. The problem is to decide which units will be on/running, and how much units that are on (committed) will produce. As mentioned above, the load that is to be met and the availability of the generating units are uncertain parameters affecting the problem.

This has been modeled as a mixed integer stochastic program and has been explored by Takriti et al. (1996), Carpentier et al. (1996), Dentcheva and Römisch (1998) and Caroe and Schultz (1998). See also Gröwe-Kuska et al. (2002), Nowak and Römisch (2000) and Gollmer et al. (2000).

Having $|\mathcal{I}|$ thermal and $|\mathcal{J}|$ hydro plants, the objective is to minimize the operating costs. Let u_{in} be unit states, i.e., a binary decision variable that represents whether thermal unit i is running ($u_{in} = 1$) or not ($u_{in} = 0$) in state n. Operating costs consist of fuel cost $FC_i(p_{in}, u_{in})$ and startup costs $SC_{in}(u_i)$ for thermal units. Startup costs may depend on the amount of time the unit was down before startup. Thus u_i is a vector consisting of u_{in} and unit states u_i for one or several predecessor states of n. Hydro plants only contribute to the objective function with the water value $V_S(l_s)$ at the end of the time horizon T, where the vector $l_s = [\{l_{js}\}_{j\in\mathcal{J}}]$. This water value function approximates the objective value for the problem (2.5)–(2.10) with T going to infinity. The objective function reads:

$$\min_{u,p,v} \sum_{n\in\mathcal{N}} \sum_{i\in\mathcal{I}} FC_i(p_{in}, u_{in}) + SC_{in}(u_i) - \sum_{s\in\mathcal{S}} V_S(l_s). \tag{2.11}$$

The local demand δ_n has to be satisfied in all states:

$$\forall\, n \in \mathcal{N}: \sum_{i \in \mathcal{I}} p_{in} + \sum_{j \in \mathcal{J}} \rho_j v_{jn} \geq \delta_n, \tag{2.12}$$

where $\rho_j v_{jn}$ is the generation of hydro unit j. Additional constraints follow (index i omitted since all constraints are for each unit i): The output of a unit should be zero if the unit is offline; otherwise it should be between a lower and an upper bound ($\underline{p}$ and $\overline{p}$):

$$\forall\, n \in \mathcal{N}: u_n \underline{p} \leq p_n \leq u_n \overline{p}. \tag{2.13}$$

Further single-unit constraints are minimum up- and down-times and additional must-on/off constraints. Minimum up- and down-time constraints are imposed to prevent thermal stress and high maintenance costs due to excessive unit cycling. Denoting by $\underline{\tau}$ the minimum down-time of the unit, the corresponding constraints are described by the inequalities (temporarily switching to time subscripts instead of states):

$$u_{t+\tau} + u_{t-1} - u_t \leq 1, \tag{2.14}$$

for all $\tau = 1, \ldots, \min\{T - t, \underline{\tau} - 1\}$. Analogous constraints can be formulated for describing minimum up-time restrictions, see e.g., Gröwe-Kuska et al. (2002).

A reserve margin $r_t \geq 0$ is often imposed via reserve constraints

$$\forall\, n \in \mathcal{N}: \sum_{i \in \mathcal{I}} (u_{in} \overline{p}_i - p_{in}) \geq r_{t(n)} \tag{2.15}$$

to ensure that the model recommends an on-line capacity that exceeds the predicted load, giving a 'spinning reserve'. This is used, particularly in deterministic models, to avoid an energy imbalance resulting from the unexpected failure of a generating unit or an unexpected increase in load, which may cause very costly brownouts or blackouts. Carpentier et al. (1996) discuss the relationship between spinning reserve in a deterministic model of the problem compared to a stochastic model without spinning reserve, and uses rolling horizon optimization to arrive at an optimal level of reserve margin. The models consider uncertainty in generator availability.

Deterministic approaches

Unit commitment is usually solved as a deterministic large scale mixed integer program (Sheble and Fahd, 1994). It is therefore interesting to learn about qualitative differences between stochastic programming solutions of the

unit commitment problem and deterministic solutions. A priori we can state that deterministic solutions will be characterized by extensive use of large plants with high start-up costs, with relatively few starts. SP solutions on the other hand, will typically use smaller units and will involve more startups of flexible but possibly high-marginal cost plants such as gas fired units. Deterministic models will know exactly how much power is needed at any time and can thus plan to run low fuel cost plants at high output for long periods of time. The gains that the model sees from such scheduling will outweigh the high startup costs that typically come with such plants.

In deterministic models, a common approach to the uncertainty regarding generator failure is to "derate" the units' maximum generation rate according to the probability of availability. However, this will underestimate the expected operations cost and the probability of load being larger than peaking capacity. Recognizing such operations cost underestimation, a class of models known as production costing models have been developed (Wood and Wollenberg, 1996). The purpose of these models is a more accurate estimation of production costs by simulating and/or optimizing the dispatch of generation under uncertainty of load and generator outages. Production costing models are used both in long-term and operations planning. These models are conceptually not much different from SP-based generation planning models, in fact, good SP models lessen the need for separate production costing and reliability models. SP contributions in this class have been made by e.g., Bloom (1983), Pereira et al. (1987, 1992) and Hobbs and Ji (1999). See also the review by Hobbs (1995).

Economic dispatch and optimal power flow

Optimal dispatch of power under uncertainty has been considered by e.g., Bunn and Paschentis (1986), Gröwe and Römisch (1992) and Gröwe et al. (1995). The models have a short time horizon, usually a day, with hourly or finer resolution. The unit commitment schedule (the unit states u_{in}) is regarded as given, and the problem is to determine a generation schedule (in the p_{in} variables) that minimizes operating costs ($FC_i(p_{in}, u_{in})$) and satisfies the demand.

Since this problem is near real-time operations, it is meaningful to include transmission issues. This is considered by Pereira et al. (1987), who solve a two-stage optimal power flow problem.

2.5 Solution methods and computations

We focus on the modeling process and not the solution methods. However, most of the SP energy papers focus on the solution method used to solve the model, not the modeling process itself. Still, there is a relationship between research on solution methods and model development, because models tend to be developed only if there is hope of solving the model. Thus, as new solution algorithms are published, new models are reported solved using twists of the

state of the art algorithms. Thus solution methods are discussed briefly in this subsection.

For hydro planning problems, stochastic dynamic programming has been used for a long time; an early reference is Massé (1946). For surveys see Yakowitz (1982), Yeh (1985) and Stedinger (1998). These methods have also been used in unit commitment and expansion planning. However, a well known problem with these methods is the curse of dimensionality. To use them, it has been necessary to aggregate and/or decompose the problems before solving them. An example of this is the aggregation of several hydro reservoirs and connected power stations into a single equivalent reservoir/power station pair. Relatively good heuristics have been developed for supporting the aggregation/de-aggregation approximation process. Important applications are presented by Terry et al. (1986) and Gjelsvik et al. (1992). A somewhat different approach for the multireservoir problem, using decomposition, is presented by Turgeon (1980) and Sherkat et al. (1985). Still, methods that could handle multidimensional problems having many state variables, were in demand. In the late 1970s, authors at Stanford University (Birge, 1985) began experimenting with nested Benders' decomposition, and in electricity models this was used and refined by Pereira and Pinto (1985), Jacobs et al. (1995) and Morton (1996).

This method was able to solve multidimensional state type problems, but was unable to match SDPs time decomposition abilities with respect to solving stochastic programs having many stages. Nested Benders' decomposition works on a scenario tree whose number of nodes explodes with the number of stages, and the size of the problem to be solved is proportional to the number of such nodes. For a comparison of the main algorithms on reservoir management problems see Archibald et al. (1999).

With this background, the algorithm of Pereira and Pinto (1991) created a lot of interest in the energy optimization community. Termed stochastic dual dynamic programming (SDDP), it effectively combines the state-time decomposition features of dynamic programming and the benefits of nested Benders' decomposition. It represents a very important expansion of nested Benders' decomposition using two important concepts of cut sharing and sampling. Commercial software based on this algorithm is in widespread use.[2]

In a deregulated setting, spot market prices become important as input to power scheduling models. Assuming the price-taker case, prices are exogenous and can be treated by ordinary linear stochastic programming.[3] Prices are autocorrelated, so the current price carries information about the likely future outcomes of price. Thus it must be treated as a state variable, which posts a problem in SDDP, because the future cost function is no longer convex in all state variables. (As is well known and probably shown in earlier chapters in

[2] The Stanford group under G. B. Dantzig worked out a similar decomposition/sampling algorithm based on importance sampling approximately at the same time (Dantzig, 1989).

[3] Ordinary is meant in contrast to game-theoretic approaches.

this volume, the recourse cost function, or future cost function, is concave in changes to objective function coefficients and convex in changes to right hand side coefficients.) Thus the future cost function can no longer be supported by cuts. This issue is discussed by Gjelsvik and Wallace (1996), who introduce an algorithm that can handle stochastic prices by not sharing cuts across price states. During the course of the algorithm the future cost function (a function of all state variables) is built for each price state at each stage. Pereira et al. (2000) approach the issue of stochastic prices causing nonconvex recourse functions by using a cut classification scheme.

Stochastic unit commitment problems are not yet in daily use, as far as we know, and for algorithmic work on these stochastic integer problems we refer to Gröwe-Kuska et al. (2002) and to other chapters in this volume.

3 Electricity in deregulated markets

This section discusses issues related to electricity production under market conditions. Researchers have studied hydrothermal scheduling, risk management, unit commitment and bidding problems in deregulated market settings. Assumptions on market form, institutional and market design and existence of derivative markets vary.

At the time of writing, electricity markets are still in transition from the old regulated regime, motivating the development of hybrid models where there is both a demand constraint and a wholesale market. The local load is to be met at each instant, but the producer can choose to serve this load by his own production capacity or by buying capacity in the market. The producer may also produce more than the local load, selling the surplus in the market. See f.ex. Takriti et al. (2000) and Gröwe-Kuska et al. (2002). A hybrid approach may also be motivated by "imperfections" causing constraints on how much the company can buy or sell in the wholesale market, or by a significant difference between the price of buying electricity in the spot market and the price of selling to the spot market.

3.1 System-wide models

Some models try to capture aspects of the whole electricity system, having the power price as endogenous variable, i.e., as a result of matching supply and demand. Examples of such models are MARKAL, MPS, which focuses on markets with a large share of hydro power, and BALMOREL. Stochastic programming efforts related to these models are reported by Fragniere and Haurie (1996), Botnen et al. (1992) and Hindsberger (2003, papers F and G). These models serve the needs of utility planners and policy makers in that they can derive scenarios of market prices of electricity. The major advantage of such models is that they capture the specific aspects of electricity and that the scenarios generated are consistent with the assumptions underlying analyses

regarding e.g., future system-wide capacity and emission allowance policies. These models are all developed for regulated markets. However, they have become very popular for generating price scenarios in deregulated markets. The reason is that in perfect markets, price will equal (long term) marginal cost. A regulated market, as described in Section 2 of this chapter, is normally based on a policy of efficiency and cost minimization so as to achieve exactly the same result—price equal to long term marginal cost. Care must be taken, however, so that the price-scenario generation does not take the form of pure scenario analysis, that is, a large number of "What-if"-questions on the external events. That would result in prices of electricity being too low, as each path of investments would be done under full knowledge of the future, underestimating the need to invest in (expensive) flexibility. The modern versions of MARKAL, like the one referenced above and Kanudia and Loulou (1998), take this into account.

It is also important to remember the setting here. These models can be used to generate price-scenarios (consistent with external events) for a *small* market participant who does not herself affect the market. A policy-maker can view the price scenarios as the *result* of her actions, but cannot use them to make other policy decisions. That would create a logical loop.

These models could of course also have been discussed in the previous section on regulated markets, as some of them represent long term stochastic investment models in the light of random demand and emission policies.

3.2 The role of futures markets

The electricity markets are developing into regional commodity markets. This can be seen in the contract market where there is decreasing use of complex physical sales contracts and increasing use of standardized financial contracts. As these derivative markets mature, they will serve an important role in risk sharing and in giving economic signals to investment and operations planning.

A common commodity contract is a forward contract, which entitles the buyer of a contract the difference between the spot price and the agreed contract price in the settlement period of the contract. In some markets these contracts are known as contracts for differences (CfDs).

If the commodity can be stored, such as coal, oil and gas, the contract price will be closely related to storage costs and interest rates, due to the arbitrage opportunities that would otherwise be present. If futures prices are higher than current spot prices compounded at the risk free rate plus storage costs, an arbitrageur can buy a unit of the good (at price S_0), finance this with a bank loan, and short sell a futures contract. The storage costs may include opportunity costs associated with the operational benefits of having the commodity immediately available in storage (convenience yield). At the time of maturity, the arbitrageur sells the good (at S_T), pays back the loan

$(-e^{rT}S_0)$, pays storage costs (C) and settles the futures contract $(F_T - S_T)$. The safe future value of this project is

$$FV = S_T - S_0\, e^{rT} - C + F_T - S_T = F_T - S_{\,0}e^{rT} - C.$$

If this value is positive, the futures price F_T is too high compared to the current spot price since the arbitrageur actually can make money on this deal. Clearly, this value must be nonpositive in a reasonable model of price dynamics. Similarly, a speculator holding the commodity in stock may arbitrage on temporarily reducing his storage by selling a unit of the commodity and buying a futures contract. This means that the above future value must be nonnegative, leading to $F_T = S_0\, e^{rT} + C$ for commodities that are stored.

Electricity can to a certain extent be stored as potential energy in reservoirs. Hydroelectric producers are thus in a position to arbitrage between the spot and futures markets using their reservoirs, hydro stations and possibly pumps. If aggregate reservoir capacity is large then such behaviour can be expected to influence the pricing of electricity futures relative to spot. In many power systems, however, the aggregate reservoir capacity is low compared to aggregate system capacity, and the price will be determined by short term equilibrium of supply and demand for the contract. Supply and demand are driven by hedging and speculation, where selling hedgers are the producers and buying hedgers are power marketers and large industry. Speculators take positions on either side depending on their capacity and willingness to take risks and their expectations on the future spot price or the future movement of the contract price.

Regardless of its determination, the contract price represents the current market value of future delivery of the commodity. This is obviously important for investment and operational planning. If an electricity company is considering an investment that will give a certain production capacity in a future time period, the current value of the revenues coming from the use of that capacity is given by multiplying the capacity with the forward price for settlement in the same future period, and discounting to present using the risk free rate of interest. This is a simple valuation procedure that will value production resources in a way that is consistent with how the market prices contracts. Rational decisions based on such valuation will contribute to maximizing the market value of the firm owning the production assets.

Valuation of future production is needed in stochastic programming models in energy. It is what many such SP models are about. These models are based on describing the uncertainty in the form of scenarios of the spot price of the commodity. However, basing the scenarios on forecasts of spot prices will not give a valuation that is consistent with the market. Price scenarios need to be adjusted for risk in order to give consistent valuation;

they must be adjusted so that the values of derivatives as calculated in the scenario tree are the same as can be observed in the market for futures, options and other contracts.[4] Once this adjustment is made, the appropriate discount interest rate to use is the risk free one, and the SP is now in the position to value the decision flexibility using a price of risk that is consistent with the market.

Adjusting for risk is necessary because expected spot prices in future periods are generally different from forward prices for the same future periods. This in turn is due to the limited capacity or willingness of speculators to trade on the mentioned difference, called the *risk premium*.

This quantity, defined more formally as $E[S_T] - F_T$ where $E[S_T]$ is the expected spot price at future time T, and F_T is the current forward price for delivery at time T.[5]

Note that if one makes optimal decisions based on price forecasts (i.e., on $E[S_T]$), the expected profit is maximized. If one makes decisions based on risk adjusted prices (i.e., on F_T), the value of cash flows is maximized. Thus, one cannot have profit maximization and shareholder value maximization at the same time. Only the latter will maximize the value of the firm.

When constructing scenario trees (or more generally, when modeling the stochastic processes involved) we must therefore make sure that the path of the expected spot price in the tree matches that of the term structure of futures prices, and that the path of the standard deviation of price returns in the tree matches the term structure of volatility (which has to be estimated, see Hull (2000)). One should possibly also match higher moments and dynamic properties of commodity prices such as mean reversion. An approach for scenario generation based on matching such statistical properties is described by Høyland and Wallace (2001). Alternatively, one may prefer modeling the stochastic processes as stochastic integrals, i.e., a parametric approach. This would have the advantage of capturing the theoretical developments in the financial commodity pricing literature, as in e.g., Schwartz (1997) and Lucia and Schwartz (2000). In this case, the scenario trees can be built using the approach of Pflug (2001) or by discretizing the continuous stochastic processes directly as in Hull (2000).

Example 1. Let us give a very simple example of how market data can be used to extract useful information for a stochastic program, and in particular to obtain an understanding of the world in which stochastic programs operate.

Assume we are facing an uncertain future price for electricity. Assume that presently, 1 MW delivered in the next period costs 100. In the next period, we know (e.g., by estimation), that the price for immediate delivery will

[4] At least approximately the same. Current market prices will change in the future. See Hull (2000, Chap. 18.6).

[5] This is positive for most commodities most of the time (Hull, 2000).

increase to 125 or decrease to 70. Each of these cases occurs with a true probability of 50%. In the next period, our production will be worth 2000 if prices go up, and 1500 if they go down. Hence, the expected value of our production, using the true probabilities, is 1750. However, we should note that $100 \neq 0.5 \times 125 + 0.5 \times 70$. Let us disregard discounting, and assume that there is a risk free asset that costs 100 now, and pays 100 in any state in the next period. We then have two instruments (price of electricity delivered in the next period and a risk free asset) in a world with two states, and we can set up the equations for the state prices π_1 and π_2.

$$100 = 125\pi_1 + 70\pi_2$$
$$100 = 100\pi_1 + 100\pi_2$$

which yields $\pi_1 = 0.5455$ and $\pi_2 = 0.4545$. Hence, the market value of our production equals

$$0.5455 \times 2000 + 0.4545 \times 1500 = 1773,$$

above its expected value using true probabilities. Someone understanding markets better could obtain arbitrage by for example buying our production for above its expected value (according to the stochastic program), say for 1751, which should make us happy, and then sell it in the forward market for its true market value, 1773, to obtain a risk free profit (arbitrage) of 22.

Again, the purpose of this example is to observe a fact about market values. To maximize the market value of our electricity production we need to use risk adjusted probabilities, and not the physically correct ones. Of course, often we do not know the true ones either, but that is not the point here—the point is that the relevant probabilities to look for are the risk adjusted ones. And they are to be found in the market prices of contracts, not in historical spot price data. This also means that if we use the true probabilities in a stochastic program, we shall not be maximizing market value, and hence, we open up our business for speculation based on the true values of risk.

As a theoretical digression, note that transforming a scenario tree, or a stochastic process with an associated probability measure $\mathcal{P}$, for the true or forecasted spot price, to a scenario tree matching the term structure of futures prices and volatility, is equivalent to changing the probability measure into an equivalent martingale measure $\mathcal{Q}$. The existence and uniqueness of such a probability measure can be analyzed via stochastic programming by setting up the problem of hedging a general contingent claim (contract) in the original ($\mathcal{P}$-measure) scenario tree. This has been done by e.g., Ross (1977), Kreps (1979), Naik (1995) and King (2002).

3.3 Energy bidding

The bidding problem can be viewed as a short term optimization problem in which the market participant offers to buy or sell capacity to the market in the form of price-quantity pairs for given time intervals that typically are 30 min or one hour long. A market operator collects such bids and calculates clearing prices and quantities for each node or zone in the network, resulting in a dispatch for the system. The price clearing process aims at maximizing the sum of consumer and producer surplus as implied by the bids, subject to transmission constraints, reserve constraints and possibly other technical constraints. In sending their bids, individual market participants try to maximize profits resulting from the dispatch (thus buyers minimize the cost of buying electricity).

The exact setup regarding market structure and market rules differs from market to market. The Electricity Pool of England and Wales was the first to be established, in 1988, and has served as a model for much of the restructuring worldwide, e.g., in Australia, New Zealand and parts of Latin America and North America. These countries use a centralized dispatch and pricing mechanism, called an electricity pool. The second country to deregulate was Norway in 1991. The Norwegian electricity trade is much more decentralized and its structure has been adopted by the other Nordic countries and in some aspects by California. Participation in the organized markets is voluntary and there is demand-side bidding. This is called a bilateral market.

Some markets have only a few or even only one round of bidding, and after these rounds the generator is assigned a generation schedule for the near future (e.g., for the next 12–36 h). In this situation, after the market operator has announced the dispatch, the traditional unit commitment models that include a demand constraint become relevant again. Furthermore, the bidding problem, i.e., determining optimal bids to send to the market operator, becomes a nontrivial task that can be supported by optimization models. Nowak et al. (2000) study this problem and present an integrated stochastic unit commitment and bidding model.

Neame et al. (1999) consider the bidding problem for a price taker in an electricity pool type market. The bids are required to be in the form of a piecewise constant increasing supply curve, i.e., a set of price-quantity pairs. If the bids could be in any form, price-taking generators maximize their profit by bidding according to the marginal cost of generation. However, since marginal cost is not generally a piecewise constant curve having a finite number of price-quantity pairs, the generator needs to optimize his bid curve. The authors study this problem and finds, among other things, that it is nonconvex. For special cases dynamic programming algorithms for computing the globally optimal bid are presented.

Anderson and Philpott (2002a) consider bidding problems in day-ahead markets for producers having market power under varying assumptions on

the allowed smoothness of the bids, and on whether there is uncertainty in demand only or also in the supply functions offered by competing generators. An important vehicle in the analysis is the "market-distribution function" Ψ encapsulating uncertainty in demand and competitor behavior. Let $\Psi(q,p)$ be the probability of not being fully dispatched by the market operator if quantity q is offered at price P. The generator is said to be fully dispatched if the whole offer was knocked down in the auction, i.e., the market operator declares it will use all of the quantity q offered. The other cases are those of not being dispatched, and of being partially dispatched if a fraction of the quantity is cleared. The problem is to find a supply curve $s = ((q(a),\ p(a), 0 \le a \le A)$ where $q(a)$ is the quantity the generator is willing to supply at a corresponding price $p(a)$. This curve is assumed to be continuous with $q(\cdot)$ and $p(\cdot)$ non-decreasing in the parameter a. If $C(q)$ is the cost associated with generation of q, the payoff resulting from a dispatch (q,p) is $R(q,p) = qp - C(q)$. If the generator has sold a quantity Q at price f via physical or financial contracts for delivery in the period in question, the payoff is $R(q,p) = qp - C(q) + Q(f-p)$. With a continuous market distribution function the expected payoff becomes a line integral given by

$$V(s) = \int_s R(q,\, p)\ \mathrm{d}\Psi(q,\, p),$$

see Anderson and Philpott (2002b). With certain (nonrestrictive) differentiability and monotonicity properties for Ψ, the problem can be formulated with respect to the parameter a as follows:

$$\begin{aligned}
\max\ & \int_0^A R(q,\, p)\left[\frac{\frac{\partial\Psi}{\partial q}(q,\, p)q'(a) + \partial\Psi}{\partial p(q,\, p)p'(a)}\right]\mathrm{d}a \\
\text{s.t.}\ & 0 \le q(a) \le q_M \\
& 0 \le p(a) \le p_M \\
& q'(a) \ge 0 \\
& p'(a) \ge 0
\end{aligned}$$

where q_M is the maximum capacity of the generator and p_M is some upper bound on price. This is a nonlinear optimal control problem, and the authors analyze its properties and various extensions, for example to the case where the generator is required to submit piecewise constant bid curves instead of smooth continuous curves.

3.4 Scheduling in a market

We first assume that this generation utility is not large enough to be able to influence electricity prices by changing the amount of generation capacity offered to the market. The market is liberalized, but not necessarily perfectly competitive.

We discuss how different classical power generation planning problems change in the face of liberalization.

Significant changes are necessary in traditional long and mid-term power scheduling, unit commitment and economic dispatch. Under the price taker assumption and that the utility does not have to worry about the transmission constraints, either because transmission is not the utility's responsibility or because there is sufficient capacity in the transmission grid, these changes affect both planning objectives and constraints.

First, the objective of the planning models should now be to maximize utility profits instead of minimizing overall system costs. The revenues are (hopefully) greater than the generation costs. From an optimization point of view, this may not amount to more than multiplying the objective function by -1 and maximize instead of minimize, but for the management focus the change is more profound.

Second, the demand constraint in these models becomes superfluous (except possibly in the very short run). Since utilities no longer have an obligation to serve demand by using only own generation resources, they now can use the spot and contract markets (i.e., other companies' resources) to meet customer obligations.

Third, reserve constraints, as used in unit commitment, also become unimportant for the utility. This happens because spinning reserve and other ancillary services become the responsibility of the system operator rather than the utilities collectively, or because well-functioning markets for different levels of reserve develop.

To see why the demand constraint becomes superfluous, consider the following problem, where e_t is the net sale (selling minus buying) in the spot market and π_t is the spot market price:

$$\min_{u,p,v} \sum_{n\in\mathcal{N}} P_n \left\{ (-\pi_n e_n) + \sum_{i\in\mathcal{I}} FC_i(p_{in}, u_{in}) + SC_{in}(u_i) \right\} - \sum_{s\in\mathcal{S}} P_s V(l_s) \tag{3.1}$$

$$\text{s.t. } \forall\, n \in \mathcal{N}: \sum_{i\in\mathcal{I}} p_{in} + \sum_{j\in\mathcal{J}} v_{jn} - e_n \geq \delta_n. \tag{3.2}$$

This formulation assumes that the cost of buying is the same as the income of selling the same energy volume. With a significant difference between

purchase price and sale price, the argument becomes invalid. Due to the presence of operating ranges $[\underline{p}_i, \overline{p}_i]$ for each unit, the demand constraint (3.2) may not be satisfied as an equality in an optimal solution, as one might expect from cost minimization. However, this rarely occurs in practice and we ignore this possibility. If there are no binding constraints on it, the net sale variable e_n and (3.2) can be substituted out to give the following model:

$$\min_{u,p,v} \sum_{n\in\mathcal{N}} P_n \sum_{i\in\mathcal{I}} FC_i(p_{in}, u_{in}) + SC_{in}(u_i) - \sum_{s\in\mathcal{S}} P_s V(l_s) - \pi_n \left(\sum_{i\in\mathcal{I}} p_{in} + \sum_{j\in\mathcal{J}} v_{jn} - \delta_n \right) \tag{3.3}$$

This is a model that is decomposable; one can solve for each thermal unit and for each group of hydrologically coupled hydro units independently.

The total implication for the models is that all or most constraints coupling the different generating units should be removed as the deregulation process is moving forward. The management is left with a set of decoupled subproblems for power operation planning, one for each unit or plant, instead of one big problem with the plants depending on each other to cover demand and spinning reserve.

In a liberalized market, the following tasks are most important for a generation utility: Risk management, hydro scheduling, unit commitment and bidding in the organized markets. In addition, short- and long term market analyses are important. Forecasting the future development of prices and other uncertain factors from now to several months or years into the future is important for trading and risk management. Short term forecasting of prices, loads and inflows is important for short term operational planning.

Hydro scheduling

Next, we present hydro scheduling. Gjelsvik and Wallace (1996), Fosso et al. (1999), Pereira et al. (2000) study hydro scheduling assuming perfect competition. For simplicity, we show a model with a single reservoir. For cascaded reservoirs, a multi-reservoir formulation is warranted. Since hydro plants are independent of each other under our assumptions, we omit the index j.

The release decisions for period t are taken after learning the realization of the stochastic variables for that period.

Decision variables and parameters for hydro scheduling are measured in energy units. The problem can be formulated as:

$$\max \sum_{n\in\mathcal{N}} ((1+\gamma)^{-N_{t(n)}} P_n \pi_n v_n) + (1+\gamma)^{-N_T} \sum_{s\in\mathcal{S}} P_s V(l_s) \tag{3.4}$$

$$\text{s.t.} \quad \forall\, n \in \mathcal{N}: l_n - l_{a(n)} + v_n + r_n =_n , \tag{3.5}$$

$$\forall\, n \in \mathcal{N}: \underline{l}_{t(n)} \leq l_n \leq \overline{l}_{t(n)}, \tag{3.6}$$

$$\forall\, n \in : \underline{v}_{t(n)} \leq v_n \leq \overline{v}_{t(n)}, \tag{3.7}$$

where $\underline{l}_{t(n)}$, $\overline{l}_{t(n)}$, $\underline{v}_{t(n)}$ and $\overline{v}_{t(n)}$ are lower and upper bounding parameters for reservoir level and discharge, and spill $r_n \geq 0$. Equation (3.5) is the energy balance in the reservoir, and (3.6) and (3.7) impose lower and upper bounds on reservoir level and discharge.

Using deterministic models for hydro scheduling in a market setting will lead to operating policies that essentially allocates the water to the periods with the highest prices. As in the case without markets, the spilling risk will be underestimated and true profit will be overestimated. There will be no extra release in the fall in case of extra inflow at near maximum reservoir levels, and no holding back water before the spring flood in case snow melting starts late and prices skyrocket.

Market power

Operations scheduling in deregulated markets when the operator has market power is discussed by Scott and Read (1996). Their focus is on imperfect competition due to the low number of suppliers in New Zealand. In particular, they develop a hydro scheduling model for a Cournot-type producer having the contract position as exogenously given. A multistage stochastic programming algorithm is developed to solve the optimization problem with a Cournot market equilibrium superimposed on it at each stage. A similar study by Kelman et al. (2001) reach the same conclusions as Scott and Read, namely that the more contracts the strategic generators have sold, the less incentive they have to withhold capacity in order to increase prices. A major limitation in these analyses is that buying and selling of contracts is in reality determined simultaneously with production. The players are also limited in the degree to which they can dynamically anticipate and react to opponents' strategies.

Unit commitment

Thermal unit commitment for price takers can be formulated as follows (index i omitted):

$$\max_{u,p} \sum_{n \in \mathcal{N}} (P_n \pi_n p_n - FC_n(p_n, u_n) - SC_n(u)) \tag{3.8}$$

$$\forall\, n \in \mathcal{N}: u_n \underline{p} \leq p_n \leq u_n \overline{p}. \tag{3.9}$$

Further, single-unit constraints are minimum up- and down-times and additional must-on/off constraints as explained in Section 2.

The decomposition that the liberalization induces should have profound implications for the organization of the utilities: Now each plant manager can be given responsibility for operating as she thinks is best. She can and should be supported by planning models that now sensibly only includes *local* generating units. Tseng and Barz (2002) consider such stochastic single-unit commitment problems.

In summary, we propose that generation utilities comprehensively revise their generation planning models. New models should include (stochastic) prices instead of using demand constraints and spinning reserve constraints. The problems become much easier to solve, thanks to the decoupling effects of the new markets.

3.5 *Risk management*

Basic financial theory implies that it is not necessary to hedge at the corporate level, since investors can do that on their own account. In practice, however, there are "market imperfections" that make the case for risk management, for example the fact that it is cheaper for a firm to operate in the power derivatives markets than for individual owners, due to the economy of scale in the risk management function.

Example 2. Let us illustrate the use of financial instruments on risk management to see, in a very simple world, how the instruments can change the risk picture. In Fig. 7, the first figure shows the distribution of profits from one unit of production without any financial contracts. Assume next that we sell 50% of our production in the forward market at the expected price of 100. That results in a new distribution of profits, given in the right-hand part of Fig. 7. The risk has clearly decreased (even if we are not very specific about what we mean by risk).

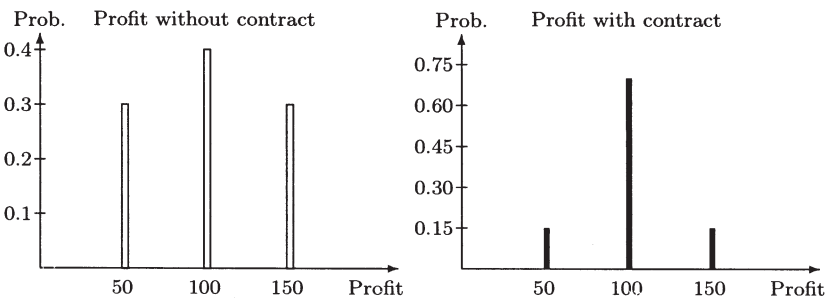

Fig. 7. Distribution of profit without and with a forward contract. The horizontal axis shows profit for one unit of production, the vertical axis probabilities.

Table 3.1
Production, prices and probabilities with and without a forward contract

Production	50	150	50	150
Price	10	10	20	20
Probability	10%	40%	40%	10%
Profit without contract	500	1500	1000	3000
Profit with contract	1000	2000	500	2500

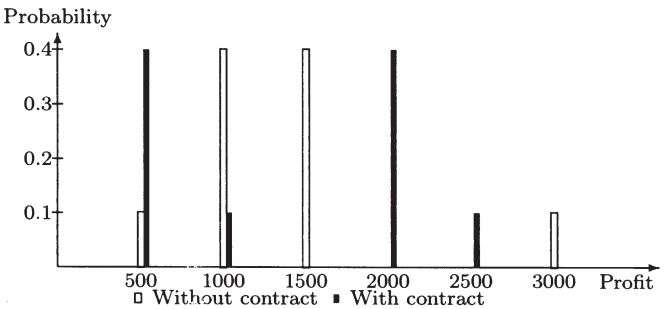

Fig. 8. Distribution of profit without and with a forward contract. The horizontal axis shows total profit, the vertical axis probabilities.

This example indicates that trading in the forward market will reduce the risk. But this may not be the case. Assume that we are facing uncertain production and uncertain prices, as outlined in Table 3.1. The most likely situations (each having 40% probability) is low production and high prices (low inflow) or high production and low prices (high inflow). But there are two other cases, representing the possibility that while we have high inflow to our reservoirs, the general picture is the reverse. Hence, there is a chance of seeing low prices and low production at the same time. The same goes for high production and high prices.

Consider the illustration in Fig. 8. The white set of columns shows the profit without any financial contracts. Assume next that we sell 100 units of production (the expected production) in the forward market at a price of 15 (the expected price). The last row in Table 3.1 shows the resulting profit. Each number is the sum of the income from the forward contract (1500) and sales or purchases in the "spot" market for what is left or what is missing relative to our forward contract. The result is the distribution in black in Fig. 8. We see that the variance has increased, and that most measures of risk will show the same. In this case, selling the expected production at expected prices increased the risk.

The purpose of these two examples is simply to illustrate the use of financial instruments for risk management, and a warning that using these markets to fix income in the future will not automatically mean reduced risks.

Mo et al. (2001) and Fleten et al. (2002) suggest that production planning and contract risk management should be integrated in order to maximize expected profit at some acceptable level of risk. However, in some circumstances (no production uncertainty or basis risk) production planning can be done independently from hedging (separation). So then it is possible to have a relatively decentralized organization, with local plant managers having much responsibility, and a centralized treasury department in charge of overall risk management. The main tasks of such a department are to speculate and hedge using derivatives in order to satisfy the goals of owners and top management regarding expected profit and risk. Of course, this requires that the relevant attitude toward risk must be expressed.

The requirements needed to invoke the separation theorem are not likely to be met 100% in practice. However, the benefits of a decoupled set of models and corresponding decentralized organizational units will probably outweigh the small theoretical gain from integrating production planning and trading.

Another argument for separating risk management is as follows: From financial theory we know that the market value of any financial contract is zero at the time it is entered into. This also holds for electricity contracts that are fairly priced, and consequently, buying a new contract will not change the market value of the electricity portfolio in question. In particular, buying and selling a range of contracts that jointly are selected in order to minimize the risk of a given electricity portfolio, will not alter the market value of that portfolio. However, operational decisions *do* change the market value of the electricity portfolio, and so generation should be allocated in order to achieve maximal market value. Any production decision that deviates from the value-maximal strategy will erode value for the owners of the generation utility. Consequently, a natural setup for the coordination of generation planning and risk management is to first schedule generation so that market value is maximized. Second, given this optimal strategy, find a set of contracts (or a trading strategy) that will minimize the risk of the total portfolio.

We model the risk management problem as follows: Given calculated (optimal) profit from hydro and thermal generation in each state in the scenario tree, summed over all plants, dynamically trade in futures and options in order to minimize some risk measure.

Let Π_t be the stochastic profit estimated from all generation activities in period t. This information must be extracted from the optimal objective function value of hydro and thermal sub-models (3.4). The scenario tree used for these sub-models is assumed to be identical to the one used for risk management.

Trading in forward contracts is modeled by the variables f_{kn}, g_{kn} and h_{kn}. Let f_{kn}, $k = 2, \ldots, T$, $n \in \{\mathcal{N} : t(n) < k\}$ be the position, measured in energy units, in state n for a contract with delivery in period k. Negative f_{kn} represent a short position in product k. Buying and selling forward contracts are represented by g_{kn} and h_{kn} (both nonnegative). Contract prices are denoted φ_{kn}, and markets are infinitely liquid and perfectly competitive.

The position accumulated in state n is

$$f_{kn} = f_{k,a(n)} + g_{kn} - h_{kn}, \tag{3.10}$$

with the initial forward position given. Contract variables and rebalancing constraints (3.10) are only defined for relevant states satisfying $t(n) < k$.

Rebalancing decisions are made in each state n, after the realizations of the random electricity prices for period $t(n)$ are known. Transaction costs are proportional to the trade volume and is T_F per unit energy bought or sold.

European-type option contracts can also be included. To conserve space, the involvement of options in rebalancing, profit measurement and objective is not shown (see Fleten (2000) for models including options).

Modeling of risk depends on the attitude toward risk in the generation utility. A simple approach that leads to a piecewise linear model is to minimize expected shortfall (Kusy and Ziemba, 1986). Shortfall is defined as profit underperformance relative to some preset profit targets at various periods. Let Π_n^{tot} be the profit to be measured. The exact definition of this profit depends on how the generation utility defines risk. A possible definition is:

$$\begin{aligned} \forall\, n \in \mathcal{N}\colon \Pi_n^{\text{tot}} &= \Pi_n + \pi_n f_{t,a(n)} \\ &+ \sum_{k<t(n)} [(\varphi_{tn} - T_F)h_{kn} - (\varphi_{kn} + T_F)g_{kn}], \end{aligned} \tag{3.11}$$

where $f_{t,a(n)}$ is the forward position in the product that has delivery in period t, during the actual delivery period.

Let C_{mt} be the marginal shortfall cost in segment (piece) m and let s_{nm} be shortfall. The following constraint defines shortfall variables:

$$\Pi_t^{\text{tot}} + \sum_m s_{mn} \geq B_t, \tag{3.12}$$

for all states n for which there is a profit target B_t.

Let W be a weight parameter. In order to avoid incurring excess transaction costs, the objective function maximizes expected profit minus the weight times expected shortfall:

$$\max_{f,g,h} \sum_{n\in\mathcal{N}} P_n(1+r)^{-N_{t(n)}} \left[\Pi_n^{\text{tot}} - W \sum_m C_{mt} s_{mn} \right]. \tag{3.13}$$

This model does not treat physical and financial forward-type contracts differently. The reason is that with the assumptions we have made, a financial contract is a perfect substitute for a physical contract. It generates the exact

same cash flow. Some generation utilities in newly liberalized markets have physical bilateral sales contracts that have a minimum energy volume that is very large and whose tariff structure is complex. The market for such wholesale consumption contracts will quickly become competitive, since small power marketers can sell such contracts and cover the liability in the spot and financial markets. The integrality of these contracts (large minimum volume) will not be a problem either, since one can always add or delete energy volume by buying or selling additional (physical or financial) contracts. The decision support tool needed for such bilateral sales contracts is thus not only a portfolio optimization model, but also a good model for pricing the specialties (e.g., embedded physical load risk) of the individual contracts. See e.g., Thompson (1995) for such an approach applied to take-or-pay contracts.

3.6 Capacity expansion

In a deregulated and well-functioning market, capacity expansion decisions should be analyzed in view of their profit and market value adding potential, and not their ability to serve growing demand at minimum cost. As such, future electricity prices, as opposed to demand, is the central object of analysis. A lot of work remains to be done on this arena, but as a starting point the readers are referred to Deng and Oren (2001), who analyze an investment in a gas-fired power plant using a stochastic dynamic programming model that includes startup costs, operating-dependent efficiency and ramping constraints.

4 Oil

4.1 Optimal field development

Haugland et al. (1988) discuss an optimization model for an oil field based on a two-dimensional reservoir model of the same type that is used in reservoir simulations (but of course much simpler). The goal is to determine platform capacity, the number of wells (and their placement and timing), plus the production profile of each well. This way of using the reservoir simulation equations within an optimization model provides a setting that spans two different fields of research. This is useful both for quality and acceptance.

But since the model is deterministic, all aspects of flexibility are gone, including the postponement of decisions. Jonsbråten (1998a) adds one type of stochasticity to these models by assuming that future oil prices are random. He describes them using scenarios. He then solves the resulting stochastic mixed integer program with scenario aggregation on the continuous variables and a heuristic for finding feasible integer solutions. He observes what is expected, namely that as soon as stochasticity is introduced, timing of

decisions, in order to take into account accumulation of information, becomes important.

It is clearly possible to expand this type of models to include other types of randomness. However, we should be aware that gathering of information about the reservoir (over time) will depend on the actual decisions made. Stochastic programming for such cases is barely developed. An initial discussion can be found in Jonsbråten's doctoral thesis (Jonsbråten, 1998b).

4.2 Scheduling arrivals of tankers at a refinery

This problem originates from Bjørstad et al. (1991), and is interesting as its randomness is different from what we have seen elsewhere. A refinery is about to receive a large ship for loading of gasoline for export. For simplicity, we shall assume that gasoline is characterized by two qualities, namely sulphur content (the lower, the better) and octane number (the higher, the better). In reality, there are many other properties, but this is enough for our example. For the arriving ship, it is known how much gasoline it needs, and there are given a minimal value for octane number and a maximal value for sulphur content. At a refinery, gasoline is not stored as final products, but rather as intermediate components, such as propane, butane etc. These are the results of the refining process, and are stored in tanks (with limited capacity). To fulfill an export order, one mixes components from the different tanks, to achieve a product with the desired properties. It is not always possible to achieve exactly the boundary values of the qualities, and in such a case it is a goal to give away as little extra quality as possible. Clearly, if one gives away very little extra quality in one shipment, one may be left in a situation in terms of stored components, such that for later shipments the quality giveaway is very high. Hence, one needs to have a somewhat long view on the production.

For many refineries, the arrival time of the exporting ships is uncertain. This is caused mainly by bad weather, but other reasons are of course possible as well. Hence, although both production of components, and requirements (quantities and qualities) relating to arriving vessels may be known for some periods into the future, the very fact that their arrival times are unknown will cause some concern. Problems may occur both with respect to production (full tanks because no ship arrived), and the mixing of gasoline for a specific ship, since production continuously change the qualities of the contents in the tanks. Even more severe effects occur if ships arrive in an unexpected order.

Assume we look four periods into the future, and that we know that during those periods three ships will arrive. The model is run when a ship arrives, so ship 1 is known to arrive in the first period. Figure 9 shows the six possible arrival sequences, with given estimated probabilities.

The scenario representing what we expect to happen is scenario 2, where ship A arrives in period 2, and ship B in period 4. This has a probability of 60%.

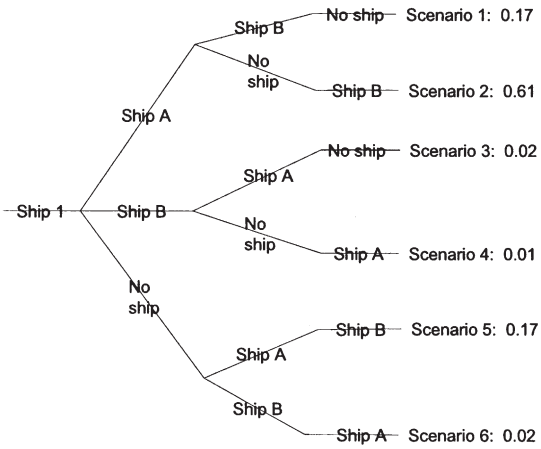

Fig. 9. The possible arriving sequence for the three ships over four periods, with given probabilities.

This *sequence* is almost certain, as it is only 5% chance that B arrives before A. So in a deterministic model, we would clearly use scenario 2.

But let us describe the problem in some detail. When a continous model is made discrete, as it is here, there is always the need to make assumptions about the order in which things happen. The assumptions here are:

- No two ships arrive in the same period.
- Gasoline for export is subtracted from the tanks *before* the production of the period is added.
- The periods are long enough to finish loading a ship.

In this example, it turns out that if we solve the problem corresponding to the most likely scenario, we may end up with all sorts of problems later on. For some scenarios we shall experience serious problems with high giveaways, tanks that fill up (resulting in stopped production) and orders that cannot be fulfilled. The optimal solution to the stochastic program takes the future appropriately into account and avoids these problems. In this case, the optimal solution is a scenario solution. But as always, this can be determined only by solving the stochastic program.

4.3 Refinery planning

In addition to the problem outlined above about the arrival of tankers to a refinery, there will always be interesting problems related to the refining process itself. Examples of short-term decisions are what qualities to produce and what tanks to use, medium-term decisions concern which crude oils to buy when, and of course there are long term investment problems. An example of a model in this area is Escudero et al. (1999).

5 Gas

5.1 Scheduling of gas fields

Haugen (1993) discusses the following question from the North Sea. Gas was at the time mostly sold on long term contracts. The income of the producer depended to a large extent on his ability to meet the contracted volumes. The market was connected to the offshore gas fields by pipelines. Some of the fields and pipelines already existed, but new ones had to be developed to satisfy the demands (i.e., the contracts). A stochastic dynamic programming model was set up to decide which fields should be developed when, and which pipelines should be constructed when. Although, as the author points out, many aspects of such a problem are random, this paper focuses on resource uncertainty. The uncertainty is described by defining a production profile, consistent with how that is normally done in the industry, and then letting the time at peak production be stochastic. The size of the peak production is a design variable (production capacity of the platform), while the time spent there is a function of field properties, and hence, random. Some small examples are given. The main result, apart from the model itself, is the fact that the author is not able to extract simple decision rules. This is not a negative result, but shows that the problem is inherently difficult, and that care must be taken (in the real world) when arguments are made on how to develop such fields and infrastructure. Simple arguments are very likely to be false.

5.2 Use of gas storage

In light of the nature of stochastic programs, storage will always be important. Storage of gas will be a way to solve many different problems of flexibility. To mention but a few:

- A gas producer has an obligation to deliver certain amounts of gas at certain points in the network at certain times. He is aware that at times there are interruptions in his production or transportation systems. By having storage facilities near the delivery points, he can reduce the chance of failing to deliver.
- At certain points in the network spot markets for gas exist. For some producers it is hard to take part in such markets, as it may take them several days from a decision about increased production is made until the gas actually reaches the point of spot delivery. A storage facility near the spot market will make it possible to take part in a potentially profitable spot market.
- A local distribution company may have as its sole goal to supply its customers according to their (random) demand at lowest possible cost. In this case storage can both help buy gas at times when it

is cheap, as well as supply gas in periods of high demand (typically cold periods) where there may be problems of delivery (in addition to high costs). The problems may be caused both by lack of available gas (limited production capacity) and lack of transportation capacity.
- Utilities producing electricity from gas will have very similar problems as above. They can save money as well as secure supply of gas by having a storage facility.

The storage facility will create both strategic and operational decisions. The strategic decisions, which normally are the ones interesting from a stochastic programming point of view, can be such as:

- Building a storage facility—in many ways a classical facility location problem.
- Renting (part of) a storage facility. If there are several possibilities, this is also a kind of facility location problem.
- Investing in equipment determining the speed by which gas can be put into and removed from storage—by some called deliverability.

There are also more indirect strategic decisions, such as changing the production capacity of a gas field (changing the number of wells, for example) to take into account the value of being able to add gas to the storage at an increased rate also in periods of high production with direct delivery.

These strategic decisions can show up in many stages of a model. For example, rental of storage capacity can be updated at times, new contracts can be entered into, old ones continued or dropped. This way, storage rental turns into a portfolio problem, where characteristics are geography, deliverability and size. As always, we should expect that the more flexible is a certain storage, the more it costs to build or rent.

Operational decisions are more obvious. In combination with purchases, production or delivery, whichever is the relevant trade, we must optimally use the storage facilities to maximize profit or minimize costs, whatever is the objective.

Useful references here are Butler and Dyer (1999), Bopp et al. (1996) and Takriti et al. (2001).

5.3 *Portfolio management of gas contracts*

Whether we are selling or buying natural gas, the chance is that we need to buy or sell the gas on contracts of different types. These may vary in price and duration. The price difference may stem from differences in forward prices, such that gas on a one-year contract may cost more or less than gas on a two-year contact. But the differences may also stem from how the gas price

depends on other entities, such as oil price, or by special rules on renegotiations of contract details.

In such a picture, we are faced with a portfolio management problem. The goal may be to buy or sell gas so as to obtain an optimal tradeoff between expected profit and some measure of risk. Haurie et al. (1992) discuss this problem for a Canadian producer. The different gas contracts have different time spans and different rules for how prices are set. They have many different models. The first is in the spirit of the Markowitz' mean-variance model, the last is a stochastic program with recourse. Risk is measured in terms of the variance of profits.

6 Conclusion

The purpose of this chapter has been to give an introduction to the use of stochastic programming in energy. Based on the available literature, the focus has naturally been on electricity production, but we have tried to provide some pointers also for natural gas (particularly the treatment of contracts) and oil. The purpose has not been to have a full overview over the literature, but to provide the reader with pointers to interesting problems and starting points for reading.

Stochastic programming used in regulated markets, that is, in monopolies, is a well-established activity. The first articles go far back, and the literature is enormous. Articles typically mix discussions of models and methods, and very often the chosen methodology is stochastic dynamic programming (SDP). We have chosen to base our presentation on models rather than methods, so as to avoid a split of papers into two arbitrary piles; those that use stochastic programming (as understood in this handbook) and those that use SDP. For regulated markets, as that is such a well established field, and since methods and models are almost always mixed, we have chosen to discuss also methodology in that section. For deregulated markets, on the other hand, we have chosen to focus very little on methodology, simply assuming that the reader will use the rest of this handbook to look for appropriate methodology. Instead, we have tried to focus on what the new markets may bring us, and tried to point to relevant theory outside stochastic programming, in particular market theory and options theory. The deregulated markets have not found their final forms, so it is impossible to provide the reader with clear-cut descriptions of where we will end up. Hence, our goal has been to assist and present ideas.

Many problems in resource management concern situations where our decisions will change the (conditional) probability distributions. Drilling exploration wells in an oil or gas field is a good example. As stochastic programming, as it stands today, cannot treat this case in any good way, we have chosen to let those problems rest, and mostly focused on problems where the uncertainty is external to the model at hand.

Acknowledgements

Much of this work was done while Stein W. Wallace headed a research group at the Centre for Advanced Study, at the Academy of Science and Letters in Oslo, during the academic year 2000/01. We would like to thank our cooperators during the last few years on energy related issues, in particular (in alphabetical order) T. Bjørkvoll, J. Lemming, T. T. Lie, M. P. Nowak, A. Tomasgard, and W. T. Ziemba. We would also like to pay tribute to J. Birge, L. F. Escudero, A. Manne, M. V. F. Pereira, A. B. Philpott, W. Römisch, R. Schultz, Y. Smeers, and J. R. Stedinger, who have helped to collect material for this chapter. Partial funding from the European Commission through EUs 5th framework programme project IST–1999–12088 is gratefully acknowledged.

References

Anderson, E.J., A.B. Philpott (2002). On supply function bidding in electricity markets, in: C. Greengard, A. Ruszczyński (eds.), *Decision Making under Uncertainty: Energy and Power. Vol. 128 of IMA volumes on Mathematics and its Applications*, Springer-Verlag, pp. 115–134.

Anderson, E.J., A.B. Philpott (2002). Optimal offer construction in electricity markets. *Mathematics of Operations Research* 27(1), 82–100.

Archibald, T.W., C.S. Buchanan, K.I.M. McKinnon, L.C. Thomas (1999). Nested Benders decomposition and dynamic programming for reservoir optimisation. *Journal of the Operational Research Society* 50(5), 468–479.

Bienstock, D., J.F. Shapiro (1988). Optimizing resource acquisition decisions by stochastic programming. *Management Science* 34(2), 215–229.

Birge, J.R. (1985). Decomposition and partitioning methods for multistage stochastic linear programs. *Operations Research* 33, 989–1007.

Birge, J.R., C.H. Rosa (1996). Incorporating investment uncertainty into greenhouse policy models. *Energy Journal* 17(1), 79–90.

Bjørstad, H., D. Haugland, R. Helming, (1991). A stochastic model for gasoline blending, in: M. Breton, G. Zaccour (eds.), *Advances in Operations Research in the Oil and Gas Industry*, pp. 137–142.

Bloom, J.A. (1983). Solving an electricity generation capacity expansion planning problem by generalized Benders' decomposition. *Operations Research* 31, 84–100.

Bopp, A.E., V.R. Kannan, S.W. Palocsay, S.P. Stevens (1996). An optimization model for planning natural gas purchases, transportation, storage and deliverability. *Omega, International Journal of Management Science* 24(5), 511–522.

Botnen, O.J., A., Johannesen, A. Haugstad, S. Kroken, O. Frøystein, June 1992. Modelling of hydropower scheduling in a national/international context, in: E. Broch, D.K. Lysne (eds.), *Proceedings of the 2nd International Conference on Hydropower Development (Hydropower '92)*, Lillehammer, Norway, Balkema, Rotterdam, pp. 575–584.

Bunn, D.W., S.N. Paschentis (1986). Development of a stochastic model for the economic dispatch of electric power. *European Journal of Operational Research* 27, 179–191.

Butler, J.C., J.S. Dyer (1999). Optimizing natural gas flows with linear programming and scenarios. *Decision Sciences* 30(2), 563–580.

Caroe, C.C., R. Schultz (1998). A two-stage stochastic program for unit commitment under uncertainty in a hydro-thermal power system. Preprint SC 98-11, Konrad-Zuse-Zentrum für Informationstechnik Berlin.

Carpentier, P., G. Cohen, J.C. Culioli, A. Renaud (1996). Stochastic optimization of unit commitment: A new decomposition framework. *IEEE Transactions on Power Systems* 11(2), 1067–1073.

Dantzig, G., G. Infanger (1993). Approaches to stochastic programming with application to electric power systems, in: K. Frauendorfer, H. Glavitsch, R. Bacher (eds.), *Optimization in Planning and Operation of Electric Power Systems*, Physica-Verlag, Heidelberg, pp. 125–138.

Dantzig, G.B. (1989). Decomposition techniques for large-scale electric power systems planning under uncertainty, in: R. Sharda, B.L. Golden, E. Wasil, O. Balci, W. Steward (eds.), *Impacts of Recent Computer Advances on Operations Research*, North-Holland, pp. 3–20.

Deng, S.-J., S.S. Oren (2001). Incorporating operational characteristics and startup costs in option-based valuaton of power generation capacity, forthcoming, Probability in the Engineering and Informational Sciences (PEIS), April 2003.

Dentcheva, D., W. Römisch (1998). Optimal power generation under uncertainty via stochastic programming, in: K. Marti, P. Kall (eds.), *Stochastic Programming Methods and Technical Applications, Lecture Notes in Economics and Mathematical Systems*, Vol. 458, Springer-Verlag, Berlin, pp. 22–56.

Escudero, L.F., I. Paradinas, F.J. Prieto (1995). Generation expansion planning under uncertainty in demand, economic environment, generation availability and book life, in: *Proceedings of the IEEE Stockholm Power Tech*. Stockholm, Sweden, pp. 226–233.

Escudero, L.F., F.J. Quintana, J. Salmeron (1999). CORO, a modeling and an algorithmic framework for oil supply, transformation and distribution optimization under uncertainty. *European Journal of Operational Research* 114(3), 638–656.

Feltenmark, S., P.O. Lindberg (1997). Network methods for head-dependent hydro power planning, in: P.M. Pardalos, D.W. Hearn, W.W. Hager (eds.), *Network Optimization. Lecture Notes in Economics and Mathematical Systems*, Springer-Verlag, Berlin, pp. 249–264.

Fleten, S.-E. (2000). Portfolio management emphasizing electricity market applications. a stochastic programming approach. Ph.D. thesis, Norwegian University of Science and Technology, Trondheim, doktor Ingeniøravhandling 2000:16.

Fleten, S.-E., S.W. Wallace, W.T. Ziemba (2002). Hedging electricity portfolios using stochastic programming, in: C. Greengard, A. Ruszczynski (eds.), *Decision Making under Uncertainty: Energy and Power. Vol. 128 of IMA Volumes on Mathematics and Its Applications*, Springer-Verlag, pp. 71–94.

Fosso, O.B., A. Gjelsvik, A. Haugstad, B. Mo, I. Wangensteen (1999). Generation scheduling in a deregulated system. The Norwegian case. *IEEE Transactions on Power Systems* 14(1), 75–80.

Fragniere, E., A. Haurie (1996). A stochastic programming model for energy/environment choices under uncertainty. International Journal of Environment and Pollution 6 (4-6), 587–603, in: A.V. Georghe, (ed.), *Integrated Regional Health and Environmental Risk Assessment and Safety Management*.

Gardner, D.T. (1996). Flexibility in electric power planning: Coping with demand uncertainty. *Energy* 21(2), 1207–1218.

Gardner, D.T., J.S. Rogers (1999). Planning electric power systems under demand uncertainty with different technology lead times. *Management Science* 45, 1289–1306.

Gjelsvik, A., T.A. Røtting, J. Røynstrand, June 1992. Long-term scheduling of hydro-thermal power system, in: E. Broch, D.K. Lysne, (eds.), Proceedings of the 2nd International Conference on Hydropower Development (Hydropower'92), Lillehammer, Norway. Balkema, Rotterdam, pp. 539–546.

Gjelsvik, A., S.W. Wallace (1996). Methods for stochastic medium-term scheduling in hydro-dominated power systems. EFI TR A4438, Norwegian Electric Power Research Institute, Trondheim.

Gollmer, R., M.P. Nowak, W. Römisch, R. Schultz (2000). Unit commitment in power generation—a basic model and some extensions. *Annals of Operations Research* 96, 167–189.

Gorenstin, B.G., N.M. Campodónico, J.P. Costa, M.V.F. Pereira (1992). Stochastic optimization of a hydrothermal system including network constraints. *IEEE Transactions on Power Systems* 7(2), 791–797.

Gorenstin, B.G., N.M. Campodónico, J.P. Costa, M.V.F. Pereira, N. Deeb (1993). Power-system expansion planning under uncertainty. *IEEE Transactions on Power Systems* 8(1), 129–136.

Grinold, R.C. (1980). Time horizons in energy planning models, in: W.T. Ziemba, S.L. Schwarz (eds.), *Energy Policy Modeling: United States and Canadian Experiences*, Vol. II, Martinus Nijhoff, Boston, pp. 216–237.

Gröwe, N., W. Römisch, (1992). A stochastic programming model for optimal power dispatch: Stability and numerical treatment, in: K. Marti (ed.), *Stochastic Optimization. Vol. 379 of Lecture Notes in Economics and Mathematical Systems*, Springer-Verlag, Berlin, pp. 111–139.

Gröwe, N., W. Römisch, R. Schultz (1995). A simple recourse model for power dispatch under uncertain demand. *Annals of Operations Research* 59, 135–164.

Gröwe-Kuska, N., K.C. Kiwiel, M.P. Nowak, W. Römisch, I. Wegner (2002). Power management in a hydro-thermal system under uncertainty by Lagrangian relaxation, in: C. Greengard, A. Ruszczynski (eds.), *Decision Making under Uncertainty: Energy and Power. Vol. 128 of IMA volumes on Mathematics and its Applications*, Springer-Verlag, pp. 39–70.

Haugen, K.K. (1993). A stochastic dynamic programming model for scheduling of offshore petroleum fields with resource uncertainty. *European Journal of Operational Research* 88, 88–100.

Haugland, D., A. Hallefjord, H. Asheim (1988). Models for petroleum field exploitatin. *European Journal of Operational Research* 37(1), 58–72.

Haurie, A., Y. Smeers, G. Zaccour (1992). Toward a contract portfolio management model for a gas producing firm. *INFOR* 30(3), 257–273.

Hindsberger, M. (2003) Interconnected hydro-thermal systems. Ph.D. thesis, Technical University of Denmark.

Hobbs, B.F. (1995). Optimization methods for electric utility resource planning. *European Journal of Operational Research* 83(1), 1–20.

Hobbs, B.F., Y.D. Ji (1999). Stochastic programming-based bounding of expected production costs for multiarea electric power systems. *Operations Research* 47(6), 836–848.

Høyland, K., S.W. Wallace (2001). Generating scenario trees for multistage decision problems. *Management Science* 47(2), 295–307.

Hull, J.C. (2000) Options, Futures and Other Derivatives, 4th Edition, Prentice Hall International, Upper Saddle River, N.J.

Jacobs, J., G. Freeman, J. Grygier, D. Morton, G. Schultz, K. Staschus, J. Stedinger (1995). SOCRATES–A system for scheduling hydroelectric generation under uncertainty. *Annals of Operations Research* 59, 99–133.

Jonsbråten, T.W. (1998). Oil field optimization under price uncertainty. *Journal of the Operational Research Society* 49(8), 811–818.

Jonsbråten, T.W. (1998b). Optimization models for petroleum field exploitation. Ph.D. thesis, Norwegian School of Economics and Business Administration, Bergen, Norway.

Kanudia, A., R. Loulou (1998). Robust responses to climate change via stochastic MARKAL: The case of Quebec. *European Journal of Operational Research* 106(1), 15–30.

Kelman, R., L.A.N. Barroso, M.V.F. Pereira (2001). Market power assessment and mitigation in hydrothermal systems. *IEEE Transactions on Power Systems* 16(3), 354–359.

King, A.J. (2002) Duality and martingales: A stochastic programming perspective on contingent claims. *Mathematical Programming* 91(3), 543–562.

Kreps, D. (1979). Multi-period securities and efficient allocation of risk: *The Economics of Information and Uncertainty*, McCall, J.J. University of Chicago Press, Chicago, pp. 203–232.

Kusy, M.I., W.T. Ziemba (1986). A bank asset and liability management model. *Operations Research* 34(3), 356–376.

Lindqvist, J. (1962). Operation of a hydrothermal electric system: A multistage decision process. *AIEE Trans. Power Apparatus and Systems* 81, 1–7.

Louveaux, F.V. (1980). A solution method for multistage stochastic programs with recourse with application to an energy investment problem. *Operations Research* 28(4), 889–902.

Lucia, J.J., E.S. Schwartz (2000) Electricity prices and power derivatives. Evidence from the Nordic Power Exchange. *Review of Derivatives Research* 5(1), 5–50.

Manne, A.S., T.R. Olsen (1996). Greenhouse gas abatement—toward Pareto-optimal decisions under uncertainty. *Annals of Operations Research* 68, 267–279.

Manne, A.S., R.G. Richels (1978). A decision analysis of the U.S. breeder reactor program. *Energy* 3(6), 747–767.

Manne, A.S., R.G. Richels (1991). *Buying greenhouse insurance Energy Policy*, Vol. 19(6).

Manne, A.S., R.G. Richels (1995). The greenhouse debate: Economic efficiency, burden sharing and hedging strategies. *The Energy Journal* 16(4), 1–37.

Massé, P. (1946). Les Réserves et la Régulation de l'Avenir dans la vie Économique. Hermann, Paris, vol I and II.

Mo, B., A. Gjelsvik, A. Grundt (2001). Integrated risk management of hydro power scheduling and contract management. *IEEE Transactions on Power Systems* 16(2), 216–221.

Morton, D.P. (1996). An enhanced decomposition algorithm for multistage stochastic hydroelectric scheduling. *Annals of Operations Research* 64, 211–235.

Murphy, F.H., S. Sen, A.L. Soyster (1982). Electric utility capacity expansion planning with uncertain load forecasts. *IIE Transactions* 14(1), 52–59.

Naik, V., (1995). Finite state securities market models and arbitrage, in: R. Jarrow, V. Maksimovic W.T. Ziemba (eds.), *Finance. Vol. 9 of Handbooks in Operations Research and Management Science*, Elsevier, Amsterdam, pp. 31–64.

Neame, P.J., A.B. Philpott, G. Pritchard (1999). Offer stack optimisation for price takers in electricity markets, in: *Proceedings of the 34th Annual Conference of ORNZ*, University of Waikato, Hamilton, New Zealand, pp. 3–12.

Nowak, M.P., R. Nürnberg, W. Römisch, R. Schultz, M. Westphalen, (2000) Stochastic programming for power production and trading under uncertainty. Preprint SM-DU-471, Fachbereich Mathematik, Universität Duisburg, submitted.

Nowak, M.P., W. Römisch (2000). Stochastic Lagrangian relaxation applied to power scheduling in a hydro-thermal system under uncertainty. *Annals of Operations Research* 100, 251–272.

Pereira, M.V.F., N. Campodónico, R. Kelman (2000). Stochastic optimization of complex hydrothermal systems in a competitive framework, not published.

Pereira, M.V.F., N.M. Campodónico, B.G. Gorenstin, J.P. Costa (1995) Application of stochastic optimization to power system planning and operation, in: Proceedings of the IEEE Stockholm Power Tech. Stockholm, Sweden, pp. 234–239.

Pereira, M.V.F., B.G. Gorenstin, M.M. Fo, J.B. Silva (1992). Chronological probabilistic production costing and wheeling calculations with transmission network modeling. *IEEE Transactions on Power Systems* 7(2), 885–891.

Pereira, M.V.F., L.M.V.G. Pinto (1985). Stochastic optimization of a multireservoir hydroelectric system—a decomposition approach. *Water Resources Research* 21(6), 779–792.

Pereira, M.V.F., L.M.V.G. Pinto (1991). Multi-stage stochastic optimization applied to energy planning. *Mathematical Programming* 52(2), 359–375.

Pereira, M.V.F., L.M.V.G. Pinto, S. Granville, A. Monticelli (1987) A decomposition approach to security constrained optimal power flow with post-contingency corrective rescheduling, in: Power Systems Computation Conference 1987.

Pflug, G.C. (2001). Scenario tree generation for multiperiod financial optimization by optimal discretization. *Mathematical Programming* 89(2), 251–271.

Philbrick, C.R., P.K. Kitanidis (1999). Limitations of deterministic optimization applied to reservoir operations. *Journal of Water Resources Planning and Management-ASCE* 125(3), 135–142.

Philpott, A.B., M. Craddock, H. Waterer (2000). Hydro-electric unit commitment subject to uncertain demand. *European Journal of Operational Research* 125(2), 410–424.

Qiu, J., A.A. Girgis (1993). Optimization of power-system reliability level by stochastic-programming. *Electric Power Systems Research* 26(2), 87–95.

Ross, S. (1977). Risk, return and arbitrage, in: I. Friend, J.L. Bicksler (eds.), *Risk and Return in Finance*, Ballinger, Cambridge, Massachusetts.

Schwartz, E.S. (1997). The stochastic behavior of commodity prices: Implications for valuation and hedging. *Journal of Finance* 52(3), 922–973.

Scott, T.J., E.G. Read (1996). Modelling hydro reservoir operation in a deregulated electricity sector. *International Transactions in Operations Research* 3(3–4), 209–221.

Sheble, G.B., G.N. Fahd (1994). Unit commitment literature synopsis. *IEEE Transactions on Power Systems* 9(1), 128–135.

Sherali, H.D., A.L. Soyster, F.H. Murphy, S. Sen (1982). Linear programming based analysis of marginal cost pricing in electric utility capacity expansion. *European Journal of Operational Research* 11, 349–360.

Sherali, H.D., A.L. Soyster, F.H. Murphy, S. Sen (1984). Intertemporal allocation of capital costs in electric utility capacity expansion planning under uncertainty. *Management Science* 30, 1–19.

Sherkat, V.R., R. Campo, K. Moslehi, E.O. Lo (1985). Stochastic long-term hydrothermal optimization for a multireservoir system. *IEEE Trans. Power Apparatus and Systems* 104(8), 2040–2049.

Smeers, Y. (1990). Traitement de l'incertain dans le calcul des plans d'équipement électrique, in: B. Cornet, H. Tulkens, (eds.), *Modélisation et Décision Economique*, De Boeck Université, Bruxelles.

Stedinger, J.R. (1998). Stochastic multi-reservoir hydroelectric scheduling, in: Proceedings 28. Internationales Wasserbau-Symposium, Wasserwirtschafliche Systeme-Konzepte, Konflikte, Kompromisse. Institut für Wasserbau und Wasserwirtschaft, Rheinisch-Westfälischen Technischen Hochschule Aachen, Aachen, Germany, pp. 89–117.

Takriti, S., J.R. Birge, E. Long (1996). A stochastic model for the unit commitment problem. *IEEE Transactions on Power Systems* 11(3), 1497–1506.

Takriti, S., B. Krasenbrink, L.S.Y. Wu (2000). Incorporating fuel constraints and electricity spot prices into the stochastic unit commitment problem. *Operations Research* 48(2), 268–280.

Takriti, S., C. Supatgiat, L.S.Y. Wu (2001). Coordinating fuel inventory and electric power generation under uncertainty. *IEEE Transactions on Power Systems* 16(4), 603–608.

Tejada-Guibert, S.A. Johnson, J.R. Stedinger (1995). The value of hydrologic information in stochastic dynamic programming models of a multireservoir system, *Water Resources Research* 31, 2571–2579.

Terry, L.A., M.V.F. Pereira, T.A.A. Neto, L.E.C.A. Pinto, P.R.H. Sales (1986). Coordinating the energy generation of the Brazilian national hydrothermal electrical generating system. *INTERFACES* 16(1), 16–38.

Thompson, A.C. (1995). Valuation of path-dependent contingent claims with multiple exercise decisions over time—the case of take-or-pay. *Journal of Financial and Quantitative Analysis* 30(2), 271–293.

Tseng, C.-L., G., Barz (2002). Short-term generation asset valuation: A real options approach. *Operations Research* 50(2), 297–310.

Turgeon, A. (1980). Optimal operation of multireservoir power systems with stochastic inflows. *Water Resources Research* 16, 275–283.

Wallace, S.W. (1986). Solving stochastic programs with network recourse. *Networks* 16, 295–317.

Wood, A.J., B.F. Wollenberg (1996). *Power Generation, Operation, and Control*, 2nd ed., John Wiley & Sons, New York.

Yakowitz, S. (1982). Dynamic programming applications in water resources. *Water Resources Research* 18(4), 673–696.

Yeh, W.W.-G. (1985). Reservoir management and operation models: a state of the art review. *Water Resources Research* 21(12), 1797–1818.

Subject Index