Chapter 1

Introduction

Most real-world problems in operations research involve uncertain data. Thus, finding optimal decisions turns into selecting a “best” random variable. Then the question comes up which criteria to use for the selection. Quickly, the matter of risk aversion becomes an issue. Further, for realistic modeling integer variables are often helpful and sometimes inevitable. This thesis suggests a way of how to make such decisions in the framework of two-stage stochastic mixed-integer programming.

Section 1.1 introduces the concept of mean-risk models, and in Section 1.2 we define the risk measures we want to analyze in this thesis. We close this chapter with Section 1.3 giving an introduction to two-stage stochastic mixed-integer programming and extending the traditional expectation-based stochastic program towards risk aversion by formulating mean-risk models with the risk measures defined in the previous section. In Chapter 2 we analyze the added model components with respect to structure and stability and in Chapter 3 computational issues are covered.

1.1 Mean-Risk Models

Throughout the thesis, we impose a cost minimization framework. Consider a probability space $(\Omega, \mathcal{A}, I\!\!P)$ and the set $\mathcal{Z}$ of all real random cost variables $Z : \Omega \longrightarrow \mathbb{R}$. Suppose, we want to find a decision $x \in X$ such that the random future costs, represented by the random variable $Z(x, \omega) \in \mathcal{Z}$, would best suit our purpose. This leads to finding a "best" random variable out of the family $\{Z(x, \omega)\}_{x \in X} \subseteq \mathcal{Z}$. We want to decide on the decision variable x and so on the corresponding random variable $Z(x, \omega)$ by comparing certain scalar characteristics of the random variables, namely by so-called mean-risk models

$$\min_{x \in X} I\!\!E(Z(x, \omega)) + \rho \mathcal{R}(Z(x, \omega)), \qquad \rho > 0, \tag{1.1}$$

where $I\!\!E : \mathcal{Z} \longrightarrow \mathbb{R}$ denotes the expected value, $\mathcal{R} : \mathcal{Z} \longrightarrow \mathbb{R}$ a risk measure, and $\rho > 0$ a suitable weight factor.

We consider the mean-risk models with the risk measures Excess Probability, Expected Excess, Value-at-Risk and Conditional Value-at-Risk applied to random variables occurring in two-stage stochastic mixed-integer programming. These random variables are essentially defined by value functions of mixed-integer linear programs being discontinuous and nonconvex, such that, in particular, convexity of the objectives in the mean-risk model is not given. Therefore, for the sake of applicability to real-world problems, it is essential to choose a risk measure such that, despite the poor properties of the random variables, the resulting stochastic integer programs are nevertheless structurally sound and amenable to algorithmic treatment. In the following we will show that the risk measures under consideration mostly satisfy these requirements.

The mean-risk model (1.1) aims at minimizing the weighted sum of two competing objectives. Viewed from a more general perspective, it is a scalarization of the multiobjective optimization problem

$$\min_{x \in X} \big(I\!\!E(Z(x, \omega)), \mathcal{R}(Z(x, \omega)) \big). \tag{1.2}$$

For an introduction to multiobjective optimization we refer to [9, 30, 56]. An accepted notion of optimality in multiobjective optimization is efficiency. A point $\bar{x} \in X$ is called efficient for (1.2) if there is no other point $x \in X$ such that $\mathbb{E}(Z(x,\omega)) \leqslant \mathbb{E}(Z(\bar{x},\omega))$ and $\mathcal{R}(Z(x,\omega)) \leqslant \mathcal{R}(Z(\bar{x},\omega))$, with at least one strict inequality. The set of all efficient points is named efficient frontier. Every optimal solution to the mean-risk model (1.1) with a weight factor $\rho > 0$ is an efficient point, a so-called supported efficient point. Due to the lacking convexity of our objective functions, not all efficient points are supported, and thus cannot be computed by solving scalarizations (1.1). However, solving the mean-risk model (1.1) for various values of $\rho > 0$ has the capability to trace the supported part of the efficient frontier. In Chapter 3 a discrete tracing method is described and carried out for a real-life optimization problem.

In the next section we introduce our four risk measures for the application in the framework of two-stage stochastic mixed-integer programming.

1.2 Risk Measures

The definition of risk is a highly subjective matter – each decision maker might have his own. Having this in mind, we just want to suggest a collection of four risk measures that are easily comprehensible. They are chosen such that they fulfill certain well accepted properties every measure of risk should have. And, last but not least, the risk measures are selected with our particular application in two-stage stochastic mixed-integer programming in mind – that is to say, the application of the measure to our setting ought to lead to optimization problems that are as "nice" as the purely expectation based problem, both in terms of structure and stability, and with respect to computational properties. Summing up, our risk measures should be "stochastically sound", "structurally sound" and "computationally sound".

We proceed by defining the measures and addressing the above issues. The first is treated completely in this section, the other two are just touched on in the next section, whereas Chapter 2 and Chapter 3 are devoted to these topics.

We want to consider two classes of risk measures – both ask for the preselection of a parameter –

$$\text{a cost threshold } \eta \in \mathbb{R},$$

or

$$\text{a probability } \alpha \in (0,1),$$

where $(0,1)$ is the open interval $\{\alpha \in \mathbb{R} \ : \ 0 < \alpha < 1\}$. The cost threshold η can be interpreted as a certain threshold of pain or ruin level, and α as the probability level of the costs the decision maker is willing to tolerate.

For fixed $x \in X$ and assuming that $\mathbb{E}(|Z(x,\omega)|) < +\infty$, we define the following risk measures:

The Excess Probability: “probability that costs exceed η”,

$$EP_\eta(Z(x,\omega)) := \mathbb{P}(\{\omega \in \Omega \ : \ Z(x,\omega) > \eta\}),$$

the Expected Excess: “expectation of costs exceeding η”,

$$EE_\eta(Z(x,\omega)) := \mathbb{E}(\max\{Z(x,\omega) - \eta, 0\}),$$

the Value-at-Risk: “minimal costs of $(1-\alpha)\cdot 100\%$ worst cases”,

$$VaR_\alpha(Z(x,\omega)) := \min\{\eta \ : \ EP_\eta(Z(x,\omega)) \leqslant 1-\alpha, \ \eta \in \mathbb{R}\},$$

and the Conditional Value-at-Risk: “expectation of costs in $(1-\alpha)\cdot 100\%$ worst cases”,

$$CVaR_\alpha(Z(x,\omega)) := \min\{\eta + \frac{1}{1-\alpha} EE_\eta(Z(x,\omega)) \ : \ \eta \in \mathbb{R}\}.$$

It holds that $EP_\eta(Z(x,\omega)) = 1 - F_x(\eta)$ where $F_x(\eta) := \mathbb{P}(\{\omega \in \Omega \ : \ Z(x,\omega) \leqslant \eta\})$ is the distribution function of the random variable $Z(x,\omega)$. Thus, the Excess Probability is well defined. We have chosen this definition since we want to minimize the probability of the worst cases. Since $\mathbb{E}(|Z(x,\omega)|) < +\infty$, the Expected Excess is also well defined.

The Value-at-Risk $VaR_\alpha := VaR_\alpha(Z(x,\omega))$ denotes an α-quantile of the random variable $Z(x,\omega)$ and by the above argument for the Excess Probability it holds that $VaR_\alpha = \min\{\eta \ : \ F_x(\eta) \geqslant \alpha, \ \eta \in \mathbb{R}\}$. The minimum is always attained, since the distribution function is nondecreasing and right-continuous in η. Consider $VaR_\alpha^+ := \inf\{\eta \ : \ F_x(\eta) > \alpha, \ \eta \in \mathbb{R}\}$. It is immediate that always $VaR_\alpha \leqslant VaR_\alpha^+$. These values are equal unless $F_x(\cdot)$ is constant at α over a certain interval. When $F_x(\cdot)$ is continuous and strictly increasing, $VaR_\alpha = VaR_\alpha^+$ is simply the unique η satisfying $F_x(\eta) = \alpha$. Otherwise, it is possible that this equation has no solution or a whole range of solutions. In the former situation $F_x(\cdot)$ has a probability atom at VaR_α, while in the latter, the graph of $F_x(\cdot)$ has a constant segment at $F_x(\cdot) = \alpha$ being either the half-open interval $[VaR_\alpha, VaR_\alpha^+)$ or the closed interval $[VaR_\alpha, VaR_\alpha^+]$, depending on whether or not $F_x(\cdot)$ has a jump at VaR_α^+.

The identity of the verbal and the mathematical definitions for the first three risk measures are self-evident. Although many authors use the same definition as we for the Conditional Value-at-Risk, cf. [2, 31, 73], we want to derive it from the verbal definition following the paper of Rockafellar and Uryasev ([84]):

If there is no probability atom at VaR_α and so $F_x(VaR_\alpha) = \alpha$, the Conditional Value-at-Risk is equal to the conditional expectation

$$\mathbb{E}(Z(x,\omega) \mid Z(x,\omega) \geqslant VaR_\alpha), \tag{1.3}$$

since VaR_α equals the "minimal costs of $(1-\alpha)\cdot 100\%$ worst cases".

Note, that (1.3) is the usual definition of the Conditional Value-at-Risk for continuous distribution functions (having no probability atoms at all) as it then coincides with the verbal definition, cf. [73, 83, 84].

But, if there is no η such that $F_x(\eta) = \alpha$ and so there is a probability atom at VaR_α, which in particular may occur for discretely distributed random variables, (1.3) does not coincide with the verbal definition of Conditional Value-at-Risk. A correct definition for the general case is

$$CVaR_\alpha(Z(x,\omega)) := \text{ mean of the } \alpha\text{-tail distribution of } Z(x,\omega),$$

where the distribution in question is the one with the distribution function defined by

$$F_x^\alpha(\eta) := \begin{cases} 0 & \text{for } \eta < VaR_\alpha, \\ [F_x(\eta) - \alpha]/[1-\alpha] & \text{for } \eta \geqslant VaR_\alpha. \end{cases} \tag{1.4}$$

For a rigorous proof of this, including graphical examples we refer the reader to [84]. The problem is, when using (1.3) for the general case, one is not taking the expectation of the upper $(1-\alpha)$-part of the full distribution, since the probability atom at VaR_α must be split to do so, but this can not be done by taking any conditional expectation. Thus one has to make the trick as in (1.4): taking the correct part of the original distribution function and rescale it onto $[0,1]$.

Another correct formalization for the general case has been worked out in [1, 2], where the Conditional Value-at-Risk is expressed as a difference of an expectation and a correcting exceeding part if there is a probability atom at VaR_α. The authors also discuss the confusion that inheres the current publications on this subject due to the latter described problems. In particular the authors mention that the name Conditional Value-at-Risk stems from the time where the continuity of the distribution function was assumed and thus the conditional expectation (1.3) was the correct definition. However, they show that in the general case there is no way to express the Conditional Value-at-Risk as a conditional expectation and thus decline the term Conditional Value-at-Risk. They suggest the name Expected Shortfall for the gain maximization framework.

In [70] the Conditional Value-at-Risk was defined for general distribution functions by means of the second quantile function being the convex conjugate function of the distribution function of order two. The authors call the risk measure Tail Value-at-Risk.

We state the last step to our definition of the Conditional Value-at-Risk without proof, cf. [2, 70, 73, 84].

Proposition 1.2.1 *For $I\!E(|Z(x,\omega)|) < +\infty$, the Conditional Value-at-Risk can be expressed by the following minimization formula:*

$$CVaR_\alpha(Z(x,\omega)) = \min\{\eta + \frac{1}{1-\alpha} EE_\eta(Z(x,\omega)) \ : \ \eta \in \mathbb{R}\}.$$

Further, $\eta + \frac{1}{1-\alpha} EE_\eta(Z(x,\omega))$ is convex in η and finite (hence continuous) and the optimal set is the nonempty closed interval $[VaR_\alpha, VaR_\alpha^+]$, reducing to VaR_α when the graph of F_x has no constant segment at $F_x(\eta) = \alpha$. In particular, VaR_α always is a minimizer, and thus $CVaR_\alpha(Z(x,\omega)) \in \mathbb{R}$.

As announced, we conclude this section with checking whether our risk measures are "stochastically sound":

The relations of stochastic dominance, cf. [58, 66], one of the fundamental concepts in decision theory, introduce partial orders in the space of real random variables. This provides a basis for selecting "best" members from families of random variables taking risk aversion preferences into account. Our risk measures are related to the first-degree stochastic dominance relation

$$Z(x_1,\omega) \preceq_1 Z(x_2,\omega) \quad :\Longleftrightarrow \quad \forall \eta \in \mathbb{R}: \ \ EP_\eta(Z(x_1,\omega)) \leqslant EP_\eta(Z(x_2,\omega)),$$

and the second-degree stochastic dominance relation

$$Z(x_1,\omega) \preceq_2 Z(x_2,\omega) \quad :\Longleftrightarrow \quad \forall \eta \in \mathbb{R}: \ \ EE_\eta(Z(x_1,\omega)) \leqslant EE_\eta(Z(x_2,\omega)).$$

Applying these partial orders directly to optimization would lead to multiobjective optimization problems with a continuum of criteria. Ogryczak and Ruszczyński have studied mean-risk models, which are single-criterion optimization problems, and their consistency with multiobjective criteria induced by stochastic dominance, see [69, 70]. A mean-risk model is called consistent with a stochastic dominance relation of a certain degree, if optimal solutions to the