

Bazovkin, Pavel; Mosler, Karl

Working Paper

Stochastic linear programming with a distortion risk constraint

Discussion Papers in Statistics and Econometrics, No. 6/11

Provided in Cooperation with:

University of Cologne, Institute of Econometrics and Statistics

Suggested Citation: Bazovkin, Pavel; Mosler, Karl (2011) : Stochastic linear programming with a distortion risk constraint, Discussion Papers in Statistics and Econometrics, No. 6/11, University of Cologne, Seminar of Economic and Social Statistics, Cologne

This Version is available at:

<https://hdl.handle.net/10419/67613>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

DISCUSSION PAPERS IN STATISTICS AND ECONOMETRICS

SEMINAR OF ECONOMIC AND SOCIAL STATISTICS
UNIVERSITY OF COLOGNE

No. 6/2011

Stochastic linear programming with a
distortion risk constraint

by

Pavel Bazovkin

Karl Mosler



DISKUSSIONSBEITRÄGE ZUR
STATISTIK UND ÖKONOMETRIE
SEMINAR FÜR WIRTSCHAFTS- UND SOZIALSTATISTIK
UNIVERSITÄT ZU KÖLN

Albertus-Magnus-Platz, D-50923 Köln, Deutschland

DISCUSSION PAPERS IN STATISTICS AND ECONOMETRICS

SEMINAR OF ECONOMIC AND SOCIAL STATISTICS
UNIVERSITY OF COLOGNE

No. 6/2011

Stochastic linear programming with a distortion risk constraint

by

Pavel Bazovkin¹

Karl Mosler²

Abstract

Linear optimization problems are investigated whose parameters are uncertain. We apply coherent distortion risk measures to capture the violation of restrictions. Such a model turns out to be appropriate for many applications and, principally, for the mean-risk portfolio selection problem. Each risk constraint induces an uncertainty set of coefficients, which comes out to be a weighted-mean trimmed region. We consider a problem with a single constraint. Given an external sample of the coefficients, the uncertainty set is a convex polytope that can be exactly calculated. If the sample is i.i.d. from a general probability distribution, the solution of the stochastic linear program (SLP) is a consistent estimator of the SLP solution with respect to the underlying probability. An efficient geometrical algorithm is proposed to solve the SLP.

Keywords: Robust optimization, data depth, weighted-mean trimmed regions, central regions, coherent risk measure, spectral risk measure.

¹Graduate School of Risk Management, University of Cologne, Albertus-Magnus-Platz, D-50931 Cologne.

²Department of Economic and Social Statistics, University of Cologne, Meister-Ekkehart-Str. 9, D-50923 Cologne.

Stochastic linear programming with a distortion risk constraint

Pavel Bazovkin¹ and Karl Mosler²

¹Graduate School of Risk Management, University of Cologne,
Albertus-Magnus-Platz, D-50931 Cologne, Germany,
bazovkin@wiso.uni-koeln.de

²Department of Economic and Social Statistics, University of Cologne,
Meister-Ekkehart-Str. 9, D-50923 Cologne, Germany,
mosler@statistik.uni-koeln.de

November 17, 2011

Abstract: Linear optimization problems are investigated whose parameters are uncertain. We apply coherent distortion risk measures to capture the violation of restrictions. Such a model turns out to be appropriate for many applications and, principally, for the mean-risk portfolio selection problem. Each risk constraint induces an uncertainty set of coefficients, which comes out to be a weighted-mean trimmed region. We consider a problem with a single constraint. Given an external sample of the coefficients, the uncertainty set is a convex polytope that can be exactly calculated. If the sample is i.i.d. from a general probability distribution, the solution of the stochastic linear program (SLP) is a consistent estimator of the SLP solution with respect to the underlying probability. An efficient geometrical algorithm is proposed to solve the SLP.

Keywords: Robust optimization, data depth, weighted-mean trimmed regions, central regions, coherent risk measure, spectral risk measure.

1 Introduction

Uncertainty in the coefficients of a linear program is often handled by constraints on probability or on some other risk measure. The random restrictions are captured by imposing risk constraints on their violation. For each linear restriction a so called uncertainty set is constructed which consists of all coefficients satisfying the risk constraint. In particular, coherent distortion (= spectral) risk measures have been used in the SLP literature (Rockafellar and Uryasev (2000), Bertsimas and Brown (2009), a.o.) as they allow a sound interpretation in terms of expected utility with distorted probabilities. The corresponding uncertainty sets are convex bodies and, as it will be made precise in the sequel, come out to equal the weighted-mean trimmed regions of Dyckerhoff and Mosler (2011b).

Consider the linear program

$$\mathbf{c}'\mathbf{x} \longrightarrow \min \quad s.t. \quad \tilde{\mathbf{A}}\mathbf{x} \geq \mathbf{b}, \quad (1)$$

and assume that $\tilde{\mathbf{A}}$ is a stochastic $m \times d$ matrix and $\mathbf{b} \in \mathbb{R}^m$. This is a stochastic linear optimization problem. To handle the stochastic side conditions a *joint chance constraint*,

$$\text{Prob}[\tilde{\mathbf{A}}\mathbf{x} \geq \mathbf{b}] \geq 1 - \alpha, \quad (2)$$

may be introduced, with some reasonable probability α of violation, resulting in a *chance-constrained linear program*. Alternatively, the side conditions may be subjected to *separate chance constraints*,

$$\text{Prob}[\tilde{\mathbf{A}}_j\mathbf{x} \geq b_j] \geq 1 - \alpha_j, \quad j = 1 \dots m, \quad (3)$$

with $\tilde{\mathbf{A}}_j$ denoting the j -th row of $\tilde{\mathbf{A}}$. Each chance constraint limits the risk of violating the side condition by imposing a maximum probability α_j of its violation. The chance constraints (2) and (3) are special cases of *risk constraints*,

$$\mu^m(\tilde{\mathbf{A}}\mathbf{x} - \mathbf{b}) \leq \mathbf{0}, \quad \text{resp.} \quad (4)$$

$$\rho^1(\tilde{\mathbf{A}}_j\mathbf{x} - b_j) \leq 0, \quad j = 1 \dots m, \quad (5)$$

where μ^m can be represented by an m -variate set-valued *coherent risk measure* as a generalization of the univariate *distortion risk* measure ρ^1 (Acerbi,

2002). An SLP that minimizes $\mathbf{c}'\mathbf{x}$ subject to one of the restrictions (4) or (5) is called a *risk-constrained stochastic linear program*.

For stochastic linear programs (SLPs) in general and risk-constrained SLPs in particular, the reader is e.g. referred to Kall and Mayer (2010). What we call a risk measure here is mentioned there as a *quality measure*, and useful representations of the corresponding constraints are given. As most of the literature, Kall and Mayer (2010) focus on classes of SLPs with chance constraints that lead to convex programming problems, since these have obvious computational advantages; see also Prékopa (1995)). Our choice of the quality measure, besides its generality, enjoys a meaningful interpretation and, as we will show later, always generates a convex program.

In the case of a single constraint ($m = 1$) we notate

$$\rho(\tilde{\mathbf{a}}'\mathbf{x} - b) \leq 0. \quad (6)$$

A practically important example of an SLP with a single risk constraint (6) is the *portfolio selection problem*. Let $\tilde{r}_1, \dots, \tilde{r}_d$ be the returns on d assets and notate $\tilde{\mathbf{r}} = (\tilde{r}_1, \dots, \tilde{r}_d)'$. A convex combination of the assets' returns is sought, $\tilde{\mathbf{r}}'\mathbf{x} = \sum_{j=1}^d \tilde{r}_j x_j$, that has maximum expectation under a risk constraint and an additional deterministic constraint,

$$\max_{\mathbf{x} \in \mathcal{C}} E[\tilde{\mathbf{r}}'\mathbf{x}], \quad s.t. \quad \rho(\tilde{\mathbf{r}}'\mathbf{x}) \leq \rho_0, \quad (7)$$

where ρ is a risk measure, $\rho_0 \in \mathbb{R}$ is a given upper bound of risk, and \mathcal{C} is a deterministic set which restricts the coefficients x_k in some way. E.g., if short sales are excluded, $\mathcal{C} = \Delta^d$, Δ^d being the unit simplex in \mathbb{R}^d ,

$$\Delta^d = \{\mathbf{x} \in \mathbb{R}^d : \sum_{k=1}^d x_k = 1, x_k \geq 0 \quad \forall k\}.$$

Two special cases are well known. First, let $\rho(\tilde{\mathbf{r}}'\mathbf{x}) = \text{Prob}[\tilde{\mathbf{r}}'\mathbf{x} \leq -v_0]$ and $\rho_0 = \alpha$. Then the optimization problem (7) says: Maximize the mean return $E[\tilde{\mathbf{r}}'\mathbf{x}]$ under the restrictions $\mathbf{x} \in \mathcal{C}$ and

$$\text{V@R}_\alpha(\tilde{\mathbf{r}}'\mathbf{x}) \leq v_0.$$

That is, the *value at risk* V@R_α of the portfolio return must not exceed the bound v_0 . Second, let

$$\rho(\tilde{\mathbf{r}}'\mathbf{x}) = -\frac{1}{\alpha} \int_0^\alpha Q_{\tilde{\mathbf{r}}'\mathbf{x}}(t) dt, \quad (8)$$

where Q_Z signifies the quantile function of a random variable Z . This means that the *expected shortfall* of the portfolio return is employed in the risk restriction.

In practice, $\tilde{\mathbf{a}}$ has to be estimated from data. If the solution of the SLP is based on a sample of observed coefficient vectors $\mathbf{a}^1, \dots, \mathbf{a}^n \in \mathbb{R}^d$, that is, on an *external sample*, the SLP is mentioned as an *empirical risk-constrained SLP*. In other words, we assume that $\tilde{\mathbf{a}}$ follows an *empirical distribution* that gives equal mass $\frac{1}{n}$ to some observed points $\mathbf{a}^1, \dots, \mathbf{a}^n \in \mathbb{R}^d$. Rockafellar and Uryasev (2000) investigate an empirical stochastic program that arises in portfolio choice when the expected shortfall of a portfolio is minimized. They convert the objective into a function that is convex in the decision vector \mathbf{x} and optimize it by standard methods.

A more complex situation is investigated by Bertsimas and Brown (2009), who discuss the risk-constrained SLP with arbitrary coherent distortion risk measures, which also include expected shortfall. They determine the uncertainty set that turns the SLP into a minimax problem, called *robust linear program*; however they provide no optimal solution of this program. Natarajan et al. (2009), on the reverse, construct similar risk measures from given polyhedral and conic uncertainty sets. Pflug (2006) has proposed an iterative algorithm for optimizing a portfolio using distortion functionals, on each step adding a constraint to the problem and solving it by the simplex method. Meanwhile, many other authors have recently contributed to the development of robust linear programs related to risk-constrained optimization problems, see, e.g. Nemirovski and Shapiro (2006), Ben-Tal et al. (2009) and Chen et al. (2010) For a review of robust linear programs in portfolio optimization the reader is referred to Fabozzi et al. (2010).

In this paper we

1. provide a geometrical characterization of the uncertainty set of an empirical risk-constrained SLP when the risk is measured by a general coherent distortion risk measure.
2. construct an *algorithm* that solves the minimax problem over the uncertainty set, hence the SLP,
3. show that, if the external sample is i.i.d. from an underlying general probability distribution, the uncertainty set and the solution of the SLP are consistent estimators of the uncertainty set and the SLP solution.

The paper is organized as follows: In Section 2 we discuss constraints on distortion risk measures and their equivalence to uncertainty sets in the parameter space; we show further that these uncertainty sets are so called weighted-mean trimmed regions which satisfy a coherency property. In Section 3 a robust linear program is investigated by which the SLP with distortion risk constraint is solved. Section 4 introduces an algorithm for this program and discusses sensitivity issues of its solution. In Section 5 we address the SLP and its solution for generally distributed coefficients. Section 6 concludes.

2 Distortion risk constraints and weighted-mean regions

Let us consider a probability space $\langle \Omega, \mathcal{F}, P \rangle$ and a set \mathcal{R} of random variables (e.g., returns of portfolios). A function $\rho : \mathcal{R} \rightarrow \mathbb{R}$ is a law invariant *risk measure* if for $Y, Z \in \mathcal{R}$ it holds:

1. *Monotonicity*: $Y \geq Z \implies \rho(Y) \leq \rho(Z)$,
2. *Translation invariance*: $\rho(Y + \gamma) = \rho(Y) - \gamma$ for all $\gamma \in \mathbb{R}$,
3. *Law invariance*: If Y and Z have the same distribution, $P_Y = P_Z$, then $\rho(Y) = \rho(Z)$.

ρ is a *coherent risk measure* if it is, in addition, positive homogeneous and subadditive,

3. *Positive homogeneity*: $\rho(\lambda Y) = \lambda \rho(Y)$ for all $\lambda \geq 0$,
4. *Subadditivity*: $\rho(Y + Z) \leq \rho(Y) + \rho(Z)$ for all $Y, Z \in \mathcal{R}$.

The last two restrictions imply that diversification is encouraged. For the theory of such risk measures, see e.g. Föllmer and Schied (2004). We will employ a fundamental representation result of Huber (1981). A function $\rho : \mathcal{R} \rightarrow \mathbb{R}$ is said to satisfy the *Fatou property* if $\liminf_{n \rightarrow \infty} \rho(Y_n) \geq \rho(Y)$ for any bounded sequence converging pointwise to Y .

Proposition 1 (Huber, 1981). ρ is a coherent risk measure satisfying the Fatou property if and only if there exists a family \mathbb{Q} of probability measures that are dominated by P (i.e. $P(S) = 0 \Rightarrow Q(S) = 0$ for any $S \in \mathcal{F}$ and $Q \in \mathbb{Q}$) such that

$$\rho(Y) = \sup_{Q \in \mathbb{Q}} E_Q(-Y).$$

We say that the family \mathbb{Q} generates ρ . In particular, let $(\Omega, \mathcal{A}) = (\mathbb{R}, \mathcal{B}^d)$ and P be the probability distribution of a random vector $\tilde{\mathbf{a}}$. Huber's Theorem implies that for any coherent risk measure ρ there exists a family \mathbb{G} of P -dominated probabilities on \mathcal{B}^d so that

$$\begin{aligned} \rho(\tilde{\mathbf{a}}'\mathbf{x} - b) \leq 0 &\Leftrightarrow \rho(\tilde{\mathbf{a}}'\mathbf{x}) \leq -b \\ &\Leftrightarrow \inf_{G \in \mathbb{G}} E_G(\tilde{\mathbf{a}}'\mathbf{x}) \geq b \\ &\Leftrightarrow E_G(\tilde{\mathbf{a}}'\mathbf{x}) \geq b \text{ for all } G \in \mathbb{G}. \end{aligned}$$

If $\tilde{\mathbf{a}}$ has an empirical distribution on n given points in \mathbb{R}^d , any subset \mathcal{Q} of Δ^n corresponds to a family of P -dominated probabilities, and thus defines a coherent risk measure ρ . Then an immediate consequence of Huber's theorem is the following (cf. Bertsimas and Brown (2009)):

Proposition 2. Let $\rho : \mathcal{R} \rightarrow \mathbb{R}$ be a coherent risk measure and let $\tilde{\mathbf{a}}$ have an empirical distribution on $\mathbf{a}^1, \dots, \mathbf{a}^n \in \mathbb{R}^d$. Then there exists some $\mathcal{Q}_\rho \subset \Delta^n$ such that

$$\begin{aligned} \rho(\tilde{\mathbf{a}}'\mathbf{x} - b) \leq 0 &\Leftrightarrow \mathbf{a}'\mathbf{x} \geq b \text{ for all} \\ &\mathbf{a} \in \mathcal{U}_\rho := \text{conv}\{\mathbf{a} \in \mathbb{R}^d : \mathbf{a} = [\mathbf{a}^1, \dots, \mathbf{a}^n]\mathbf{q} : \mathbf{q} \in \mathcal{Q}_\rho\}. \end{aligned}$$

Here, $\text{conv}(W)$ denotes the convex closure of a set W . Proposition 2 says that a deterministic side condition $\mathbf{a}'\mathbf{x} \geq b$ holding uniformly for all \mathbf{a} in the uncertainty set \mathcal{U}_ρ is equivalent to the above risk constraint (6) on the stochastic side condition. This will be used below in providing an algorithmic solution of the risk-constrained SLP.

2.1 Distortion risk measures

An important subclass of coherent risk measures is the following (Acerbi (2002)). Again, let Q_Y denote the quantile function of a random variable Y .

Definition 1 (Distortion risk measure). *Let r be an increasing function $[0, 1] \rightarrow [0, 1]$. The risk measure ρ given by*

$$\rho(Y) = - \int_0^1 Q_Y(t) dr(t) \quad (9)$$

is a distortion risk measure with weight generating function r . It is coherent if and only if r is concave.

For example, with $r(t) = 0$ if $t < \alpha$ and $r(t) = 1$ if $t \geq \alpha$, the *value at risk* $V@R_\alpha(Y) = -Q_Y(\alpha)$ is obtained, which is a non-coherent distortion risk measure. An important example of a coherent distortion risk measure is the *expected shortfall*, which is yielded by $r(t) = t/\alpha$ if $t < \alpha$ and $r(t) = 1$ otherwise. Note that with $r(t) = t$, the risk measure becomes the expectation of $-Y$. A general distortion risk measure $\rho(Y)$ can thus be interpreted as the expectation of $-Y$ with respect to a probability distribution that has been distorted by the function r . In particular, a concave function r distorts the probabilities of lower outcomes of Y in positive direction (the lower the more) and conversely for higher outcomes (the higher the less). In empirical applications, coherent distortion risk measures other than expected shortfall have been recently used by many authors; see, e.g., Adam et al. (2008) for a comparison of various such measures in portfolio choice.

An equivalent characterization of a coherent distortion risk measure is that it is coherent and comonotonic (Acerbi (2002)). ρ is *comonotonic* if

$$\rho(Y + Z) = \rho(Y) + \rho(Z) \text{ for all } Y \text{ and } Z \text{ that are comonotonic,}$$

i.e., that satisfy $(Y(\omega) - Y(\omega'))(Z(\omega) - Z(\omega')) \geq 0$ for every $\omega \in \Omega$. If Y has an empirical distribution on $y_1, \dots, y_n \in \mathbb{R}$, a coherent distortion risk measure has the form

$$\rho(Y) = - \sum_{i=1}^n q_i y_{[i]}, \quad (10)$$

with some weights q_i satisfying $0 \leq q_1 \leq \dots \leq q_n$ and $\sum_{i=1}^n q_i = 1$, that is, $\mathbf{q} \in \Delta_{\leq}^n$, while $y_{[i]}$ are the values ordered from above.

2.2 Uncertainty sets are weighted-mean regions

If ρ is a coherent distortion risk measure, the uncertainty set \mathcal{U}_ρ has a special geometric structure, which we will explore now in order to visualize the optimization problem and to provide the basis for an algorithm. We will demonstrate that \mathcal{U}_ρ equals a so called *weighted-mean (WM) region* of the distribution of $\tilde{\mathbf{a}}$.

Given a general probability distribution on \mathbb{R}^d , weighted-mean regions form a nested family of convex compact sets that are affine equivariant and, by this, describe the distribution with respect to its location, dispersion and shape. Weighted-mean regions have been introduced in Dyckerhoff and Mosler (2011b) for empirical distributions, and in Dyckerhoff and Mosler (2011a) for general ones.

For an empirical distribution on $\mathbf{a}^1, \dots, \mathbf{a}^n \in \mathbb{R}^d$, a weighted-mean region is a polytope in \mathbb{R}^d and defined as

$$D_{\mathbf{w}_\alpha}(\mathbf{a}^1, \dots, \mathbf{a}^n) = \text{conv} \left\{ \sum_{j=1}^n q_{\alpha,j} \mathbf{a}^{\pi(j)} \mid \pi \text{ permutation of } \{1, \dots, n\} \right\}. \quad (11)$$

Here $\mathbf{w}_\alpha = [w_{\alpha,1}, \dots, w_{\alpha,n}]'$ is a vector of ordered weights, i.e. $\mathbf{w}_\alpha \in \Delta_{\leq}^n$, indexed by $0 \leq \alpha \leq \alpha_{max}$ that for $\alpha < \beta$ satisfies

$$\sum_{j=1}^k w_{\alpha,j} \leq \sum_{j=1}^k w_{\beta,j}, \quad k = 1, \dots, n. \quad (12)$$

Any such family of *weight vectors* $(\mathbf{w}_\alpha)_{0 \leq \alpha \leq \alpha_{max}}$ specifies a particular notion of weighted-mean regions. The class of weighted-mean regions is rather large. It contains well known trimmed regions like the zonoid regions, the expected convex hull regions and many others. For example,

$$w_{\alpha,j} = \begin{cases} \frac{1}{n\alpha} & \text{if } j > n - \lfloor n\alpha \rfloor, \\ \frac{n\alpha - \lfloor n\alpha \rfloor}{n\alpha} & \text{if } j = n - \lfloor n\alpha \rfloor, \\ 0 & \text{if } j < n - \lfloor n\alpha \rfloor, \end{cases}$$

$0 \leq \alpha \leq 1$, defines the *zonoid regions*. However some popular types of trimmed regions, such as Mahalanobis or halfspace regions, are no weighted-mean regions.

A WM region is characterized by its projections on lines. Note that each $\mathbf{p} \in S^{d-1}$, where S^{d-1} is a $(d-1)$ -variate unit sphere, yields a projection of the data $\mathbf{a}^1, \dots, \mathbf{a}^n$ on the line generated by \mathbf{p} and thus induces a permutation $\pi_{\mathbf{p}}$ of the data,

$$\mathbf{p}'\mathbf{a}^{\pi_{\mathbf{p}}(1)} \leq \mathbf{p}'\mathbf{a}^{\pi_{\mathbf{p}}(2)} \leq \dots \leq \mathbf{p}'\mathbf{a}^{\pi_{\mathbf{p}}(n)}.$$

The permutation is not necessarily unique, and let $H(\mathbf{a}^1, \dots, \mathbf{a}^n)$ denote the set of all directions $\mathbf{p} \in S^{d-1}$ that induce a non-unique permutation $\pi_{\mathbf{p}}$. Dyckerhoff and Mosler (2011b) have shown that the support function h_{α} of $D_{\mathbf{w}_{\alpha}} = D_{\mathbf{w}_{\alpha}}(\mathbf{a}^1, \dots, \mathbf{a}^n)$ amounts to

$$h_{\alpha}(\mathbf{p}) = \sum_{j=1}^n w_{\alpha,j} \mathbf{p}'\mathbf{a}^{\pi_{\mathbf{p}}(j)}, \quad \mathbf{p} \in S^{d-1}. \quad (13)$$

It follows that, whenever $\pi_{\mathbf{p}}$ is unique, the polytope $D_{\mathbf{w}_{\alpha}}$ has an extremal point in direction \mathbf{p} , which is given by

$$\sum_{j=1}^n w_{\alpha,j} \mathbf{a}^{\pi_{\mathbf{p}}(j)} = \sum_{i=1}^n w_{\alpha,\pi_{\mathbf{p}}^{-1}(i)} \mathbf{a}^i, \quad \mathbf{p} \in S^{d-1} \setminus H(\mathbf{a}^1, \dots, \mathbf{a}^n). \quad (14)$$

Now we are moving to the main result of this section, which will be the Theorem 1. From (10) and (14) we see that, with $q_i = w_{\alpha,\pi_{\mathbf{p}}^{-1}(i)}$ and $y_{[i]} = -\mathbf{p}'\mathbf{a}^i$, the extreme point of the projection of $D_{\mathbf{w}_{\alpha}}$ on the \mathbf{p} -line is obtained by applying a \mathbf{q} -distortion risk measure to the projected data points. Now, setting

$$\mathcal{Q}_{\rho} = \{\mathbf{q} \in \Delta^n : \mathbf{q} = (w_{\alpha,\pi_{\mathbf{p}}^{-1}(1)}, \dots, w_{\alpha,\pi_{\mathbf{p}}^{-1}(n)}), \mathbf{p} \in S^{d-1} \setminus H(\mathbf{a}^1, \dots, \mathbf{a}^n)\}$$

and $\mathcal{U}_{\rho} = \text{conv}\{\mathbf{a} : \mathbf{a} = \sum_{i=1}^n q_i \mathbf{a}^i, \mathbf{q} \in \mathcal{Q}_{\rho}\}$ we obtain that all extreme points of $D_{\mathbf{w}_{\alpha}}$ are in \mathcal{U}_{ρ} , hence $D_{\mathbf{w}_{\alpha}} \subseteq \mathcal{U}_{\rho}$. On the other hand, for every $\mathbf{q} \in \mathcal{Q}_{\rho}$ it holds that $\sum_{i=1}^n q_i \mathbf{a}^i \in D_{\mathbf{w}_{\alpha}}$, which implies $\mathcal{U}_{\rho} \subseteq D_{\mathbf{w}_{\alpha}}$. We conclude $\mathcal{U}_{\rho} = D_{\mathbf{w}_{\alpha}}$.

Thus we have proven the equality between the distortion risk constraint feasible set and a properly chosen WM region, which is a reformulation of Theorem 4.3 from Bertsimas and Brown (2009):

Theorem 1.

$$\{\mathbf{x} \in \mathbb{R}^d : \rho(\tilde{\mathbf{a}}'\mathbf{x} - b) \leq 0\} = \{\mathbf{x} \in \mathbb{R}^d : \mathbf{a}'\mathbf{x} \geq b \forall \mathbf{a} \in D_{\mathbf{w}_{\alpha}}(\mathbf{a}^1, \dots, \mathbf{a}^n)\}, \quad (15)$$

where $\mathbf{a}^1, \dots, \mathbf{a}^n$ is an external sample from the parameter vector $\tilde{\mathbf{a}}$.

Recall that $D_{\mathbf{w}_\alpha}(\mathbf{a}^1, \dots, \mathbf{a}^n)$ is a d -dimensional *convex polytope*, and thus the convex hull of a finite number of points (its vertices) or, equivalently, a bounded nonempty intersection of a finite number of closed halfspaces (that contain its facets). By this the calculation and representation of such a polytope can be done in two ways: either by its vertices or by its facets. Recall that a nonempty intersection of the polytope's boundary with a hyperplane is a *facet* if it has an affine dimension $d - 1$, and a *ridge* if it has an affine dimension $d - 2$. It is called an *edge* if it is a line segment, and a *vertex* if it is a single point. In general, each facet of a polytope in \mathbb{R}^d is itself a polytope of dimension $d - 1$ and has at least d vertices. With WM regions the number of a facet's vertices can vary considerably; it ranges between d and $d!$ (Bazovkin and Mosler, 2010a). That is why in calculating WM regions a representation by facets is preferable.

The vertices of a polytope are its extreme points. From above we know that the directions $\mathbf{p} \in S^{d-1} \setminus H(\mathbf{a}^1, \dots, \mathbf{a}^n)$ belong to vertices, while the directions $\mathbf{p} \in H(\mathbf{a}^1, \dots, \mathbf{a}^n)$ belong to parts of the boundary that have affine dimension ≥ 1 .

In the context of risk measurement it is crucial that the WM regions possess two properties that enable them to generate coherent risk measures: *monotonicity* and *subadditivity*. The following proposition is due to Dyckerhoff and Mosler (2011b):

Proposition 3. (Coherency properties of WM regions)

1. *Monotonicity: If $\mathbf{z}_k \leq \mathbf{y}_k$ holds for all k (in the componentwise ordering of \mathbb{R}^d), then*

$$\begin{aligned} D_{\mathbf{w}_\alpha}(\mathbf{y}_1, \dots, \mathbf{y}_n) &\subset D_{\mathbf{w}_\alpha}(\mathbf{z}_1, \dots, \mathbf{z}_n) \oplus \mathbb{R}_+^d, \quad \text{and} \\ D_{\mathbf{w}_\alpha}(\mathbf{z}_1, \dots, \mathbf{z}_n) &\subset D_{\mathbf{w}_\alpha}(\mathbf{y}_1, \dots, \mathbf{y}_n) \oplus \mathbb{R}_-^d. \end{aligned}$$

2. *Subadditivity:*

$$D_{\mathbf{w}_\alpha}(\mathbf{y}_1 + \mathbf{z}_1, \dots, \mathbf{y}_n + \mathbf{z}_n) \subset D_{\mathbf{w}_\alpha}(\mathbf{y}_1, \dots, \mathbf{y}_n) \oplus D_{\mathbf{w}_\alpha}(\mathbf{z}_1, \dots, \mathbf{z}_n).$$

In this Proposition the symbol \oplus is the *Minkowski addition*, $A \oplus B = \{a + b : a \in A, b \in B\}$ for A and $B \subset \mathbb{R}^d$.

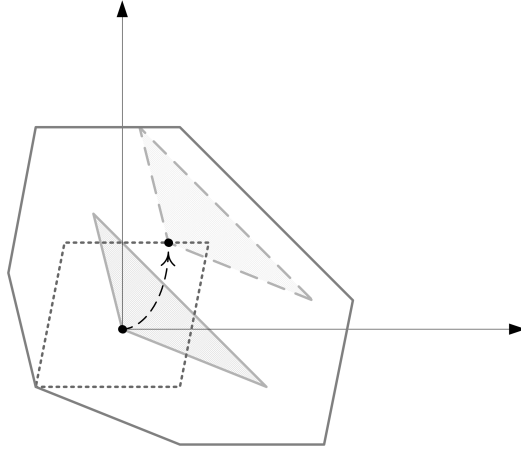


Figure 1: An illustration of the Minkowski addition.

The subadditivity property has a nice interpretation as a dilation of one trimmed region by the other. To understand this better let us consider the simple example of Minkowski addition given in Figure 1. The figure exhibits a solid triangle with one vertex at the origin and a dotted-border quadrangle. Now move the triangle in a way that its lower left corner passes all points of the quadrangle. At each point of the quadrangle we get a copy of the initial triangle (with a dashed border) shifted by coordinate of the point. The union of all these triangles gives us the Minkowski sum of the initial two sets, which is the big heptagon in the picture. Observe that, if we move the rectangle around the triangle, the same sum is obtained. The subadditivity states that if, e.g., these two figures are WM regions $D_{\mathbf{w}_\alpha}(\mathbf{y}_1, \dots, \mathbf{y}_n)$ and $D_{\mathbf{w}_\alpha}(\mathbf{z}_1, \dots, \mathbf{z}_n)$ respectively, the $D_{\mathbf{w}_\alpha}(\mathbf{y}_1 + \mathbf{z}_1, \dots, \mathbf{y}_n + \mathbf{z}_n)$ is contained by the heptagon.

3 Solving the SLP with distortion risk constraint

3.1 Calculating the uncertainty set

In the previous section we have shown that the uncertainty set \mathcal{U}_ρ equals the weighted-mean (WM) region $D_{\mathbf{w}_\alpha}$ for a properly chosen weight vector \mathbf{w}_α .

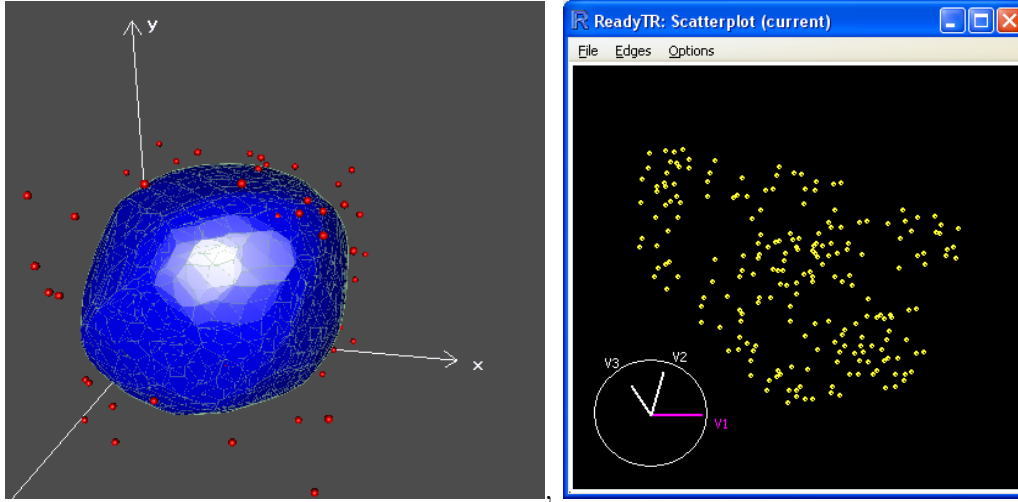


Figure 2: Visualization of WM regions by the R package *WMTregions*. Left panel: Facets of a three-dimensional region in \mathbb{R}^3 . Right panel: Vertices of a four-dimensional region projected on a subspace of \mathbb{R}^3 .

Bazovkin and Mosler (2010a) provide an algorithm by which this WM region can be exactly calculated in any dimension d . The algorithm is implemented as an R package (<http://www.r-project.org/>), *WMTregions* (Bazovkin and Mosler, 2010b). The R package allows representing a WM region either by its vertices or by an intersection of halfspaces, giving for each of them the normal vector and the intercept. The results can be visualized in dimensions two and three; for examples see Figure 2.

It has been also shown in Bazovkin and Mosler (2010a) that the number of vertices of a facet can be as much as $d!$. Therefore the representation of a WM region by its vertices appears to be less efficient than that by its facets. In the sequel, we will use the facet representation for solving the SLP.

3.2 The robust linear program

The robust linear program to be solved is

$$\mathbf{c}'\mathbf{x} \longrightarrow \min \quad s.t. \quad \mathbf{a}'\mathbf{x} \geq b \quad \text{for all } \mathbf{a} \in \mathcal{U}, \quad (16)$$

where the subscript ρ has been dropped for convenience. We rewrite the side condition as

$$\mathbf{x} \in \mathcal{X} = \bigcap_{\mathbf{a} \in \mathcal{U}} X_{\mathbf{a}}, \quad X_{\mathbf{a}} = \{\mathbf{x} : \mathbf{a}'\mathbf{x} \geq b\}. \quad (17)$$

Note that \mathcal{X} is a convex polyhedron. So, a linear goal function is to be minimized on a convex polyhedron. Obviously, any optimal solution will lie on the surface of \mathcal{X} .

3.3 Finding the optimum on the uncertainty set

In constructing an algorithm for the robust linear program, we will explore the set \mathcal{X} of feasible solutions and relate it to the uncertainty set \mathcal{U} in the parameter space. We will show that the space of solutions \mathbf{x} and the space of coefficients \mathbf{a} are *dual* to each other, however the duality is not reflexive. The following two lemmas provide the connection between \mathcal{X} and \mathcal{U} . First we demonstrate that \mathcal{X} is the intersection of those halfspaces whose normals are extreme points of \mathcal{U} .

Lemma 1. *It holds that*

$$\mathcal{X} = \bigcap_{\mathbf{a} \in \mathcal{U}} \{\mathbf{x} : \mathbf{a}'\mathbf{x} \geq b\} = \bigcap_{\mathbf{a} \in \text{ext}\mathcal{U}} \{\mathbf{x} : \mathbf{a}'\mathbf{x} \geq b\}.$$

Proof. We show that $\bigcap_{\mathbf{a} \in \text{ext}\mathcal{U}} X_{\mathbf{a}} \subset X_{\mathbf{u}}$ for all $\mathbf{u} \in \mathcal{U}$; then $\bigcap_{\mathbf{a} \in \text{ext}\mathcal{U}} X_{\mathbf{a}} \subset \bigcap_{\mathbf{a} \in \mathcal{U}} X_{\mathbf{a}}$. The opposite inclusion is obvious. Assume $\mathbf{u} \in \mathcal{U}$. Then, as \mathcal{U} is convex and compact, \mathbf{u} is a convex combination of some points $\mathbf{a}^1, \dots, \mathbf{a}^\ell \in \text{ext}\mathcal{U}$, i.e. $\mathbf{u} = \sum_{i=1}^{\ell} \lambda_i \mathbf{a}^i$ with $\lambda_i \geq 0$ and $\sum_{i=1}^{\ell} \lambda_i = 1$, and for any $\mathbf{x} \in \bigcap_{\mathbf{a} \in \text{ext}\mathcal{U}} X_{\mathbf{a}}$ holds $\mathbf{x} \in X_{\mathbf{a}^j}$ and $\mathbf{a}^{j'}\mathbf{x} \geq b$ for all j , hence $\mathbf{u}'\mathbf{x} = \sum_{i=1}^{\ell} \lambda_i \mathbf{a}^{i'}\mathbf{x} \geq b$, that is, $\mathbf{x} \in X_{\mathbf{u}}$. \square

Lemma 1 says that each facet of the set \mathcal{X} of feasible solutions corresponds to a vertex of the uncertainty set \mathcal{U} . Hence it is sufficient to consider the extreme points of the uncertainty set.

As a generalization of Lemma 1, we may prove by recursion on k : Each k -dimensional face of the feasible set corresponds to a $(d - k)$ -dimensional face of the uncertainty set in the solution space. This resembles the dual correspondence between convex sets and their *polars* (cf. e.g. Rockafellar

(1997)). However, in contrast to polars, in our case the correspondence is not reflexive.

From Lemma 1 it is immediately seen, how the robust optimization problem contrasts with a deterministic problem, where the empirical distribution of $\tilde{\mathbf{a}}$ concentrates at some $\mathbf{u} \in \mathcal{U}$. Observe that the deterministic feasible set \mathcal{X} is just a halfspace, $\mathcal{X}_{\mathbf{u}} = \{\mathbf{x} : \mathbf{u}'\mathbf{x} \geq b\}$. In the general robust case a halfspace is obtained for each $\mathbf{a} \in \text{ext}\mathcal{U}$, and the robust feasible set \mathcal{X} is their intersection. The halfspaces are bounded by hyperplanes with normals equal to $\mathbf{a} \in \text{ext}\mathcal{U}$, and their intercepts are all the same and equal to b . Obviously, it follows that the robust feasible set \mathcal{X} is always included in the deterministic feasible set $\mathcal{X}_{\mathbf{u}}$,

$$\mathcal{X} \subseteq \mathcal{X}_{\mathbf{u}} \quad \text{for any } \mathbf{u} \in \mathcal{U}.$$

Moreover, the two feasible sets cannot be equal unless each element of \mathcal{U} is a scalar multiple of \mathbf{u} with a factor greater than one, $\mathcal{U} \subseteq \{\mathbf{a} : \mathbf{a} = \lambda\mathbf{u}, \lambda > 1\}$. Consequently, the minimum value of the robust stochastic LP cannot be smaller than the value of an LP with any deterministic parameter \mathbf{u} chosen from the uncertainty set. Figure 3 (left panel) illustrates how a deterministic feasible set in dimension two compares to a general robust one: The line that bounds the halfspace $\mathcal{X}_{\mathbf{u}}$ ‘folds’ into a piecewise linear curve delimiting \mathcal{X} .

Let

$$U_{\mathbf{x}} = \{\mathbf{a} \in \mathbb{R}^d : \mathbf{a}'\mathbf{x} \geq b\}, \quad \mathbf{x} \in \mathbb{R}^d.$$

Lemma 2. *It holds that*

$$\mathcal{U} \subset \bigcap_{\mathbf{x} \in \mathcal{X}} U_{\mathbf{x}} \subset \bigcap_{\mathbf{x} \in \text{ext } \mathcal{X}} U_{\mathbf{x}}.$$

Moreover, each vertex $\mathbf{x} \in \text{ext } \mathcal{X}$ corresponds to a facet of \mathcal{U} .

Proof. By Lemma 1 we have $\mathbf{x} \in \mathcal{X} \Leftrightarrow \mathbf{a}'\mathbf{x} \geq b$ for all $\mathbf{a} \in \mathcal{U}$. Now let $\mathbf{a} \in \mathcal{U}$; then for any $\mathbf{x} \in \mathcal{X}$ it holds that $\mathbf{a}'\mathbf{x} \geq b$, hence $\mathbf{a} \in U_{\mathbf{x}}$. Conclude $\mathcal{U} \subset \bigcap_{\mathbf{x} \in \mathcal{X}} U_{\mathbf{x}}$. Further, it is clear that an extreme point $\mathbf{x} \in \text{ext } \mathcal{X}$ yields a facet of \mathcal{U} . \square

Remark. While \mathcal{U} is always compact, \mathcal{X} is in general not. Therefore neither inclusion holds with equality.

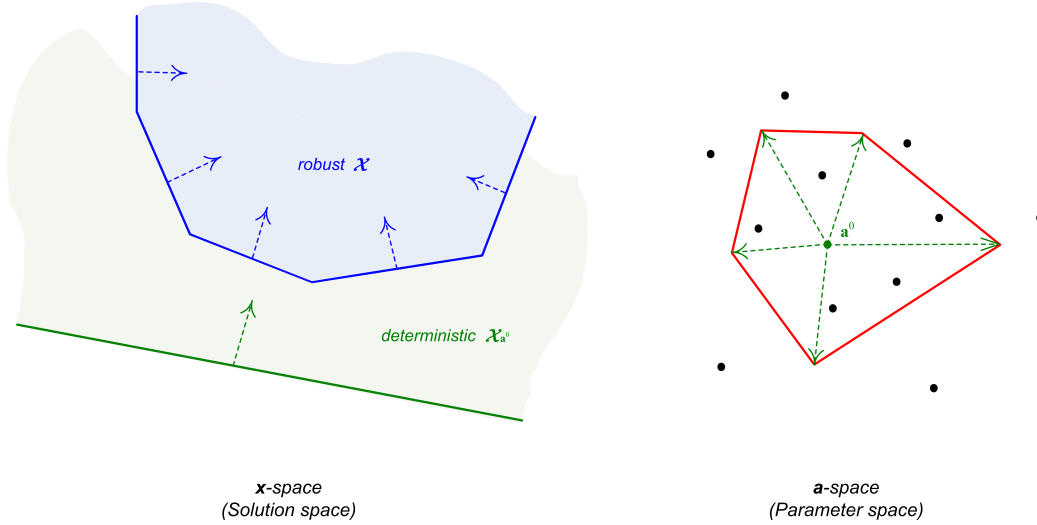


Figure 3: Deterministic and robust cases: feasible set (left panel), uncertainty set (right panel).

The ordinary simplex algorithm, operating on the vertices of \mathcal{X} , constructs a chain of adjacent facets in the space of parameters. The chain ends at the solution of the optimization task. Notice that this chain corresponds to a chain of facets of the uncertainty set. So, in principle we could try to calculate this chain of facets in the parameter set. However, in our algorithm, another way is pursued to find the optimal solution.

To manage this task let us consider the goal function $\mathbf{c}'\mathbf{x}$. In the parameter space \mathbf{c} corresponds to a point or a direction. In the solution space it corresponds to all hyperplanes that have \mathbf{c} as their normal. To produce all these hyperplanes in the parameter space we have to multiply \mathbf{c} with some scaling factor. Hence the hyperplanes are obtained by passing through a straight line φ that contains the origin and \mathbf{c} .

Next we search the intersection of \mathcal{U} with the line φ . Note that finding the intersection of a line and a polyhedron in \mathbb{R}^3 is an important problem in computer graphics (cf. Kay and Kajiya (1986)). We employ the same principle for a general dimension d . The uncertainty set \mathcal{U} is the finite intersection of halfspaces \mathcal{H}_j , $j = 1 \dots J$, each being defined by a hyperplane H_j with normal \mathbf{n}_j pointing into \mathcal{H}_j and an intercept d_j .

Consider some point \mathbf{u} on the line φ that is not in \mathcal{U} . Compute $\frac{d_j}{\mathbf{u}'\mathbf{n}_j}$ for all

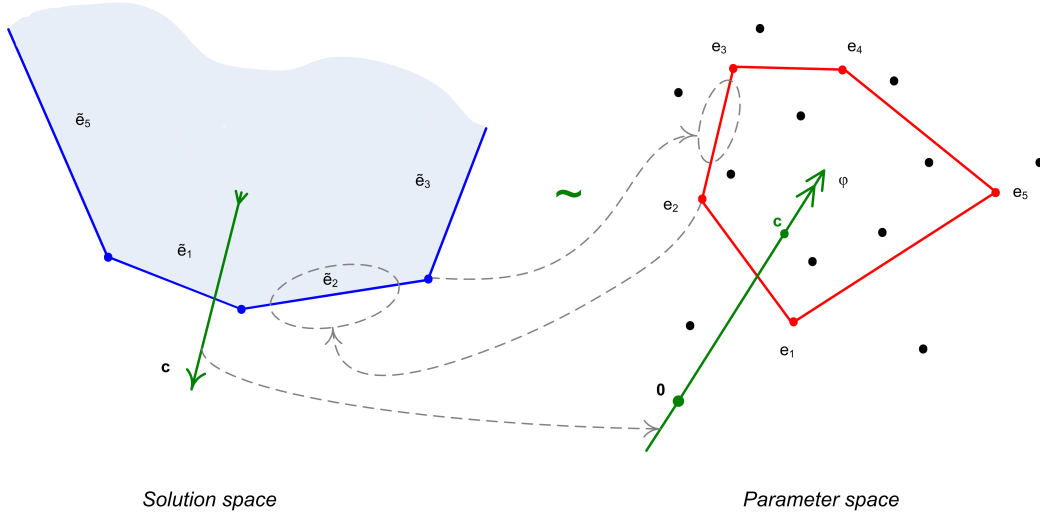


Figure 4: Duality between spaces.

halfspaces \mathcal{H}_j that do *not* include \mathbf{u} , i.e. where $(\mathbf{u}'\mathbf{n}_j - d_j) < 0$ holds. (In other words, H_j is *visible* from \mathbf{u} .) Find j_* at which this value is largest. Recall that moving a point \mathbf{u} along φ is equivalent to multiplying \mathbf{u} by some constant. The furthest move is given by the biggest constant. The *optimal solution* \mathbf{x}^* of the robust SLP has to satisfy $\mathbf{a}'\mathbf{x}^* \geq b$, which is equivalent to

$$\mathbf{a}' \left(\frac{d_{j_*}}{b} \mathbf{x}^* \right) \geq d_{j_*} .$$

Hence, to obtain \mathbf{x}^* , we have to scale the normal \mathbf{n}_{j_*} with the constant $\frac{b}{d_{j_*}}$,

$$\mathbf{x}^* = \frac{b}{d_{j_*}} \mathbf{n}_{j_*} .$$

Finally, we like to point out that we need not calculate the whole polytope \mathcal{U} but only a part of it which intersects the line φ . In searching for the optimum not all F facets need to be checked, but only a subset of the surface where the intersection will happen. Such a filtration makes the procedure more efficient. The search for a proper subset can be driven by *geometrical* considerations.

A subset \mathcal{U}_{eff} of \mathcal{U} will be mentioned as an *efficient parameter set* if

$$\bullet \quad \mathcal{X} \supset \mathcal{X}_{\text{eff}} = \bigcap_{\mathbf{a} \in \mathcal{U}_{\text{eff}}} \{ \mathbf{x} : \mathbf{a}'\mathbf{x} \geq b \} \ni \mathbf{x}^* \quad \text{and}$$

- $\mathbf{a}, \mathbf{a}_* \in \mathcal{U}_{\text{eff}}, \mathbf{a}'\mathbf{x} \geq b, \mathbf{a}'_*\mathbf{x} \geq b,$ implies $\mathbf{a} = \mathbf{a}_*$.

That is to say, \mathcal{U}_{eff} is the minimal subset of \mathcal{U} containing all facets that can be optimal for some \mathbf{c} .

Proposition 4. \mathcal{U}_{eff} is the union of all facets of \mathcal{U} for which $d_j \geq 0$ holds.

In other words, an efficient parameter set \mathcal{U}_{eff} consists of that part of the surface of \mathcal{U} that is visible from the origin $\mathbf{0}$. The proof is obvious.

To visualize the efficient parameter set we will use the *augmented uncertainty set*, which is defined as

$$\{\mathbf{a} : \mathbf{a} = \lambda \mathbf{a}^*, \lambda > 1, \mathbf{a}^* \in \mathcal{U}_{\text{eff}}\}.$$

It includes all parameters that are dominated by \mathcal{U}_{eff} ; see the shaded area in the right panel of Figure 5.

4 The algorithm

In this part we give an accurate procedure of obtaining the optimal solution.

Input:

- a vector $\mathbf{c} \in \mathbb{R}^d$ of coefficients of the goal function,
- an external sample $\{\mathbf{a}^1, \dots, \mathbf{a}^n\} \subset \mathbb{R}^d$ of coefficient vectors of the restriction,
- a right-hand side $b \in \mathbb{R}$ of the restriction,
- a distortion risk measure ρ (defined either by name or by a weight vector).

Output:

- the uncertainty set \mathcal{U} of parameters given by
 - facets (i.e. normals and intercepts),

- vertices,
- the optimal solution \mathbf{x}^* of the robust LP and its value $\mathbf{c}'\mathbf{x}^*$.

Steps:

- A. Calculate the subset $\mathcal{U}_{\text{eff}} \subset \mathcal{U}$ consisting of facets $\{(\mathbf{n}_j, d_j)\}_{j \in J}$.
- B. Create a line φ passing through the origin $\mathbf{0}$ and \mathbf{c} .
- C. Search for a facet H_{j_*} of \mathcal{U}_{eff} that is intersected by φ :
 - a. Select a subset $\mathcal{U}_{\text{sel}} \subseteq \mathcal{U}_{\text{eff}}$ of facets: This may be either \mathcal{U}_{eff} itself or its part where the intersection is expected; $\mathcal{U}_{\text{sel}} = \{(\mathbf{n}_j, d_j) : j \in J_{\text{sel}}\}$. For example, we can search the best solution on a pre-given subset of parameters. The other possible filtration is iterative transition to a facet with better criterion value.
 - b. Take a point $\mathbf{u} = \lambda \mathbf{c}$, $\lambda \geq 0$, outside the augmented uncertainty set. Find the $j_* = \arg \max_j \left\{ \frac{d_j}{\mathbf{u}'\mathbf{n}_j} \right\}_{j \in J_{\text{sel}} \subseteq J}$.
 - c. $\mathbf{x}^* = \frac{b}{d_{j_*}} \mathbf{n}_{j_*}$ is the optimal solution of the robust LP.

In fact, the line φ consists of points that correspond to hyperplanes whose normal is the vector \mathbf{c} in the dual space. One part of φ is dominated by points from \mathcal{U}_{eff} , while the other is not (which results from Proposition 4). The crossing point \mathbf{a}^* defines the hyperplane that touches the feasible set at the optimum as its dual.

Finally, we should point out special cases that may arise:

1. φ does not intersect \mathcal{U}_{eff} . It means that we get an infinite solution.
2. The inner part of \mathcal{U}_{eff} contains the origin. It means that there is no solution.

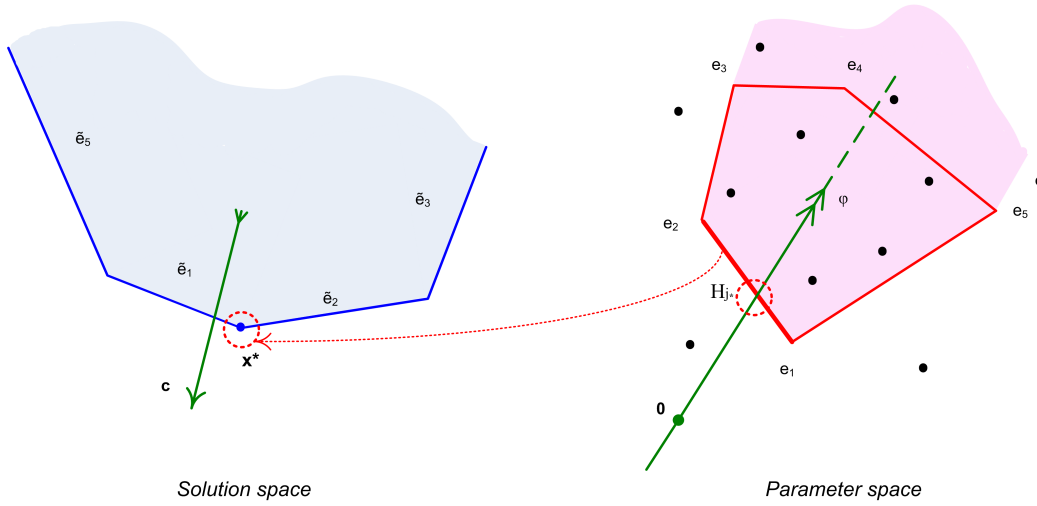


Figure 5: Finding the optimal solution on the uncertainty set.

4.1 Sensitivity issues

Next we like to discuss how the robust SLP and its optimal solution behave when the data $\{\mathbf{a}^1, \dots, \mathbf{a}^n\}$ on the coefficients are slightly changed. From (13) it is immediately seen that the support function $h_{\mathcal{U}}$ of the uncertainty set is continuous in the data \mathbf{a}^j as well as in the weight vector \mathbf{w}_α . (Note that the support function $h_{\mathcal{U}}$ is even *uniformly continuous* in $\mathbf{a}^1, \dots, \mathbf{a}^n$ and \mathbf{w}_α , which is tantamount saying that the uncertainty set \mathcal{U} is *Hausdorff continuous* in the data and the risk weights.) Consequently, a slight perturbation of the data will only slightly change the value of the support function of \mathcal{U} , which is a practically useful result regarding the sensitivity of the uncertainty set with respect to the data. The same is true for a small change in the weights of the risk measure.

We conclude that the point \mathbf{a}^{j*} where the line through the origin and \mathbf{c} cuts \mathcal{U} depends continuously on the data and the weights. However this is not true for the optimal solution \mathbf{x}^* , which may ‘jump’ when the cutting point moves from one facet of \mathcal{U} to a neighboring one.

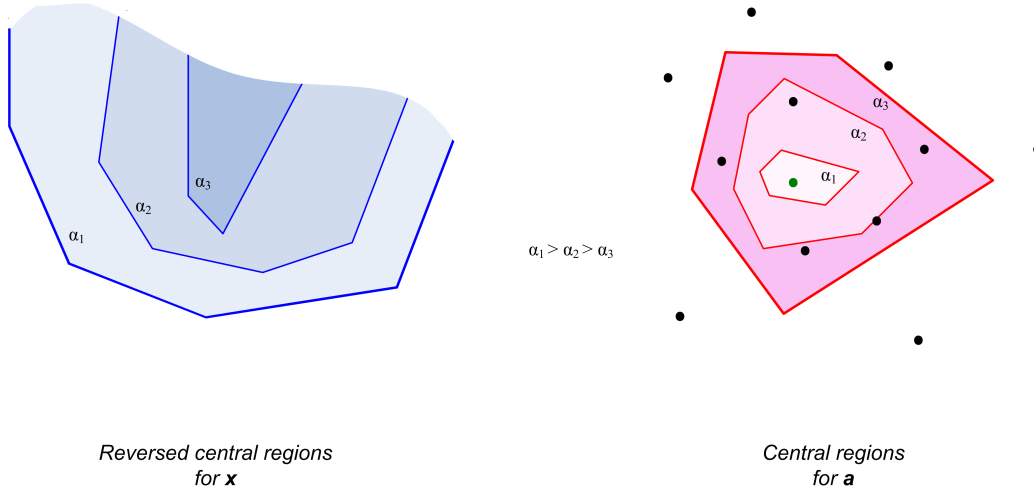


Figure 6: Example of the ‘reversed’ central regions in the dimension 2.

4.2 Ordered sensitivity analysis

We may also compare alternative uncertainty sets that are ordered by inclusion. From Lemma 1 it is clear that the respective sets of feasible solutions are then ordered in the reverse direction; see e.g. Figure 6). In particular we may consider the robust LP for two alternative distortion risk measures which are based on weight vectors \mathbf{w}_α and \mathbf{w}_β , respectively, that satisfy the monotonicity restriction (12). Then the resulting uncertainty sets are nested, $\mathcal{U}_\beta \subset \mathcal{U}_\alpha$ and so are, reversely, the feasible sets, $\mathcal{X}_\beta \supset \mathcal{X}_\alpha$. This is a useful approach for visualizing the sensitivity of the robust LP against changes in risk evaluation.

5 Robust SLP for generally distributed coefficients

So far we have considered the SLP (1) when the coefficient vector $\tilde{\mathbf{a}}$ follows an empirical distribution and have solved it on the basis of an external sample

$\{\mathbf{a}^1, \dots, \mathbf{a}^n\}$. In this section the SLP will be addressed with a general probability distribution P of $\tilde{\mathbf{a}}$. We formulate the robust SLP in the general case and demonstrate that the solution of this SLP can be consistently estimated by random sampling from P .

Consider a distortion risk measure ρ (9) that measures the risk of a general random variable Y and has weight generating function r , $\rho(Y) = -\int_0^1 Q_Y(t)dr(t)$. Similarly as in Section 2.2 a convex compact \mathcal{U} in \mathbb{R}^d is constructed through its support function $h_{\mathcal{U}}$,

$$h_{\mathcal{U}}(\mathbf{p}) = \int_0^1 Q_{\mathbf{p}'\tilde{\mathbf{a}}}(t)dr(t).$$

For details, see Dyckerhoff and Mosler (2011a).

Now, let a sequence $(\tilde{\mathbf{a}}^n)_{n \in \mathbb{N}}$ of independent random vectors be given that are identically distributed with P , and consider the sequence of random uncertainty sets \mathcal{U}_n based on $\tilde{\mathbf{a}}^1, \dots, \tilde{\mathbf{a}}^n$. Dyckerhoff and Mosler (2011b) have shown:

Proposition 5 (Dyckerhoff and Mosler (2011b)). *\mathcal{U}_n converges to \mathcal{U} almost surely in the Hausdorff sense.*

The proposition implies that by drawing an independent sample of $\tilde{\mathbf{a}}$ and solving the robust LP based on the observed empirical distribution we obtain a consistent estimate of the uncertainty set \mathcal{U} . Moreover, the cutting point \mathbf{a}^{j^*} , where the line through the origin and \mathbf{c} hits the uncertainty set, is consistently estimated by our algorithm. But, in particular for a discretely distributed $\tilde{\mathbf{a}}$, the optimal solution \mathbf{x}^* need not be a consistent estimate, as it may perform a jump when \mathbf{a}^{j^*} moves from one facet of \mathcal{U} to neighboring one.

6 Discussion and conclusions

We have shown that any constraint on a coherent distortion risk measure corresponds to an uncertainty set of parameters, which is a weighted-mean trimmed region. The respective risk-constrained SLP is equivalent to a robust linear program with parameters restricted to the weighted-mean trimmed

region. Its solution is essentially obtained by calculating the relevant parts of the weighted-mean trimmed region. An algorithm has been proposed to solve the robust linear optimization problem. It is based on constructing the relevant part of the uncertainty set, which builds on an algorithm given in (Bazovkin and Mosler, 2010a).

Our approach can be used very flexibly. E.g., non-sample information may be introduced into the procedure in an interactive way by explicitly changing and modifying the uncertainty set.

A realization of the algorithm is available as an R package *StochaTR* (Bazovkin and Mosler, 2011). It is able to solve rather large problems efficiently. For example, for the simulated data with $d = 15$ and $n = 1500$ using zonoid regions of $\alpha = 0.05$ the solution is found in forty seconds, while for $n = 2500$ sixty seconds are needed. For the dimension $d = 20$ and 1500 points the algorithm needs approximately five minutes.

The algorithm can be modified to allow for additional deterministic constraints. In particular, a nonnegativity constraint, $\mathbf{x} \geq \mathbf{0}$, can be handled by replacing the uncertainty set \mathcal{U} with $\mathcal{U} \oplus \mathbb{R}_+^d$. Obviously with this uncertainty set the algorithm selects only facets that have nonnegative normals.

Future research will tackle the SLP with multiple constraints of the form (4), as well as a stochastic right-hand side $\tilde{\mathbf{b}}$ in the constraints and a random vector $\tilde{\mathbf{c}}$ in the goal function.

References

- Acerbi, C. 2002. Spectral measures of risk: A coherent representation of subjective risk aversion. *Journal of Banking and Finance* **26** 1505–1581.
- Adam, Alexandre, Mohamed Houkari, Jean-Paul Laurent. 2008. Spectral risk measures and portfolio selection. *Journal of Banking and Finance* **32** 1870–1882.
- Bazovkin, Pavel, Karl Mosler. 2010a. An exact algorithm for weighted-mean trimmed regions in any dimension. Working Paper, University of Cologne.
- Bazovkin, Pavel, Karl Mosler. 2010b. *WMTregions: Exact calculation of weighted-mean trimmed regions*. R package version 3.1.0, URL <http://CRAN.R-project.org/package=WMTregions>.

- Bazovkin, Pavel, Karl Mosler. 2011. *StochaTR: Solving stochastic linear programs with a single risk constraint*. R package version 1.0.4, URL <http://CRAN.R-project.org/package=StochaTR>.
- Ben-Tal, A., L. El Ghaoui, A. Nemirovski. 2009. *Robust Optimization*. Princeton University Press.
- Ben-Tal, A., A. Nemirovski. 2000. Robust solutions of linear programming problems contaminated with uncertain data. *Math. Programming Ser. A* **88** 411–424.
- Bertsimas, Dimitris, David B. Brown. 2009. Constructing uncertainty sets for robust linear optimization. *Operations Research* **57**(6) 1483–1495.
- Chen, Wenqing, Melvyn Sim, Jie Sun, Chung-Piaw Teo. 2010. From cvar to uncertainty set: Implications in joint chance-constrained optimization. *Oper. Res.* **58** 470–485. doi:<http://dx.doi.org/10.1287/opre.1090.0712>. URL <http://dx.doi.org/10.1287/opre.1090.0712>.
- Delbaen, F. 2002. Coherent risk measures on general probability spaces. *Advances in Finance and Stochastics*. Springer, Berlin, 1–37.
- Dyckerhoff, R., K. Mosler. 2011a. Weighted-mean regions of a probability distribution. Submitted.
- Dyckerhoff, Rainer, Karl Mosler. 2011b. Weighted-mean trimming of multivariate data. *Journal of Multivariate Analysis* **102** 405–421.
- Fabozzi, F. J., D. Huang, G. Zhou. 2010. Robust portfolios: contributions from operations research and finance. *Annals of Operations Research* **176** 191–220.
- Föllmer, H., A. Schied. 2004. *Stochastic Finance: An Introduction in Discrete Time*. Walter de Gruyter, Berlin.
- Huber, P. J. 1981. *Robust statistics*. Wiley, New York.
- Kall, P., J. Mayer. 2010. *Stochastic Linear Programming. Models, Theory, and Computation*. 2nd ed. Springer, New York.

- Kay, Timothy L., James T. Kajiya. 1986. Ray tracing complex scenes. *SIG-GRAPH Comput. Graph.* **20** 269–278. doi:<http://doi.acm.org/10.1145/15886.15916>. URL <http://doi.acm.org/10.1145/15886.15916>.
- Natarajan, K., D. Pachamanova, M. Sim. 2009. Constructing risk measures from uncertainty sets. *Operations Research* **57**(5) 1129–1141.
- Nemirovski, A., A. Shapiro. 2006. Convex approximations of chance constrained programs. *SIAM Journal on Optimization* **17** 969–996.
- Pflug, Georg Ch. 2006. On distortion functionals. *Statistics & Decisions* **24**(1) 45–60. URL <http://dx.doi.org/10.1524/std.2006.24.1.45>.
- Prékopa, A. 1995. *Stochastic Programming*. Kluwer Academic Publ.
- Rockafellar, R. T. 1997. *Convex analysis*. Princeton Univ. Press, Princeton, NJ.
- Rockafellar, R.T., S. Uryasev. 2000. Optimization of conditional value-at-risk. *Journal of Risk* **2** 21–41.