DISCUSSION PAPERS IN ECONOMETRICS AND STATISTICS

INSTITUTE OF ECONOMETRICS AND STATISTICS

No. 01/14

Geometrical Framework for Robust Portfolio Optimization

by

Pavel Bazovkin

UNIVERSITY OF COLOGNE

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

DISCUSSION PAPERS IN ECONOMETRICS AND STATISTICS

INSTITUTE OF ECONOMETRICS AND STATISTICS

No. 01/14

Geometrical Framework for Robust Portfolio Optimization

by

Pavel Bazovkin[1]

Abstract

We consider a vector-valued multivariate risk measure that depends on the user's profile given by the user's utility. It is constructed on the basis of weighted-mean trimmed regions and represents the solution of an optimization problem. The key feature of this measure is convexity. We apply the measure to the portfolio selection problem, employing different measures of performance as objective functions in a common geometrical framework.

Keywords: Multivariate risk measure, robust portfolio optimization, weighted-mean trimmed regions, data central regions, convex risk measure, distortion risk measure.

AMS Subject Classification: 62P05, 91G10, 52A20, 91B30, 52A41, 62C20.

[1]Institute of Econometrics and Statistics, University of Cologne, D-50923 Köln, Germany

Geometrical Framework for Robust Portfolio Optimization

Pavel Bazovkin[1]

[1]Institute of Econometrics and Statistics, University of Cologne,
Meister-Ekkehart-Str. 9, D-50923 Cologne, Germany,
bazovkin@wiso.uni-koeln.de

September 12, 2014

Abstract: We consider a vector-valued multivariate risk measure that depends on the user's profile given by the user's utility. It is constructed on the basis of weighted-mean trimmed regions and represents the solution of an optimization problem. The key feature of this measure is convexity. We apply the measure to the portfolio selection problem, employing different measures of performance as objective functions in a common geometrical framework.

Keywords: Multivariate risk measure, robust portfolio optimization, weighted-mean trimmed regions, data central regions, convex risk measure, distortion risk measure.

1 Introduction

Quantifying risk is one of the most important problems in modern economics. Classical tools of mathematical finance include *risk measures.* These functions, as their name states, assess the risk of some financial positions, which are traditionally modelled by some random vector X. The basic idea of a risk measure is to indicate a critical value of a (monetary) deposit, or reserve, that, being added to an uncertain position, does cancel its risk in some sense.

The latter means that the location of the distribution of the corresponding random vector satisfies certain formal requirements that are provided by, say, a regulator.

For example, if X is univariate, he or she may require that some α-quantile of the distribution be non-negative. If we add a constant $-Q_X(\alpha)$, which can be interpreted as an insurance deposit, to the distribution, where Q_X denotes the quantile function of X, we make the condition hold. In other words, only the worst $\alpha \cdot 100\%$ of outcomes of the insured position are expected to be negative. In such a manner we get a famous and widely used (cf. Jorion, 2006) risk measure called *value-at-risk* (V@R). Actually, there is a plenty of different notions of risk measures, each controlling particular aspects of the outcome distribution. There is also a list of desired properties of such functions: to some of them we will refer below. As further examples of univariate risk measures one can recall the *expected shortfall*[1], *expected minimum*, textitentropic risk measure and others.

A univariate risk measure concerns an investment into one asset. However, in practice a user is usually operating with several different assets. In this case, measuring the risk becomes a much more complicated problem than just doing it for each asset individually. The issue lies in a dependence between the assets, which can be rather complex and lead to an asymmetric joint distribution of the assets' returns.

The higher dimension is, namely the number of assets, the more importance has to be given to the dependency information. This is similar to modelling returns by a d-dimensional random vector X instead of d separate univariate random variables.

At this point, immediately we get an issue: again, the risk of X could not be comprehensively described only by risks of its marginals. To tackle this problem, a rather natural idea has been proposed. If a univariate monetary risk measure describes the minimal deterministic amount of money that, being added to the investment, compensates its risk, one could do the same in the multivariate case. In other words, to find such 'minimal' deterministic vectors in $\mathbb{R}^d$ that compensate the risk of X. To get rid of the ambiguous 'minimal' qualifier, we can just take *all* deterministic vectors compensating the risk. It is easy to see that these vectors form a set in $\mathbb{R}^d$, of an affine

[1]It is also called the *average value-at-risk* or the *tail value-at-risk* in the literature.

dimension d in general. Obviously, 'minimal' vectors are lying on the surface of this d-dimensional body. In general, if there are transaction costs, there are many incomparable 'minimal' vectors.

The above gives way to *set-valued* risk measures, which are nowadays rather common in considering multivariate risks.

Working with such measures is a rather complex task, however, it becomes simpler if all such sets, that is to say, values of a set-valued risk measure, are convex. Further on, we consider this property as an advantage.

The development of coherent risk measures (Artzner et al., 1999; Delbaen, 2002) and of the pertaining machinery (see, e.g., Föllmer and Schied (2004)) as well as tightening of economic standards have led to considering multivariate risks and extending the notion of the risk measure as a real-valued function to a class of set-valued functions (Jouini et al., 2004). Recently, the corresponding theory has been deeply developed both generally (see, e.g., Hamel and Heyde (2010); Hamel et al. (2011); Rüschendorf (2013)) and concerning specific exemplars of risk measures (e.g., Cousin and Di Bernardino (2013); Hamel et al. (2013)).

Such literature proposes several ways of defining set-valued risk measures. In this paper, we pursue the approach of Cascos and Molchanov (2007), who explore a direct connection of such measures to data central regions. This gives us an advantage of applying geometrical algorithms for these regions to calculating set-valued risk measures. In fact, computability of set-valued risk measures is usually a hard issue (cf. Hamel et al. (2013, 2014)).

The investigation of multivariate risk measures develops in several *major* tasks. The first one, the representation, is connected with a discussion of a set of *desirable properties* for a risk measure, which the reader can find, e.g., in Rachev et al. (2008). A specialized analysis of *comonotonic* risk measures is given by Ekeland et al. (2012). A widely used *dual representation* of risk measures via acceptance sets is comprehensively described, for example, in Hamel and Heyde (2010). The second task, again, computability, is a very recent one and concerns mostly applying methods from vector optimization, such as Benson's algorithm (cf. Schrage and Löhne (2013), Hamel et al. (2014)). In our research, we concentrate on the computability via efficient *geometrical representations*. Besides this, we the measures can be applied not only in the sphere of finance but also in completely different ones (see, e.g., Bazovkin and Mosler (2014)).

2 Vector-valued multivariate risk measure based on data trimmed regions

In this section, we define a measure combining the objective evaluation of the risk by means of a set-valued risk measure, and the subjective preferences of the user, which are modelled by the user's admissable set.

2.1 The measure

According to Cascos (2009), using the ideas from Cascos and Molchanov (2007), a risk measure μ^d based on some data trimmed region D^*_α can be defined as follows:

$$\mu^d(X) = -\Big(D^*_\alpha(X) \oplus \mathbb{R}^d_+\Big),$$

meaning a reflection of the set $D^*_\alpha(X) \oplus \mathbb{R}^d_+$. In simple words, it states that for all $\mathbf{z} \in \mu^d$ the trimmed region $D^*_\alpha(X + \mathbf{z})$ does not lie in the positive orthant. For example, if D^*_α is a *halfspace region*, we obtain a *multivariate quantile*, which enables us to get a set-valued generalization of the *value-at-risk*. The *subadditivity* property and the analytical simplicity of *zonoid regions* enable us to use them for generalization of the *expected shortfall*, which is a *coherent* risk measure. In turn, the *expected minimum*, also a coherent measure, is generalized by means of *expected convex hull regions*.

In the same manner, we define a special class of multivariate risk measures based on weighted-mean trimmed regions $D_{\mathbf{w}_\alpha}$ given by a weight vector $\mathbf{w}_\alpha$.

Definition 1. *The multivariate set-valued distortion risk measure is defined as follows:*

$$\mu^d(X) = -\Big(D_{\mathbf{w}_\alpha}(X) \oplus \mathbb{R}^d_+\Big) \subset \mathbb{R}^d. \tag{1}$$

A detailed consideration of distortion risk measures the reader can find, for instance, in Mosler and Bazovkin (2014). In this paper, we are only interested in a measure with desirable properties, such as the subadditivity, which encourages diversification and is crucial in risk management.

We should mention that this is not a unique way of defining a multivariate distortion risk measure. For comparison, Rüschendorf (2013) gives a different notion of such a measure, which is scalar-valued: For a d-variate distribution having p.d.f. F, he considers the level set $Q(t)$ of F at level t and defines some scalar measure of $Q(t)$ as the t-quantile. Then, based on these scalar-valued quantiles, he introduces multivariate risk measures in the same way as univariate ones. Obviously, much information is lost in this case and the choice of the scalar measure is not straightforward.

To flexibilize our definition by incorporating the information about the user's preferences, described by his or her utility function $U(\cdot)$, we introduce the *admissable set* $\mathcal{F}$. This set collects all such returns that are perceived positively by the user. To relate it to the utility function, we assume $\mathcal{F} = \{\mathbf{y} \in \mathbb{R}^d : U(\mathbf{y}) \geq u_0\}$. Thus, the surface of $\mathcal{F}$ is the u_0-level set of the utility function.

We take the natural assumption of the user's risk aversion, which is equivalent to possessing a convex admissable set $\mathcal{F}$ (see, e.g., Föllmer and Schied (2004)). As an approximation, we suppose $\mathcal{F}$ to have the following form:

$$\mathcal{F} = \{\mathbf{y} \in \mathbb{R}^d : \mathbf{p}_k'\mathbf{y} \geq \delta_k, \quad k = 1 \dots K\} \tag{2}$$

with some $\mathbf{p}_1, \dots, \mathbf{p}_K \in \mathbb{R}^d_+$ and $\delta_1, \dots, \delta_K \in \mathbb{R}$, that is, $\mathcal{F}$ is an *upper convex polytope.*

E.g., a market with proportional transaction costs $\mathcal{F}$ is a cone with the apex at $\mathbf{0}$. Each level set of $U(\cdot)$ is the same (but translated) cone.

Our idea lies in a comparison of the position of the set-valued measure μ^d with that of the admissable set $\mathcal{F}$.

Definition 2. $\nu(X)$, *a real-valued risk measure of a risky position X given the user's utility $U(\cdot)$, is defined as follows:*

$$\nu(X) = \underset{\mathbf{z} \in \mathbb{R}^d}{\arg\min} \|z\|_U : \{-\mu^d(X) + \mathbf{z}\} \subset \mathcal{F}, \tag{3}$$

where $\|\cdot\|_U$ denotes a proper norm.

In other words, $\nu(X)$ is (in the sense of the norm $\|\cdot\|_U$) the shortest vector $\mathbf{z}$ that brings the set-valued measure $\rho(X)$ into the admissable set $\mathcal{F}$. The conventional Euclidean norm $\|\cdot\|_2$ is a natural choice for $\|\cdot\|_U$, however, a weighting of dimensions is possible due to their different importance in the user's subjective perception. If this mutual weighting is described by some positive definite matrix Γ_U, then for any $\mathbf{z} \in \mathbb{R}^d$ it holds $\|\mathbf{z}\|_U = \|\Gamma_U \mathbf{z}\|_2$.

$\nu(\cdot)$ enjoys a clear interpretation as a *monetary* measure: The minimal reserve to be added to the position to make it acceptable. While the measure states on the optimal decision of the user, we will call $\nu(\cdot)$ a *best-decision risk measure.*

The transition from the set-valued measure $\mu^d(\cdot)$ to the vector-valued $\nu(\cdot)$ is realized by solving an *optimization problem.* In fact, what we are doing is a specific scalarization of a set-valued risk measure (cf. Hamel and Heyde (2010), or Schrage (2012)). Our approach consists in the most broad employment of the user profile information (given by the utility function $U(\cdot)$) in doing this.

It is easy to see that (3) in Definition 2 is equivalent to the following:

$$\nu(X) = \underset{\mathbf{z} \in \mathbb{R}^d}{\arg\min} \|z\|_U : \{D_{\mathbf{w}_\alpha}(X) + \mathbf{z}\} \subset \mathcal{F}. \tag{4}$$

Finally, we like mention that using the measure $\nu(\cdot)$, we can define an order on a set of appropriate risky positions $\mathcal{X}$.

Definition 3 (Ordering risks)**.** *The preference relation $\succcurlyeq_\nu$ on $\mathcal{X}$ is given as follows:*

$$\forall Y, Z \in \mathcal{X} \qquad Y \succcurlyeq_\nu Z \qquad \iff \qquad \|\nu(Z)\|_U \geq \|\nu(Y)\|_U.$$

3 Portfolio choice as a special case

A portfolio choice problem can be stated using risk measures. Unlike standard portfolio theory, where variances are used as proxies for risk, the risk measures machinery allows to treat risk more comprehensively. It is worth to mention that the disadvantage of representing risk by the variance has become a vital issue in the literature of the last decade. Besides this, two-stage mean-variance procedures, where on the first stage parameters of a model should be estimated, such as covariance matrix of random returns, are subject to estimation risk (cf. Meucci, 2009). Authors from various mathematical fields propose approaches for solving the problem. For instance, Fabozzi et al. (2010) give a detailed review of robust methods emerged in portfolio optimization and the corresponding literature. These methods are usually based on modelling uncertainty either in parameters (e.g., Tütüncü and Koenig (2004); Costa and Paiva (2002); El Ghaoui et al. (2003)) or in the whole distribution (e.g., Calafiore, 2007) and appropriate modifications of the variance. A part of recent approaches consider risk measures (e.g., Rockafellar et al. (2006); Bion-Nadal and Kervarec (2012); Drapeau and Kupper (2013)). A qualitatively new algorithm, which efficiently combines robust optimization with coherent risk measures, contributing to this trend, has been proposed by Mosler and Bazovkin (2014). In this paper, we solve the portfolio choice problem using optimization of either the multivariate risk measure $\nu(\cdot)$ or some *performance measure*. In our approach, we get rid of usual distributional assumptions on returns, namely their *ellipticity*.

Let $\tilde{r}_1, \ldots, \tilde{r}_d$ be random return rates on d assets. We will notate $\tilde{\mathbf{r}} = (\tilde{r}_1, \ldots, \tilde{r}_d)'$. A convex combination of the assets' returns is sought, $\tilde{\mathbf{r}}'\boldsymbol{\omega} = \sum_{j=1}^{d} \tilde{r}_j \omega_j$, that maximizes some performance measure. Let us have a portfolio of d assets and the historical information about its returns $\{\mathbf{r}^1, \ldots, \mathbf{r}^n\} \subset \mathbb{R}^d$. Now we can consider a task of finding a portfolio with the lowest risk possible or a portfolio optimized by means of some generalized performance measure, for example, a Sharpe ratio.

To solve the task, we use the multivariate measure $\nu(\cdot)$ given by Definition 2 in the previous section. The considered problem has a certain form of the admissable set: a halfspace, that is, a special case of (2). The border of the halfspace, a hyperplane, is determined by a portfolio vector, or simply a *portfolio*, $\boldsymbol{\omega} \in \Delta^d = \{\boldsymbol{\delta} \in \mathbb{R}^d : \boldsymbol{\delta} \geq \mathbf{0},\ \mathbf{1}'\boldsymbol{\delta} = 1\}$. This fact enables us to control the admissable set by means of varying $\boldsymbol{\omega}$ and find one that produces the minimal risk in such a way. Thus we obtain a parametric optimization problem in the sense of optimizing the risk measure $\nu(\cdot)$ or some function dependent on it. In the following subsection we propose an efficient geometric procedure of finding the optimal $\boldsymbol{\omega}^{\text{opt}}$ in the space $\mathbb{R}^d$.

3.1 Minimal risk portfolio

Further on, a $d \times d$ matrix $\mathbf{\Omega}$ denotes diag($\boldsymbol{\omega}$). We are minimizing the risk of a portfolio, that is, are employing the following criterion $g(\boldsymbol{\omega})$:

$$g(\boldsymbol{\omega}) = \|\nu(\mathbf{\Omega}\tilde{\mathbf{r}})\|_U \to \min_{\boldsymbol{\omega}\in\Delta^d} . \tag{5}$$

Note that later the restriction to $\boldsymbol{\omega} \in \Delta^d$ will be relaxed concerning the non-negativity of components.

For a specified distortion risk measure, namely a given weight vector $\mathbf{w}_\alpha$, and an empirical sample $\mathbf{r}^1, \ldots, \mathbf{r}^n$, we construct a trimmed region $D_{\mathbf{w}_\alpha}(\mathbf{r}^1, \ldots, \mathbf{r}^n)$.

Taking the Euclidean norm as $\|\cdot\|_U$, the value of the objective $g(\boldsymbol{\omega})$ for some $\boldsymbol{\omega}$ is the Euclidean length of the minimal shift $\mathbf{s}_{\boldsymbol{\omega}} \in \mathbb{R}^d$ of the admissable set $\mathcal{F} = \{\mathbf{x} : \mathbf{1}'\mathbf{x} \geq 0\}$ such that for the data weighted by $\boldsymbol{\omega}$ it holds:

$$D_{\mathbf{w}_\alpha}(\mathbf{\Omega}\mathbf{r}^1, \ldots, \mathbf{\Omega}\mathbf{r}^n) \subset \mathcal{F} - \mathbf{s}_{\boldsymbol{\omega}}. \tag{6}$$

Obviously, $\mathbf{s}_{\boldsymbol{\omega}} = \nu(\mathbf{\Omega}\tilde{\mathbf{r}})$, where $\tilde{\mathbf{r}}$ is empirically distributed on $\mathbf{r}^1, \ldots, \mathbf{r}^n$. For convenience, we will denote $\mathcal{F} - \mathbf{s}_{\boldsymbol{\omega}}$ by $\widehat{\mathcal{F}}_{\boldsymbol{\omega}}$.

Let us now do an inverse transform of the space, that is, a linear transform by $\mathbf{\Omega}^{-1}$. Then, we get $\mathbf{\Omega}^{-1}\mathcal{F}$ instead of $\mathcal{F}$ and, respectively, the condition (6) becomes equivalent to the following:

$$D_{\mathbf{w}_\alpha}(\mathbf{r}^1, \ldots, \mathbf{r}^n) \subset \mathbf{\Omega}^{-1}\widehat{\mathcal{F}}_{\boldsymbol{\omega}}. \tag{7}$$

Because of the budget constraint $\mathbf{1}'\boldsymbol{\omega} = 1$, it is easy to show that the harmonic mean of axes intersections with the hyperplane $\partial\widehat{\mathcal{F}}_{\boldsymbol{\omega}}$ does not change after getting to $\partial\{\mathbf{\Omega}^{-1}\widehat{\mathcal{F}}_{\boldsymbol{\omega}}\}$. It equals $g(\boldsymbol{\omega})\sqrt{d}$, where $\partial\{\cdot\}$ denotes the border of a set. Now, let some $\boldsymbol{\omega}^1, \boldsymbol{\omega}^2$ produce the same objective values $g(\boldsymbol{\omega}^1) = g(\boldsymbol{\omega}^2) = g$ and form the borders $\partial\{\mathbf{\Omega}_1^{-1}\widehat{\mathcal{F}}_{\boldsymbol{\omega}^1}\}$ and $\partial\{\mathbf{\Omega}_2^{-1}\widehat{\mathcal{F}}_{\boldsymbol{\omega}^2}\}$, respectively. It can be shown that these borders intersect at the point $-\mathbf{s}_{\boldsymbol{\omega}^1} = -\mathbf{s}_{\boldsymbol{\omega}^2} = (-\frac{g}{\sqrt{d}}, \ldots, -\frac{g}{\sqrt{d}})'$. This point delimits the interval $(\mathbf{0}; -s_{\boldsymbol{\omega}^1})$ on the bisector, which has length g.

Thus, we see that there is a bijection between all plausible $\boldsymbol{\omega}$-s and $\mathbf{\Omega}^{-1}\widehat{\mathcal{F}}_{\boldsymbol{\omega}}$. Moreover, $g(\boldsymbol{\omega})$ is the length of the interval cut off by the surface $\partial\{\mathbf{\Omega}^{-1}\widehat{\mathcal{F}}_{\boldsymbol{\omega}}\}$ on the bisector. Hence we have to find a hyperplane $\partial\{\mathbf{\Omega}^{-1}\widehat{\mathcal{F}}_{\boldsymbol{\omega}^{\mathrm{opt}}}\}$ that 'covers' $D_{\mathbf{w}_\alpha}$ and cuts the shortest interval on the bisector. It is easy to show

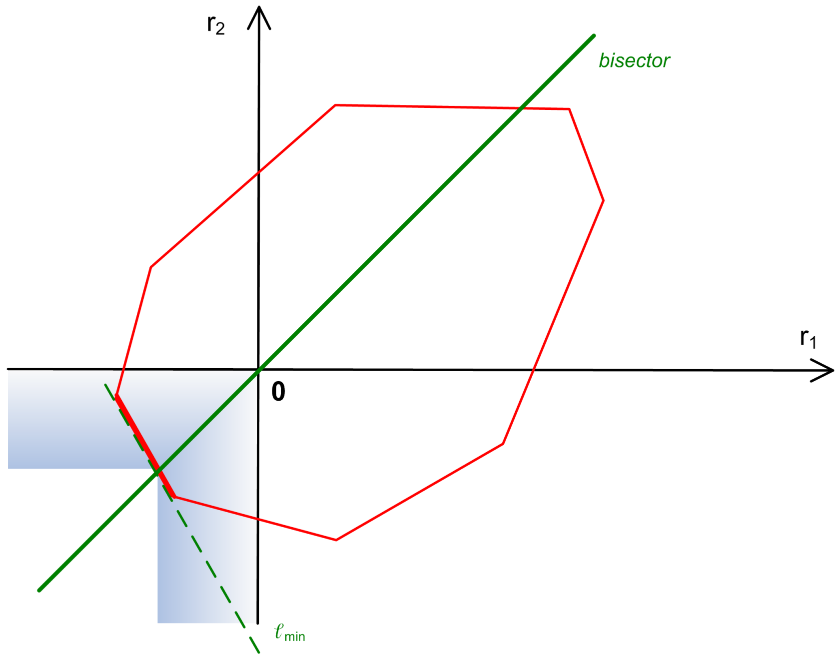

Figure 1: Searching the minimal risk portfolio.

that it is a hyperplane $\ell_{\min}$ containing the facet intersected by the bisector (see Figure 1). Hence the solution is the following: the sought-for $\boldsymbol{\omega}^{\text{opt}}$ is the normalized (to the component sum of 1) normal to the facet intersected by the bisector.

3.2 Portfolio selection with a generalized Sharpe ratio

Now we solve the problem of maximizing the ratio of expected returns to the risk taken. It stands on the same principle as the well-known Sharpe ratio (Sharpe, 1966) and will be denoted by $\mathsf{SR}_{\tilde{\mathbf{r}}}$:

$$\mathsf{SR}_{\tilde{\mathbf{r}}}(\boldsymbol{\omega}) = \frac{\boldsymbol{\omega}' \cdot \boldsymbol{\mu}(\tilde{\mathbf{r}})}{\|\nu(\boldsymbol{\Omega}\tilde{\mathbf{r}})\|_U}\,, \tag{8}$$

where $\boldsymbol{\mu}(\tilde{\mathbf{r}})$, or simply $\boldsymbol{\mu}$, is the expected return $E(\tilde{\mathbf{r}})$ of the investment.

3.2.1 Finding the optimum

Under the standard restriction on $\boldsymbol{\omega}$, $\boldsymbol{\omega} \in \Delta^d$, we have to solve the following optimization task:

$$g(\boldsymbol{\omega}) = \mathsf{SR}_{\tilde{\mathbf{r}}}(\boldsymbol{\omega}) \to \max_{\boldsymbol{\omega} \in \Delta^d} . \tag{9}$$

In the inverse transformed space, a hyperplane parallel to $\partial\{\boldsymbol{\Omega}^{-1}\widehat{\mathcal{F}}_{\boldsymbol{\omega}}\}$ is a set of same-return outputs $\mathbf{x}$. The value of this return equals the length of the origin-started segment of the bisector cut off by the hyperplane, because this segment is not effected by the transformation. Hence, the expected return of a portfolio $\boldsymbol{\omega}$ is equal to the length of a segment cut off by such a hyperplane containing $\boldsymbol{\mu}$. Consider Figure 2: for a case of minimal risk (see Subsection 3.1), this segment corresponds to $\mathbf{0}\mathrm{E}'$ ($\boldsymbol{\mu} \in \ell'_{\min}$, $\ell'_{\min} \| \ell_{\min}$, where '$\|$' means that $\ell'_{\min}$ and $\ell_{\min}$ are parallel). Let us now draw a line through the points $\boldsymbol{\mu}$ and $\mathbf{0}$ and find its intersection with the hyperplane $\ell_{\min}$ (the solution hyperplane for the minimal risk problem, a green dashed line on the Figure 2) - the point A. $\ell_{\min}$ cuts off the segment $\mathbf{0}\mathrm{E}$ with length equal to the risk estimate. $\ell'_{\min}$ cuts off the segment $\mathbf{0}\mathrm{E}'$ with length equal to the expected return. Thus, we obtain $\mathsf{SR}_{\tilde{\mathbf{r}}}(\boldsymbol{\omega}) = \frac{\mathbf{0}\mathrm{E}'}{\mathbf{0}\mathrm{E}}$.[2]

Let us now rotate $\ell_{\min}$ in $\mathbb{R}^d$ arbitrarily around the point A to some position ℓ. Of course, ℓ must not intersect $D_{\mathbf{w}_\alpha}$. $\ell'_{\min}$ is rotated parallelly around the point $\boldsymbol{\mu}$ to some hyperplane $\ell' \| \ell$. The rotation corresponds to browsing through different portfolios $\boldsymbol{\omega}$.

Thus, the points E and E$'$ move: $\mathrm{E} \mapsto \mathrm{B}$, $\mathrm{E}' \mapsto \mathrm{B}'$. We immediately get $\triangle \mathbf{0}\mathrm{EA} \sim \triangle \mathbf{0}\mathrm{E}'\boldsymbol{\mu}$ and $\triangle \mathbf{0}\mathrm{BA} \sim \triangle \mathbf{0}\mathrm{B}'\boldsymbol{\mu}$, where under '$\sim$' we understand the similarity relationship. Hence $\mathsf{SR}_{\tilde{\mathbf{r}}}(\boldsymbol{\omega}') = \frac{\mathbf{0}\mathrm{E}'}{\mathbf{0}\mathrm{E}} = \frac{\mathbf{0}\boldsymbol{\mu}}{\mathbf{0}\mathrm{A}} = \frac{\mathbf{0}\mathrm{B}'}{\mathbf{0}\mathrm{B}} = \mathrm{const}$. For each rotation $\boldsymbol{\omega}$ there is a hyperplane $\ell_0 \| \ell$ which touches $D_{\mathbf{w}_\alpha}$ and intersects the bisector at D, that is, gives the actual estimation $\|\mathbf{0}\mathrm{D}\|_U$ of risk of the portfolio $\boldsymbol{\omega}$.

To maximize $\mathsf{SR}_{\tilde{\mathbf{r}}}(\boldsymbol{\omega}) = \mathsf{SR}_{\tilde{\mathbf{r}}}(\boldsymbol{\omega}') \cdot \frac{\mathbf{0}\mathrm{B}}{\mathbf{0}\mathrm{D}}$, we should maximize $\frac{\mathbf{0}\mathrm{B}}{\mathbf{0}\mathrm{D}}$ by a rotation. Let C be an intersection of the line $(\mathbf{0}, \boldsymbol{\mu})$ with the hyperplane ℓ_0. Then it holds $\triangle \mathbf{0}\mathrm{DC} \sim \triangle \mathbf{0}\mathrm{BA}$, leading to $\frac{\mathbf{0}\mathrm{B}}{\mathbf{0}\mathrm{D}} = \frac{\mathbf{0}\mathrm{A}}{\mathbf{0}\mathrm{C}}$. At the same time, $\mathbf{0}\mathrm{A}$ remains constant, which means that we should just minimize $\mathbf{0}\mathrm{C}$. It is easy to see that the shortest possible $\mathbf{0}\mathrm{C}$ is the interval with the point C lying on the

[2]Further in this paper the name of a segment in a formula implies the length of the segment.

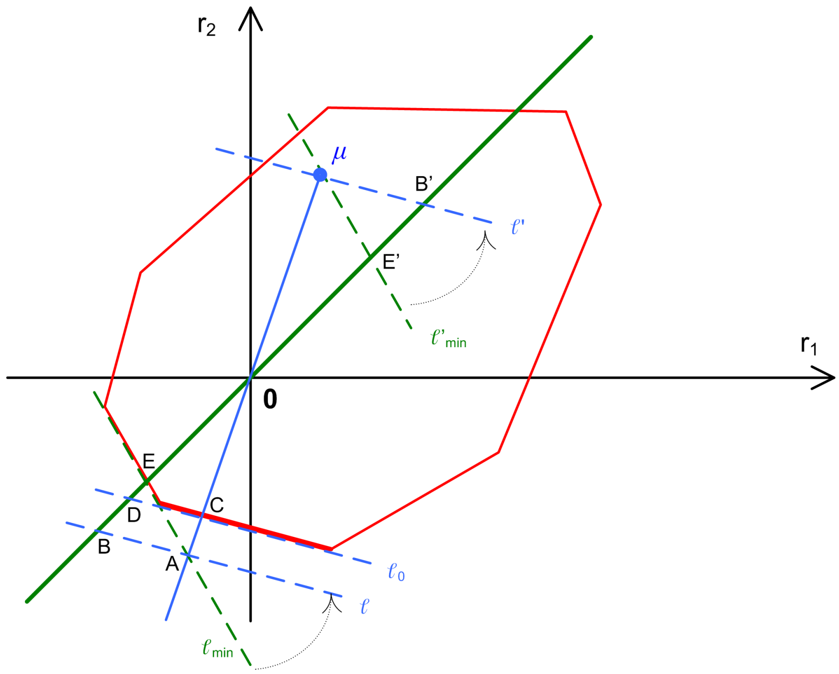

Figure 2: Searching the Sharpe ratio optimized portfolio.

border of $D_{\mathbf{w}_\alpha}$. In turn, it means that the sought-for optimal hyperplane ℓ_0^{opt} is that containing the facet intersected by the line $(\mathbf{0}, \boldsymbol{\mu})$.

Hence, the sought-for *solution* $\boldsymbol{\omega}^{\text{opt}}$ is a normalized (to the component sum of 1) normal to the facet of the lower boundary of $D_{\mathbf{w}_\alpha}$ intersected by the line $(\mathbf{0}, \boldsymbol{\mu})$.

It is easy to check, that the above considerations hold for all $d \geq 2$, although being illustrated in $\mathbb{R}^2$.

The reader may make the following observations, which are quite important:

1. If all assets yield similar expected returns, the procedure calculates the minimal risk portfolio (because $\boldsymbol{\mu}$ lies on the bisector), which is intuitively natural. In this case, the procedure degenerates to one from Subsection 3.1.

2. The procedure can be enlarged to the case of data following a general probability distribution. The solution will be the similarly normalized

vector tangent to the lower surface of $D_{\mathbf{w}_\alpha}$ at the point of its intersection with the line $(\mathbf{0}, \boldsymbol{\mu})$.

3. $\boldsymbol{\mu}$ can be replaced, for example, by a median or a shrinkage location estimator (cf. Meucci (2009)).

If we have some risk-free asset with the risk-free rate r_f, we put a point $\mathbf{u}_0$ on the bisector so that the length of the interval $[\mathbf{0}, \mathbf{u}_0]$ equals r_f. Then, it is easy to show that one should apply the same procedure as above, just replacing the line $(\mathbf{0}, \boldsymbol{\mu})$ by $(\mathbf{u}_0, \boldsymbol{\mu})$, likewise changing the focus of the intersecting ray.

At this step, it is interesting to observe how the procedure works in the special case of elliptically distributed returns with some covariance matrix $\boldsymbol{\Sigma}$. While in this case WM regions asymptotically converge (see Dyckerhoff and Mosler (2012); Mosler (2002)) to ellipsoids with the shape matrix $\boldsymbol{\Sigma}$ for any choice of α and type of the region, it can be easily shown that the solution will, in turn, converge to the *tangential portfolio*:

$$\boldsymbol{\omega}^{\text{opt}} = \frac{\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}}{\mathbf{1}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}}. \tag{10}$$

Really, the normal to the tangent hyperplane at the ellipsoid's point intersected by the line of direction $\boldsymbol{\mu}$ is $\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$. Having normalized the vector, we get the above formula (10).

It is immediately seen that replacing $\boldsymbol{\mu}$ with $\mathbf{1}$ above gives the portolio $\frac{\boldsymbol{\Sigma}^{-1}\mathbf{1}}{\mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1}}$, which, in turn, is the minimal variance portfolio. It means that the latter is defined in a standard way by the intersection of a line parallel to the bisector and passing through $\boldsymbol{\mu}$ with the trimmed region.

These facts demonstrate that our approach is a generalization of a typical mean-variance procedure, where standard distributional assumptions are avoided and a comprehensive non-parametric risk measure is employed.

3.2.2 The algorithm

Input

- $\{\mathbf{r}^1, \dots, \mathbf{r}^n\} \subset \mathbb{R}^d$ - the given empirical data about returns.

- Risk parameter α.
- The type of distortion risk measure μ^d to be used.
- *Optionally:* The risk-free rate r_f.

Output

- The optimal portfolio $\boldsymbol{\omega}^{\text{opt}}$.
- Value of the criterion.

Steps (SR-algorithm)

SR1. Define the weight vector $\mathbf{w}_\alpha$. Construct a focus point $\mathbf{u}_0 = (\frac{r_f}{\sqrt{d}}, \dots, \frac{r_f}{\sqrt{d}})'$ or take, by default, the origin $\mathbf{0}$. Construct a line $\varphi = (\mathbf{u}_0, \boldsymbol{\mu})$ or $\varphi = (\mathbf{0}, \boldsymbol{\mu})$ respectively.

SR2. Calculate (Bazovkin and Mosler, 2012) a part of $D_{\mathbf{w}_\alpha}(\mathbf{r}^1, \dots, \mathbf{r}^n)$ in the place of a probable intersection with φ (cf. the *efficient set* in Mosler and Bazovkin (2014)). The type of $D_{\mathbf{w}_\alpha}$ corresponds to the selected distortion risk measure.

SR3. Find the facet of $D_{\mathbf{w}_\alpha}$ intersected by φ. Get the normal $\vec{n}_{\text{opt}}$ to the hyperplane containing it. The sought-for $\boldsymbol{\omega}^{\text{opt}} = \frac{\vec{n}_{\text{opt}}}{\mathbf{1}'\vec{n}_{\text{opt}}}$.

A special consideration is needed for a case when there is some negative component in $\boldsymbol{\omega}^{\text{opt}}$, namely $\exists i : \omega_i^{\text{opt}} < 0$. If it occurs, one sets $\omega_i^{\text{opt}} = 0$ and solves the task without the i-th asset (namely projecting onto $\mathbb{R}^{d-1}$). However, this situation can be managed more flexibly, which is the topic of Subsection 3.4 below.

The intersected facet from Step SR3. of the algorithm can be realized as follows:

A. Construct the first facet of $D_{\mathbf{w}_\alpha}(\mathbf{r}^1, \dots, \mathbf{r}^n)$ with the normal close to the direction of φ.

B. Find the neighboring facet with the best criterion (Mosler and Bazovkin, 2014) describing its distance from φ.

C. Jump to the facet found and go to step B..

It can be seen that on each step of this subalgorithm we get a better solution. Furthermore, the tactics of the *"long jump"* can be used, where a jump over some neighbors in a criterion-enhancing direction is made at one step.

The main complexity-contributing issues are the following:

1. Calculating some facets of the trimmed region $D_{\mathbf{w}_\alpha}$: much simpler than calculating the whole region (since knowing φ).

2. Finding an intersection of a line with a convex surface in $\mathbb{R}^d$.

3.3 Optimization with a generalized certainty equivalent

In this subsection we pursue the same optimization problem but with a performance measure given by the certainty equivalent, which is commonly used in modern portfolio theory (cf. Markowitz (1952)). Again, the difference is that we replace the variance by the risk measure $\nu(\cdot)$. Then the criterion is the following:

$$\mathsf{CE}_{\tilde{\mathbf{r}}}(\boldsymbol{\omega}) = \boldsymbol{\omega}'\boldsymbol{\mu} - \lambda \cdot \|\nu(\boldsymbol{\Omega}\tilde{\mathbf{r}})\|_U \,, \tag{11}$$

where λ is a given positive constant describing the risk aversion of the user.

3.3.1 Finding the optimum

We will maximize $\frac{1}{\lambda}\mathsf{CE}_{\tilde{\mathbf{r}}}$, namely:

$$g(\boldsymbol{\omega}) = \frac{1}{\lambda}\mathsf{CE}_{\tilde{\mathbf{r}}}(\boldsymbol{\omega}) \to \max_{\boldsymbol{\omega}\in\Delta^d} \,. \tag{12}$$

First, we create a point $\boldsymbol{\mu}_\lambda = -\frac{1}{\lambda}\boldsymbol{\mu}$. Now consider Figure 3. If we have a portfolio $\boldsymbol{\omega}^1$ given by the hyperplane ℓ_1, the corresponding risk $\|\nu(\boldsymbol{\Omega}X)\|_U$ equals the length of the segment $\mathrm{A}_1\mathbf{0}$. Analogously to the previous subsection, $\frac{1}{\lambda}\boldsymbol{\omega}^{1\prime}\boldsymbol{\mu}$ equals $\mathrm{B}_1\mathbf{0}$, where $\mathrm{B}_1 = \ell_1' \cap \{\text{bisector}\}$ and ℓ_1' is a hyperplane parallel to ℓ_1 and containing $\boldsymbol{\mu}_\lambda$. Now, it is directly seen that $\frac{1}{\lambda}\mathsf{CE}_{\tilde{\mathbf{r}}}$ equals

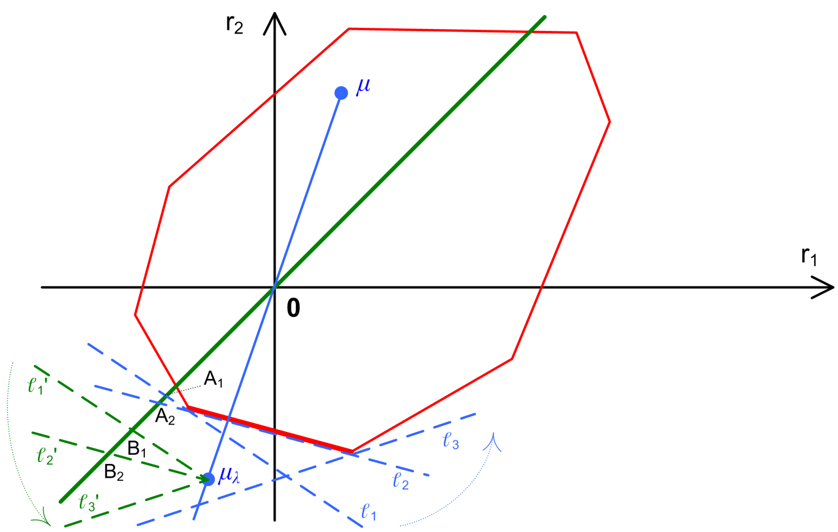

Figure 3: Searching the certainty equivalent optimized portfolio.

$\mathrm{A}_1\mathrm{B}_1 = \mathrm{B}_1\mathbf{0} - \mathrm{A}_1\mathbf{0}$. The same principle is applied to a portfolio $\boldsymbol{\omega}^2$, yielding the criterion value $\mathrm{A}_2\mathrm{B}_2$ for the latter.

We see that $\mathrm{A}_1\mathrm{B}_1 < \mathrm{A}_2\mathrm{B}_2$, because ℓ_1 rotates to the position ℓ_2 with a smaller shoulder relatively to (i.e. distance to) the bisector as ℓ_1' to ℓ_2'. That is to say, $\mathrm{A}_i\mathrm{B}_i$ increases while rotating ℓ_i if ℓ_i' has a larger shoulder relatively to the bisector and vice versa. Thus, starting from the minimal risk position, we first increase $\mathsf{CE}_{\tilde{\mathbf{r}}}$ until ℓ gets a pivot more distant from the bisector as $\boldsymbol{\mu}_\lambda$. It can happen when leaving a position containing a facet that, in turn, contains points equidistant with $\boldsymbol{\mu}_\lambda$ from the bisector. Hence the optimal hyperplane is one containing a facet intersected by a hypercylinder with the bisector as its axis and $\boldsymbol{\mu}_\lambda$ lying on its surface.

To construct an algorithm, we first find a facet intersected by a line parallel to the bisector and containing $\boldsymbol{\mu}_\lambda$. It is the first candidate. Then we move along the ring (intersection with the hypercylinder) and check the values of $\mathsf{CE}_{\tilde{\mathbf{r}}}$ for each of the facets. A facet ℓ_{j^*} with the maximum $\mathsf{CE}_{\tilde{\mathbf{r}}}$ defines the optimal portfolio.

Finally, consider a special case when $\lambda \to \infty$. Maximizing the criterion (11) becomes equivalent to optimizing $\nu(\boldsymbol{\Omega}\tilde{\mathbf{r}})$. Thus, we obtain the minimal risk problem. While $\boldsymbol{\mu}_\lambda \to \mathbf{0}$, the hypercylinder degenerates into a line. Hence the sought-for facet is the facet intersected by the bisector. Obviously, we get the same solution as in Subsection 3.1. Another extreme case occurs when λ is small enough, so that the hypercylinder contains $D_{\mathbf{w}_\alpha}$. In this case, we can rotate ℓ until it becomes parallel to the bisector ($\mathbf{1}'\vec{n} = 0$). Clearly, that from all such hyperplanes the optimum is given by the one that is most remote from $D_{\mathbf{w}_\alpha}$. This optimum is a vector that has a single positive component for the maximal expected return and others are negative. It means purchasing only the asset j with $\mu_j = \max\{\mu_1, \dots, \mu_d\}$, where $(\mu_1, \dots, \mu_d)\ equiv \boldsymbol{\mu}$.

3.3.2 The algorithm

Input

- $\{\mathbf{r}^1, \dots, \mathbf{r}^n\} \subset \mathbb{R}^d$ - the given empirical data about returns.
- Risk parameter α.
- The type of distortion risk measure μ^d to be used.
- The risk aversion constant λ.

Output

- The optimal portfolio $\boldsymbol{\omega}^{\text{opt}}$.

Steps (CE-algorithm)

CE1. Calculate the relevant part of the trimmed region $D_{\mathbf{w}_\alpha}(\mathbf{r}^1, \dots, \mathbf{r}^n)$.

CE2. Construct the point $\boldsymbol{\mu}_\lambda = -\frac{1}{\lambda}\boldsymbol{\mu}$.

CE3. Build a line parallel to the bisector and containing $\boldsymbol{\mu}_\lambda$. Find its intersection with $D_{\mathbf{w}_\alpha}$ similarly to the step SR3. of the SR-algorithm.

CE4. Calculate the criterion $\frac{1}{\lambda}\mathsf{CE}_{\tilde{\mathbf{r}}}$ for the current facet with the index j. It equals the length of the segment $\mathrm{A}_j\mathrm{B}_j$. If it is the best currently, store the facet.

CE5. Find appropriate neighboring facets for the newly stored facet. For each of them, go to the step CE4. If there is no new neighbors, *stop*.

   a. A neighbor is appropriate if it contains points equidistant with $\boldsymbol{\mu}_\lambda$ to the bisector, which means being intersected by the hypercylinder. In doing this, calculate the min and max distances of the facet's points to the bisector.

CE6. Get a normal $\vec{n}_{\text{opt}}$ to the current best facet. The sought-for $\boldsymbol{\omega}^{\text{opt}} = \frac{\vec{n}_{\text{opt}}}{\mathbf{1}'\vec{n}_{\text{opt}}}$.

A case of negative weights can be solved as proposed in Subsection 3.4. If negative weights are not allowed, we pursue them analogously to Subsection 3.2.2.

3.4 Negative weights and short sellings

It is well-known that an estimated negative value of ω_i for the i-th asset actually proposes to do a short selling of that asset. We can use such a strategy as an alternative to just fixing corresponding weights to 0 and solving the similarly stated subproblem for the remaining assets. The approach given in this subsection is common for both the SR-algorithm and the CE-algorithm.

3.4.1 Optimum with shorting permitted

First, we modify the derivation of $\boldsymbol{\omega}^{\text{opt}}$ from a found $\vec{n}_{\text{opt}}$ due to the relaxation of the restriction $\boldsymbol{\omega} \in \Delta^d$. Namely only the sum of component absolute values $\|\boldsymbol{\omega}\|_1$ is set to 1, resulting in $\boldsymbol{\omega}^{\text{opt}} = \frac{\vec{n}_{\text{opt}}}{\|\vec{n}_{\text{opt}}\|_1}$. Let the user possess stores of the d assets available for allocating at the rates of $S_1, \dots, S_d$ units. The idea is to solve the task recursively.

We start from all d assets and on each stage allocate a finite number of units Z_k and eliminate those assets whose store is fully exhausted on the current stage. This filtering implies setting weights to 0 for the 'bottleneck' assets on next stages. We solve the filtered task recursively until we get some stage T with an optimal solution without negative components, or there is nothing more to allocate. Now consider a stage k: let J_k be a set of indices corresponding to negative components of the optimal portfolio on this stage, $\boldsymbol{\omega}^{\text{opt}_k}$. We determine the volume of units to be allocated on this step:

$$Z_k = \min_{j \in J_k} \frac{S_j^k}{|\omega_j^{\text{opt}_k}|},$$

where S_j^k denotes an available store of the asset j at the beginning of the stage k.

The structure of the problem is typical for *dynamic programming*, and each stage is pursued optimally. It means that the recursive procedure yields the overall optimal solution.

As a result, we obtain a 'ladder' of allocated units (see Figure 4) $Z_1, \dots, Z_T$, which yields the optimal allocation after an aggregation. If we want to invest some V units, we find such K that $\sum_{i=1}^{K-1} Z_i \le V \le \sum_{i=1}^{K} Z_i$. Then invest[3] Z_i into $\boldsymbol{\omega}^{\text{opt}_k}$ for all $k = 1, \dots, K-1$. The rest of V we invest into $\boldsymbol{\omega}^{\text{opt}_K}$. For example, on Figure 4, for $V = V_1$ we have $K = 2$, while for $V = V_2$, K equals T.

This simple example shows that the optimal aggregate portfolio depends on V (without shorting permitted, it is independent).

3.4.2 The algorithmic supplement

Input

- An aggregate number of units V to be allocated.
- Available stores $S_1, \dots, S_d$ of the assets.
- Standard inputs for either SR- or CE-algorithm.

Output

- The optimal allocation $\{V_1, \dots, V_d\}$.

Steps (NW-supplement)

[3] Investing into a negatively weighted asset means shorting it.

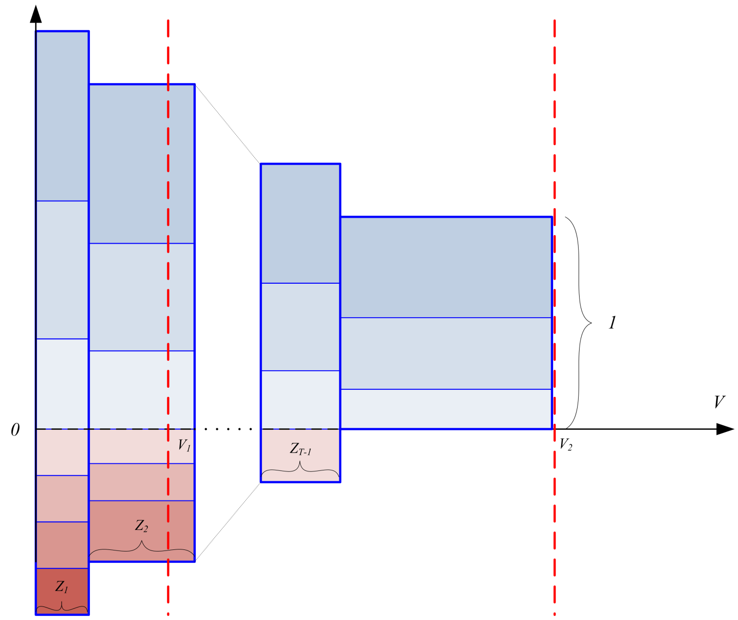

Figure 4: The recursive procedure for negative weights.

NW1. The first step (d assets, nothing invested): $k := 1$, $\hat{V} = 0, V_j = 0$ $\forall j = 1 \ldots d$.

NW2. Find the optimal portfolio $\boldsymbol{\omega}^{\text{opt}_k}$ by the SR- or CE-algorithm, while fixing weights of eliminated assets at zero.

NW3. $\boldsymbol{\omega}^{\text{opt}_k} := \frac{\boldsymbol{\omega}^{\text{opt}_k}}{\|\boldsymbol{\omega}^{\text{opt}_k}\|_1}$; $J_k = \{j : \omega_j^{\text{opt}_k} < 0\}$.

NW4. If $\forall j$ holds $\omega_j^{\text{opt}_k} \geq 0$, then $Z_k = V - \hat{V}$; else $Z_k = \min_{j \in J_k} \frac{S_i}{|\omega_j^{\text{opt}_k}|}$.

NW5. $Z_k := \min\{Z_k, V - \hat{V}\}$.

NW6. $V_j := V_j + Z_k \omega_j^{\text{opt}_k}, \forall j$.

NW7. $S_j := S_j + Z_k \omega_j^{\text{opt}_k}, \forall j$.

NW8. $\hat{V} := \hat{V} + Z_k$. If $\hat{V} = V$, go to Step NW10.

NW9. Eliminate assets with indices in J_k. $k := k + 1$. Go to Step NW2.

NW10. V_j is the final investment into the j-th asset. $V = \sum_j V_j$.

4 Discussion

In this paper we have shown a connection between set-valued distortion risk measures and weighted-mean trimmed regions. The former can be calculated using the algorithms for weighted-mean trimmed regions (Bazovkin and Mosler, 2012). We have considered the multivariate vector-valued risk measure $\nu(\cdot)$ that, firstly, aggregates the information from a set-valued coherent distortion risk measure and, at the same time, employs the user's risk posture information.

In a special case of substitutable components, we have applied the measure $\nu(\cdot)$ to solving a portfolio choice problem with different performance measures as objective functions. As a result, the efficient algorithms for the minimal risk, the generalized Sharpe ratio and the generalized certainty equivalent were proposed.

As a possible extension to be regarded, the shape of a trimmed region can be modified explicitly or via visual tools. The minimal risk and the SR-algorithm are realized in an R package PortfolioTR (Bazovkin, 2013). Besides this, the framework is flexible for incorporating further possible performance measures.

One more potential way of development of the framework lies in extending it to markets with transaction costs with admissable sets in form of convex cones or convex upper polytopes (2). An application of the risk measure $\nu(\cdot)$ for such situations is considered in Bazovkin and Mosler (2014).

Acknowledgments

The author is grateful to Karl Mosler and Oliver Grothe for deep discussions of the topics covered in this paper as a whole and in parts.

References

Artzner, P., Delbaen, F., Eber, J.-M., and Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, 9(3):203–228.

Bazovkin, P. (2013). *PortfolioTR: Portfolio Selection with Multivariate Distortion Risk Measures*. *R* package version 2.0, URL `http://CRAN.R-project.org/package=PortfolioTR`.

Bazovkin, P. and Mosler, K. (2012). An exact algorithm for weighted-mean trimmed regions in any dimension. *Journal of Statistical Software*, 47(13).

Bazovkin, P. and Mosler, K. (2014). A general solution for robust linear programs with distortion risk constraints. Submitted to Annals of Operations Research. Revised and Resubmitted.

Bion-Nadal, J. and Kervarec, M. (2012). Risk measuring under model uncertainty. *The Annals of Applied Probability*, 22(1):213–238.

Calafiore, G. C. (2007). Ambiguous risk measures and optimal robust portfolios. *Siam Journal on Optimization*, 18(3):853–877.

Cascos, I. (2009). Data depth: Multivariate statistics and geometry. In Kendall, W. and Molchanov, I., editors, *New Perspectives in Stochastic Geometry*. Clarendon Press, Oxford University Press, Oxford.

Cascos, I. and Molchanov, I. (2007). Multivariate risks and depth-trimmed regions. *Finance and Stochastics*, 11:373–397.

Costa, O. L. V. and Paiva, A. C. (2002). Robust portfolio selection using linear-matrix inequalities. *Journal of Economic Dynamics and Control*, 26(6):889–909.

Cousin, A. and Di Bernardino, E. (2013). On multivariate extensions of value-at-risk. *Journal of Multivariate Analysis*, 119:32–46.

Delbaen, F. (2002). Coherent risk measures on general probability spaces. In *Advances in Finance and Stochastics*, pages 1–37. Springer-Verlag, Berlin.

Drapeau, S. and Kupper, M. (2013). Risk preferences and their robust representation. *Mathematics of Operations Research*, 38(1):28–62.

Dyckerhoff, R. and Mosler, K. (2012). Weighted-mean trimming of a probability distribution. *Statistics and Probability Letters*, 82:318–325.

Ekeland, I., Galichon, A., and Henry, M. (2012). Comonotonic measures of multivariate risks. *Mathematical Finance*, 22(1):109–132.

El Ghaoui, L., Oks, M., and Oustry, F. (2003). Worst-case value-at-risk and robust portfolio optimization: A conic programming approach. *Operations Research*, 51(4):543–556.

Fabozzi, F. J., Huang, D., and Zhou, G. (2010). Robust portfolios: contributions from operations research and finance. *Annals of Operations Research*, 176:191–220.

Föllmer, H. and Schied, A. (2004). *Stochastic Finance: An Introduction in Discrete Time*. Walter de Gruyter, Berlin.

Frittelli, M. and Gianin, E. R. (2002). Putting order in risk measures. *Journal of Banking & Finance*, 26(7):1473–1486.

Hamel, A. and Heyde, F. (2010). Duality for set-valued measures of risk. *SIAM Journal on Financial Mathematics*, 1(1):66–95.

Hamel, A. H., Heyde, F., and Rudloff, B. (2011). Set-valued risk measures for conical market models. *Mathematics and Financial Economics*, 5(1):1–28.

Hamel, A. H., Löhne, A., and Rudloff, B. (2014). Benson type algorithms for linear vector optimization and applications. *Journal of Global Optimization*, 59(4):811–836.

Hamel, A. H., Rudloff, B., and Yankova, M. (2013). Set-valued average value at risk and its computation. *Mathematics and Financial Economics*, 7(2):229–246.

Holz, H. and Mosler, K. (1994). An interactive decision procedure with multiple attributes under risk. *Annals of Operations Research*, 52:151–170.

Jorion, P. (2006). *Value at Risk: The New Benchmark for Managing Financial Risk*. McGraw-Hill, New York, 3rd edition.

Jouini, E., Meddeb, M., and Touzi, N. (2004). Vector-valued coherent risk measures. *Finance and Stochastics*, 8(4):531–522.

Maccheroni, F., Marinacci, M., and Rustichini, A. (2006). Ambiguity aversion, robustness, and the variational representation of preferences. *Econometrica*, 74(6):1447–1498.

Markowitz, H. (1952). Portfolio selection. *Journal of Finance*, 7(1):77–91.

Meucci, A. (2009). *Risk and asset allocation*. Springer-Verlag, Berlin Heidelbergs.

Mosler, K. (2002). *Multivariate Dispersion, Central Regions and Depth: The Lift Zonoid Approach*. Springer-Verlag, New York.

Mosler, K. and Bazovkin, P. (2014). Stochastic linear programming with a distortion risk constraint. *OR Spectrum*, 36(4):949–969.

Rachev, S., Ortobelli Lozza, S., Stoyanov, S., and Fabozzi, F. J. (2008). Desirable properties of an ideal risk measure in portfolio theory. *International Journal of Theoretical and Applied Finance*, 11(1):19–54.

Rockafellar, R. T., Uryasev, S., and Zabarankin, M. (2006). Optimality conditions in portfolio analysis with general deviation measures. *Mathematical Programming*, 108(2-3):515–540.

Rüschendorf, L. (2013). *Mathematical Risk Analysis*. Springer-Verlag, Berlin Heidelberg.

Schrage, C. (2012). Scalar representation and conjugation of set-valued functions. *Optimization*, (ahead-of-print):1–27.

Schrage, C. and Löhne, A. (2013). An algorithm to solve polyhedral convex set optimization problems. *Optimization*, 62(1):131–141.

Sharpe, W. F. (1966). Mutual fund performance. *Journal of Business*, 39:119–138.

Tütüncü, R. H. and Koenig, M. (2004). Robust asset allocation. *Annals of Operations Research*, 132(1-4):157–187.