Robust risk management

Apostolos Fertis*, Michel Baes, Hans-Jakob Lüthi

Institute for Operations Research (IFOR), Eidgenössische Technische Hochschule Zürich, Rämistrasse 101, HG G 21.3, 8092 Zürich, Switzerland

A R T I C L E   I N F O

Article history:
Received 12 May 2011
Accepted 19 March 2012
Available online 28 March 2012

Keywords:
Convex programming
Robust optimization
Risk management

A B S T R A C T

Estimating the probabilities by which different events might occur is usually a delicate task, subject to many sources of inaccuracies. Moreover, these probabilities can change over time, leading to a very difficult evaluation of the risk induced by any particular decision. Given a set of probability measures and a set of nominal risk measures, we define in this paper the concept of robust risk measure as the worst possible of our risks when each of our probability measures is likely to occur. We study how some properties of this new object can be related with those of our nominal risk measures, such as convexity or coherence. We introduce a robust version of the Conditional Value-at-Risk (CVaR) and of entropy-based risk measures. We show how to compute and optimize the Robust CVaR using convex duality methods and illustrate its behavior using data from the New York Stock Exchange and from the NASDAQ between 2005 and 2010.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

In quantitative risk management, risk measures are used to determine a preferential order among financial positions with random outcome. Each financial position is seen as a random variable that maps each state of nature \( \omega \) to a real number. This number corresponds to the reward ensured by the financial position when the state \( \omega \) occurs. Risk measures are designed to take into account the trade-off between the magnitudes of the values that a position can take, and the risk or variability in these values. Mathematically, they are mappings of a space of random variables to the extended real line. The choice of the risk measure determines the investment risk profile.

A portfolio is a linear combination of some available assets, each characterized by a cost and a random variable representing their income, in the limits of a fixed budget. Markowitz (1952) defines the risk of a portfolio as a weighted sum of its expected return and its variance. The ratio between the two weights determines the risk profile of the investor. Since Markowitz’s breakthrough paper, many other risk measures have been introduced. Most notably, Value-at-Risk (VaR), a quantile of the position’s probability distribution, has been extensively used (RiskMetrics, 1995). However, VaR has been criticized for not detecting unfavorable behavior in the tails of the probability distribution (Donnelly and Embrechts, 2010). This observation triggered the introduction of classes of risk measures that satisfy some desirable properties. For instance, the class of convex risk measures (Föllmer and Schied, 2002) gathers monotone and convex mappings that satisfy a translation invariance property (see in Section 2). Every convex risk measure can be expressed as the conjugate of some “penalty” function defined in a space of signed measures. This representation can be used to compute optimal portfolios and to assess their value of flexibility (Lüthi and Doege, 2005). Moreover, convex risks are closely connected with the concept of the optimized certainty equivalent introduced in Ben-Tal and Teboulle (2007). Coherent risk measures (Artzner et al., 1999) forms the subclass of convex risk measures that are positively homogenous (see Section 2). They can be expressed as the worst-case expectation of the portfolio outcome when the probability measure of the assets returns varies in some uncertainty set (Artzner et al., 1999). For instance, Conditional Value-at-Risk (CVaR), that is, the expected value of a portfolio if its loss lies beyond some quantile of its distribution, is such a coherent risk measure. The composition of a portfolio optimal with respect to its CVaR can be computed using the dual representation of this risk measure (Rockafellar and Uryasev, 2000; Shapiro et al., 2009).

Due to the intrinsic uncertainty of the environment they describe, it can happen that the data defining a problem is not known exactly. As a result, it is possible that the optimal solution computed for the erroneous problem we have is far from optimal, or not even feasible, for the actual problem. Robust optimization is now increasingly used to tackle his issue. It considers that the actual data of a problem belongs to a predefined uncertainty set \( \mathcal{S} \), then assigns to every feasible point the worse objective value among all the problems with data in \( \mathcal{S} \). The optimal point it returns is then the feasible point with the best of those worse values, and is thereby immune to data uncertainties. In linear optimization problems, Soyster (1973) considers box-type uncertainty sets and (Ben-Tal and Nemirovski, 1998, 1999, 2000) ellipsoidal ones. In linear and in mixed integer optimization, problems with budgeted...
uncertainty sets (Ben-Tal et al., 2009) for their constraints are efficiently solved in Bertsimas and Sim (2003, 2004). Interestingly, minimizing the coherent risk measure of an affine combination of random variables can be reformulated as a robust optimization problem; an explicit description of the uncertainty set is given in Bertsimas and Brown (2009).

Not only the data of the problem can be subjected to errors, but also the probability distribution model for the random positions, as it is constructed, among other sources, from possibly corrupted historical data. Several approaches deal with this issue. A first possibility consists in defining a class of parameterized probability distributions, among which the actual one is determined by standard parameter estimation procedures. Moreover, these procedures can yield confidence intervals, which can be used as uncertainty sets in a robust optimization framework (Bertsimas and Pachamanova, 2008). Robust solutions are unavoidably conservative: due to the typically infinite number of extra constraints, the obtained return is often much lower than the non-robust return. Among other techniques to tackle this issue, let us point the one developed in Zymler et al. (2011). There, the set of constraints ensures a certain minimal return when the probability parameters belong to a certain small set, and another minimal return when they belong to another, larger, set.

In this paper, we consider the problem of assessing risk when the probability measure driving the underlying random process is not known exactly, but resides in some uncertainty set, called here the scenarios set. Instead of using a single risk measure for our whole problem, we have one risk measure per probability function from our scenario set. In other words, we use a family of risk measures, each indexed by a probability measure from the scenarios set. We define our robust risk measure as an appropriate combination of them. Some of their properties, such as convexity and coherence can be shown to be transferred to our robust risk measure (see Propositions 2.1 and 2.2). Our definition is then particularized to define a robust counterpart to Conditional Value-at-Risk (CVaR) and to entropy-based risks in the context of two-stage structured uncertainty sets (see SubSection 3.1 and Section 4 for precise definitions). We also provide efficient algorithms to compute these risks. Robust CVaR has been successfully used in hydro-electric pumped storage plant management (Fertis and Abbeg, 2010), and has been connected with regularization in portfolio optimization (Fertis et al., 2011).

Special cases for the Robust CVaR under two-stage structured uncertainty sets, namely the cases when the probability measures are discrete or when the probability measure uncertainty set is the whole set of probability measures on the considered space, have been already studied (Zhu and Fukushima, 2009). In this paper, we deal with continuous or discrete probability measures, and generic norm-bounded uncertainty sets. In the case of Robust CVaR for continuous distributions, our result enables portfolio optimization through the stochastic average approximation method, which discretizes the sample space (Shapiro et al., 2009). Robust risk measures with different probability distribution uncertainty sets have been considered in the past as well. The worst-case CVaR when certain moments of the assets’ probability distribution are known can be expressed as a finite-dimensional robust optimization problem, and can be efficiently computed if the distribution is discrete or continuous of a special kind (Natarajan et al., 2009). The worst-case CVaR can be computed through linear optimization when the probability distribution uncertainty is structured (Pflug et al., 2012). The worst-case CVaR under certain moment information has been considered in the framework of chance constrained optimization, and has been compared to the worst-case VaR under certain moment information (Zymler et al., in press).

The paper is structured as follows:

- In Section 2, we define the robust risk measure with reference to a family of nominal risk measures and an uncertainty set for the probability measure that drives the random process. We investigate the structure of the robust risk measure when the family of nominal risk measures contains convex or coherent risks.
- In Section 3, we define Robust CVaR, as the robust risk measure corresponding to CVaR. When the scenario set is structured in two stages, and uncertainty is limited in the second-stage probability distribution, we show how to compute the Robust CVaR of a position, and how to compute portfolios that optimize the Robust CVaR. The complexity of the proposed algorithms is almost the same as the complexity of the corresponding algorithms for CVaR.
- In Section 4, we define the robust entropy-based risks, as the robust risk measures corresponding to the entropy-based risks. We show that these risks can be computed using convex optimization methods.
- In Section 5, we compare the performance of the Robust CVaR-optimal and CVaR-optimal portfolios under various probability measures. The probability distribution models were constructed using historical data of 20 stocks from various sectors traded in the New York Stock Exchange (NYSE) and the National Association of Securities Dealers Automated Quotations (NASDAQ) for the period between 2005 and 2010.

2. Robust risk measures and its representation

It is important to know whether the risk measure defined through the application of the robust optimization paradigm to deal with probability distribution uncertainty follows certain principles of consistent decision making, as the ones required in the denition of convex and coherent risk measures. The used terminology and mathematical background can be found in the appendix.

First, we present the definition of convex risk measures, and the representation theorem for them (Föllmer and Schied, 2002; Shapiro et al., 2009).

Definition 2.1. Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and consider the random variables classes $L^1(\Omega, \mathcal{F}, \mathbb{P})$. A mapping $\rho : L^1(\Omega, \mathcal{F}, \mathbb{P}) \to \mathbb{R}$ is called a convex risk measure if it satisfies the following properties:

- **Monotonicity:** If $X_1, X_2 \in L^1(\Omega, \mathcal{F}, \mathbb{P})$, and $X_1 \leq X_2, \mathbb{P}$-a.s., then, $\rho(X_1) \geq \rho(X_2)$.
- **Translation invariance:** If $X, A \in L^1(\Omega, \mathcal{F}, \mathbb{P})$, and $A = \alpha, \mathbb{P}$-a.s., $\alpha \in \mathbb{R}$, then, $\rho(X + A) = \rho(X) - \alpha$.
- **Convexity:** If $X_1, X_2 \in L^1(\Omega, \mathcal{F}, \mathbb{P}), 0 \leq \lambda \leq 1$, then, $\rho(\lambda X_1 + (1 - \lambda)X_2) \leq \lambda \rho(X_1) + (1 - \lambda)\rho(X_2)$.

Let us fix once forever a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We view the probability $\mathbb{P}$ as a reference probability measure. To allow for probability changes in our model, and in contrast with the standard approach, we do not assume that the probability that drives the random process of our problem is $\mathbb{P}$, but merely one that is only minimally related to $\mathbb{P}$. Specifically, with $\mathbb{P}^\infty$ be the set of all probability measures on $(\Omega, \mathcal{F})$, let

$$\mathcal{P} = \left\{ P \in \mathbb{P}^\infty \mid P \ll \mathbb{P} \text{ and } \frac{dP}{d\mathbb{P}} \in L^\infty(\Omega, \mathcal{F}, \mathbb{P}) \right\},$$

(2.1)

where $P \ll \mathbb{P}$ means that $P$ is absolutely continuous with respect to $\mathbb{P}$, that is, that every $\mathbb{P}$-negligible set is also $P$-negligible, and $dP/d\mathbb{P}$ is a Radon-Nikodym derivative. In this case, we can associate a probability measure $\mathbb{P}^*$ to any probability measure $P \in \mathcal{P}$, which satisfies $\mathbb{P} = \mathbb{P}^*/\mathbb{P}$, $\mathbb{P}$-a.s.
The following propositions determine essential properties of the robust risk measure, and its subdifferentials, when the risk measures in \( \mathcal{S} \) are convex or coherent. The subdifferentials determine descent directions of \( \rho^u \) in \( L^1(\Omega, \mathcal{F}, P) \), which, according to the standard duality theory in convex analysis (Rockafellar, 1970), can be computed by solving a convex optimization problem. In the following proposition, the standard notation \( \text{lscc} \) and \( \text{cv} \) is defined in the appendix.

**Proposition 2.1.** Suppose that \( \mathcal{S} \) is weakly-* compact and that, for every \( P \in \mathcal{S} \), the risk measure \( \rho_P \) restricted to \( L^1(\Omega, \mathcal{F}, P) \) is proper, lower semicontinuous, and convex. Denote by \( \rho^u \) the penalty function of the restriction of \( \rho_P \) to \( L^1(\Omega, \mathcal{F}, P) \) provided by Theorem 2.1. Define, for every \( Q \in \mathcal{P} \) the mapping

\[
K_Q : \mathcal{M} \to \mathbb{R}
\]

where

\[
P \mapsto K_Q(P) := \text{lscc} \rho^u(Q) \text{ if } P \in \mathcal{S}.
\]

\[
K_Q(P) := +\infty \text{ otherwise.}
\]

If \( K_Q \) is proper and weakly-* lower semicontinuous, then \( \rho^u \) is a proper, lower semicontinuous mapping and for every \( X \in L^1(\Omega, \mathcal{F}, P) \):

\[
\rho^u(X) = \sup_{Q \in \mathcal{P}} \left( E_Q[-X] - \text{lscc} \rho^u(Q) \right).
\]

In addition, if \( \rho^u(X) \in \mathbb{R} \) and \( X \in L^1(\Omega, \mathcal{F}, P) \), the subdifferential of \( \rho^u \) at \( X \) is given by:

\[
\partial \rho^u(X) = \text{argmax}_{Q \in \mathcal{P}} \left( E_Q[-X] - \text{lscc} \rho^u(Q) \right).
\]

**Proof.** By definition, we can write for every \( X \in L^1(\Omega, \mathcal{F}, P) \):

\[
\rho_P(X) = \sup_{Q \in \mathcal{P}} \left( E_Q[-X] - \text{lscc} \rho^u(Q) \right).
\]

Thus,

\[
\rho^u(X) = \sup_{Q \in \mathcal{P}} \sup_{P \in \mathcal{S}} \left( E_Q[-X] - \text{lscc} \rho^u(Q) \right) = \sup_{Q \in \mathcal{P}} \left( E_Q[-X] - \text{inf}_{P \in \mathcal{S}} \rho_P(Q) \right).
\]

(2.8)

We will prove that the mapping \( Q \to \text{inf}_{P \in \mathcal{S}} \rho_P(Q) \) is proper, \( Q \in \mathcal{P} \). To prove this, it suffices to prove the following two claims:

(a) For any \( Q \in \mathcal{P} \), \( \text{inf}_{P \in \mathcal{S}} \rho_P(Q) > -\infty \). Since for any \( P \in \mathcal{S} \), \( Q \to \rho_P(Q) \) is a proper mapping, we have that for any \( P \in \mathcal{S} \), and any \( Q \in \mathcal{P} \), \( \rho_P(Q) > -\infty \). Observe that

\[
\text{inf}_{P \in \mathcal{S}} \rho_P(Q) = \inf_{P \in \mathcal{P} \cap \mathcal{M}} \rho_P(Q) \in \mathbb{R}.
\]

Given that \( K_Q \) is weakly-* lower semicontinuous, we conclude that \( \{ P \in \mathcal{M} | K_Q(P) \in \mathbb{R} \} \) is weakly-* closed. Since \( \mathcal{S} \) is weakly-* compact, the intersection \( \mathcal{S} \cap \{ P \in \mathcal{M} | K_Q(P) \in \mathbb{R} \} \) is weakly-* compact. For any \( P \in \mathcal{S} \cap \{ P \in \mathcal{M} | K_Q(P) \in \mathbb{R} \} \), \( \rho^u(Q) \in \mathbb{R} \). The mapping \( P \to \rho^u(Q) \), \( P \in \mathcal{S} \cap \{ P \in \mathcal{M} | K_Q(P) \in \mathbb{R} \} \), is the restriction of \( K_Q \) to \( \mathcal{S} \cap \{ P \in \mathcal{M} | K_Q(P) \in \mathbb{R} \} \), and has been assumed to be weakly-* lower semicontinuous. Thus, \( \text{inf}_{P \in \mathcal{S}} \rho^u(Q) > -\infty \).

(b) There exists \( Q \in \mathcal{P} \) such that \( \text{inf}_{P \in \mathcal{S}} \rho^u(Q) \in \mathbb{R} \).

Consider any \( Q \in \mathcal{P} \). Since \( K_Q \) is a proper mapping, there exists \( P \in \mathcal{S} \) such that \( \rho_P(Q) \in \mathbb{R} \). Consequently, \( \mathcal{S} \cap \{ P \in \mathcal{M} | K_Q(P) \in \mathbb{R} \} \) is not empty, and \( \text{inf}_{P \in \mathcal{S}} \rho^u(Q) \in \mathbb{R} \).

Given Theorem 2.1, Eq. (2.8), and the fact that mapping \( Q \to \text{inf}_{P \in \mathcal{S}} \rho_P(Q) \) is proper, we conclude that \( \rho^u \) is a proper, lower semicontinuous mapping, and a convex risk measure. Proposition A.1 implies that \( \rho^u(Q) \) and \( \text{lscc}(\text{cv}(\text{inf}_{P \in \mathcal{S}} \rho_P(Q))) \) is pair of conjugate mappings. Hence, Eqs. (2.6) and (2.7) follow (Shapiro et al., 2009, p. 407).
Proposition 2.2. Assume that the set of scenarios $S$ is weakly-" compact and that for every $P \in S$, the restriction of the risk measure $\rho_P$ to $L^1(\Omega, F, P)$ is a real-valued coherent risk measure. Denote by $Q_P$ the subset of $P$ provided by Theorem 2.2. If $Q_P$ is non-empty, weakly-" compact, and convex, then the robust risk measure $\rho^*_P : L^1(\Omega, F, P) \to \mathbb{R}$ is real-valued and coherent. For every $X \in L^1(\Omega, F, P)$, we have:

$$\rho^*_P(X) = \max_{Q \in Q_P} E_Q[-X]$$

s.t. $Q \in \text{cl}(\text{conv}(\bigcup_{P \in S} Q_P))$. \hfill (2.9)

and

$$\partial \rho^*_P(X) = \arg\max_{Q \in Q_P} E_Q[-X]$$

s.t. $Q \in \text{cl}(\text{conv}(\bigcup_{P \in S} Q_P))$. \hfill (2.10)

Proof. By Theorem 2.2, since $\rho_P$ is a coherent risk measure with test probability set $Q_P$, it is known that:

$$\rho_P(X) = \max_{Q \in Q_P} E_Q[-X].$$

Define $\alpha^*(Q) = \frac{1}{m} \sum_{i=1}^m \alpha_i(Q)$ ($\alpha_i$ is the indicator function of the set $A_i$, see Appendix). Since $S$ is weakly-" compact, we have that:

$$\sup_{P \in S} \rho_P(X) = \max_{P \in S} \rho_P(X).$$

Given the representation of $\rho_P(X)$, we have

$$\rho^*_P(X) = \max_{P \in S} \max_{Q \in Q_P} E_Q[-X] = \max_{Q \in \text{conv}(\bigcup_{P \in S} Q_P)} E_Q[-X]$$

$$= \max \{ E_Q[-X] - \alpha^*(Q) : Q \in Q_P \}. \hfill (2.11)$$

Thus, $\rho^*_P$ is a real-valued mapping. Since $S$ is non-empty, and for any $P \in S$, $Q_P$ is non-empty, $\bigcup_{P \in S} Q_P$ is non-empty too. Hence, $\alpha^*(Q)$, which is a support function, is a proper mapping. Due to Proposition A.1, and Eq. (2.11), we conclude that

$$\rho^*_P(X) = \max_{Q \in Q_P} \{ E_Q[-X] - \text{isc( conv(} \bigcup_{P \in S} Q_P) \text{) (Q))} \}$$

$$= \max_{Q \in Q_P} \left( E_Q[-X] - \text{isc( conv(} \bigcup_{P \in S} Q_P) \text{) (Q)} \right). \hfill (2.12)$$

Given Eq. (2.12), and because $\text{isc( conv(} \bigcup_{P \in S} Q_P) \text{)}$ is an indicator function, we conclude that $\rho^*_P$ is a coherent risk measure (Shapiro et al., 2009, p. 267). Eqs. (2.9) and (2.10) follow immediately (Shapiro et al., 2009, p. 407). \hfill \Box

Observe that any coherent risk measure $\rho_P$ is itself a robust risk measure, if we consider the nominal risk measure family to contain the expectation risk measures, and all the sets of scenarios to be the test probability set $Q_P$.

3. Robust Conditional Value-at-Risk

In this section, we define the Robust Conditional Value-at-Risk (Robust CVaR), the robust risk measure corresponding to Conditional Value-at-Risk (CVaR) under a two-stage structured uncertainty set. We show how to compute the Robust CVaR of a position, and how to compute the optimal-Robust CVaR portfolio, using convex duality, when the scenario set is defined by dividing randomness in two stages, and considering uncertainty only in the first stage.

3.1. Definition

We apply the general framework developed in the previous section to define a robust counterpart of the standard Conditional Value-at-Risk. Let us recall the formal definition of this well-known risk measure. We fix a constant $\beta \in [0,1]$ and we set, for every $X \in L^1(\Omega, F, P)$:

$$\text{CVaR}_\beta(X) = \max_{Q} E_Q[-X] = \max_{Q} \ E_P[-G|X]$$

s.t. $Q \subseteq P$ s.t. $E[Q[G]] = 1$. \hfill (3.1)

$$G \in L^\infty(\Omega, F, P).$$

The set $P$ used above was defined in (2.1). Now, we need to specify a scenario set $S \subseteq P$. Let $\{P_1, \ldots, P_n\} \subset P$ and denote by $A_r = \{\xi \in R | \frac{1}{n} \sum_{i=1}^n \xi_i = 1\}$ the standard $r-1$-dimensional simplex. Fix $\xi_0 \in A_r$, a norm $\| \cdot \|_P$ on $R^n$, and $\phi > 0$, and set:

$$S := \{\frac{1}{n} \sum_{i=1}^n \xi_i | \xi \in A_r, \| \xi - \xi_0 \|_P < \phi\}. \hfill (3.2)$$

Note that we could consider more generally that $\xi$ is restricted to a closed, hence compact, subset $\Xi$ of $A_r$.

Observe that:

$$\text{CVaR}_\beta^*(X) = \max_{Q \in L^\infty(\Omega, F, P)} E_Q[-G|X]$$

s.t. $G \subseteq P$ s.t. $E_{[G]} = 1$. \hfill (3.3)

The mapping $\xi \to \sum_{k=1}^n \xi_k dP_k/dP$ is continuous, as it is a linear function between two finite-dimensional spaces. Thus, $\text{CVaR}_\beta^*$ is a continuous function of $\xi$.

According to Definition 2.3, the Robust CVaR corresponding to the scenario set presented in Eq. (3.2) is defined by

$$\text{RCVaR}_\beta^*(X) = \sup_{\xi \in \Xi} \text{CVaR}_\beta^*(X) = \max_{\sum_{k=1}^n \xi_k P_k} \text{CVaR}_\beta^*(X). \hfill (3.4)$$

The maximum in Eq. (3.3) is always attained, since the feasible $\xi$ form a compact set, and $\text{CVaR}_\beta^*$ is a continuous function of $\xi$.

Then, combining Eqs. (3.1) and (3.2), we conclude that the Robust CVaR is given by

$$\text{RCVaR}_\beta^*(X) = \max_{\xi \in \Xi} E_{\xi}[G|X]$$

s.t. $G \subseteq P$ s.t. $E_{[G]} = 1$. \hfill (3.4)

$$\xi \in A_r, \| \xi - \xi_0 \|_P < \phi$$

$$G \in L^\infty(\Omega, F, P).$$

The above problem is convex. If $X$ is a discrete random variable with $n$ states, then, we can choose $\Omega = \{1,2,\ldots,n\}$, which sets the problem in space $R^n$. In addition, if $\| \cdot \|$ is the $l_1$-norm or the $l_\infty$-norm, the problem is linear, and if $\| \cdot \|$ is the $l_2$-norm or any quadratic norm, it is a second-order cone programming problem. If $X$ is a continuous random variable, then the problem becomes infinite dimensional. It can be approximately solved by discretizing the probability distribution of $X$.

3.2. Dual problem

Consider a set of assets $\{X_1, X_2,\ldots,X_m\}$. A portfolio $\theta \in \Delta_m$ gives rise to random variable $\sum_{i=1}^m \theta_i X_i$. We would like to compute
the portfolio with the minimum Robust CVaR. If $A \subseteq A \times L^\infty(\Omega, \mathcal{F}, \mathbb{P})$ is the feasible set of Problem (3.4), the Robust CVaR portfolio problem can be expressed as

$$\min_{\theta \in \mathbb{R}^n} \max_{(\xi, G) \in A} \left[ -\sum_{j=1}^{m} \phi_j X_j \right]$$

(3.5)

or, equivalently,

$$\min_{\theta \in \mathbb{R}^n} Y \geq E_{P}[ -\sum_{j=1}^{m} \phi_j X_j ], \quad \forall (\xi, G) \in A$$

(3.6)

$$\theta \in A_m.$$

Even if $\Omega$ is finite, which implies that $G$ is finite dimensional, the set $A$ is infinite. To remove this source of infinity, we will compute a dual of Problem (3.4) with no duality gap. Up to now, only the special cases when $\Omega$ is finite, and norm $\| \cdot \|$ defines box-type or ellipsoidal uncertainty, or when $\phi$ is large enough for the constraint on $\xi$ to reduce to $\xi \in A_m$ have been solved (Zhu and Fukushima, 2009). We prove the duality result in the generic framework we have described, which covers the cases of an infinite $\Omega$, and a generic $\| \cdot \|$ norm for the probability measure uncertainty. Note that the case of a finite $\Omega$ is simple, since it is covered by linearized optimization duality, but this is not true for the case of an infinite $\Omega$. We denote here the conjugate norm of $\| \cdot \|$ by $\| \cdot \|_\ast$.

**Theorem 3.1.** A dual of Problem (3.4) with no duality gap is

$$\min_{\theta \in \mathbb{R}^n} \left[ -\sum_{j=1}^{m} \phi_j X_j \right]$$

s.t.

$$\forall (\xi, G) \in A, \quad \theta \in A_m.$$  

(3.7)

**Proof.** We will compute a dual of

$$\min_{E_{P}[G]}
\left[ t + \sum_{j=1}^{m} \frac{1}{\xi_j} \frac{d P_j}{d P} \right]$$

s.t.

$$E_{P}[G] = 1, \quad \xi \in A_\ast,$$

$$G \in L^\infty(\Omega, \mathcal{F}, \mathbb{P}),$$

(3.8)

whose optimal value is opposite to the optimal value of Problem (3.4). Let $K_0$ be the generalized Lorentz cone in $\mathbb{R}^{n+1}$ with reference to norm $\| \cdot \|$ in $\mathbb{R}^n$, and $K_1$ its dual, defined in Appendix A. By adding auxiliary variables $Z \in L^\infty(\Omega, \mathcal{F}, \mathbb{P})$, $y \in \mathbb{R}^r$, $Y_{r+1} \in \mathbb{R}$, Problem (3.8) is expressed in the canonical form

$$\min_{E_{P}[G]}
\left[ t + \frac{1}{\xi_0} \sum_{k=1}^{r} \xi_k \frac{d P_k}{d P} - G - Y \right]$$

s.t.

$$E_{P}[G] = 1, \quad \xi \in A_\ast,$$

$$\xi - \xi_0 = \phi,$$

$$y_{r+1} = \phi,$$

$$\sum_{k=1}^{r} \xi_k = 1, \quad \xi_0 \in \mathbb{R}^+, \quad \xi \in \mathbb{R}^n \setminus (\Omega, \mathcal{F}, \mathbb{P}),$$

$$G \in L^\infty(\Omega, \mathcal{F}, \mathbb{P}),$$

$$y \in \mathbb{R}^r, \quad Y_{r+1} \in \mathbb{R},$$

(3.9)

We are going to derive the dual of Problem (3.9) using the dual expression for a generic canonical form problem (Barvinok, 2002, p. 166) (Rockafellar, 1974). We consider dual variables $Z \in L^1(\Omega, \mathcal{F}, \mathbb{P}), \eta \in \mathbb{R}, v = (v_1, v_2, \ldots, v_l)^T \in \mathbb{R}^l, v_{r+1}, t \in \mathbb{R}$. The dual cone of $\mathbb{R}^r \times L^1(\Omega, \mathcal{F}, \mathbb{P}) \times L^1(\Omega, \mathcal{F}, \mathbb{P}) \times K_0$ is $\mathbb{R}^r \times L^1(\Omega, \mathcal{F}, \mathbb{P}) \times L^1(\Omega, \mathcal{F}, \mathbb{P}) \times K_1$, (see Appendix for the definition of $K_1$). The inner product of the left-hand side constraint functions with the dual variables is given by

$$E_{P} \left[ \left( \frac{1}{1 - \beta} \sum_{k=1}^{r} \frac{d P_k}{d P} - G - Y \right) Z \right] + E_{P}[G]\eta + \langle \xi - \eta, y \rangle + y_{r+1} Y_{r+1}$$

$$+ \sum_{k=1}^{r} \xi_k t = \sum_{k=1}^{r} \left( t + v_k + E_{P} \left[ \left( \frac{1}{1 - \beta} \frac{d P_k}{d P} \right) Z \right] \right) \xi_k + E_{P}[-Y + \xi_0 G]$$

$$+ E_{P}[-Y_{r+1} Z] - v^T y + Y_{r+1} Y_{r+1}.$$  

Hence, a dual of Problem (3.9) is

$$\sup_{\xi, \eta, t, y, y_{r+1}} \left[ t + \frac{1}{\xi_0} \sum_{k=1}^{r} \xi_k \frac{d P_k}{d P} - G - Y \right]$$

s.t.

$$t + v_k + E_{P} \left[ \frac{1}{1 - \beta} \frac{d P_k}{d P} Z \right] \leq 0, \quad k = 1, \ldots, r$$

$$-Z + \eta \leq X, \quad P\text{-a.s.}$$

$$-Z \leq 0, \quad P\text{-a.s.}$$

$$\eta, t \in \mathbb{R}, \quad y \in \mathbb{R}^r, \quad Y_{r+1} \in \mathbb{R},$$

$$\langle \xi - \eta, y \rangle + y_{r+1} Y_{r+1} \in K_1.$$  

Define

$$K_2 = \left\{ \left( t + v_k + E_{P} \left[ \frac{1}{1 - \beta} \frac{d P_k}{d P} Z \right], \xi_k \right) \right\}$$

$$-Z + \eta \leq X, \quad P\text{-a.s.}$$

$$-Z \leq 0, \quad P\text{-a.s.}$$

$$\eta, t \in \mathbb{R}, \quad y \in \mathbb{R}^r, \quad Y_{r+1} \in \mathbb{R}.$$  

The primal, Problem (3.9), has a finite optimal value. To prove that the dual, Problem (3.10), has the same optimal value, it suffices to prove that cone $K_0 \times K_1 \times K_{r+1} \times K_0$ is closed in space $L^\infty(\Omega, \mathcal{F}, \mathbb{P}) \times \mathbb{R} \times \mathbb{R}$ under the product of weak-* topology in $L^\infty(\Omega, \mathcal{F}, \mathbb{P})$, and the standard topology in $\mathbb{R}$ (Barvinok, 2002, p. 168). Using Tikhonov’s theorem (Barvinok, 2002, p. 110), to prove the latter, it suffices to prove that cones $K_0, K_1, K_{r+1}, K_0$ are closed in spaces $L^\infty(\Omega, \mathcal{F}, \mathbb{P}), \mathbb{R}, \mathbb{R}^r, \mathbb{R}, \mathbb{R}$, respectively, under the weak-* and the standard topologies, respectively.

Set $\{ \xi \in \mathbb{R}^n, \sum_{k=1}^{r} \xi_k = 1 \}$ is a compact convex base for $\mathbb{R}^n$. (see Definition A.6 in the Appendix). Set $\{ (y^T, y_{r+1})^T \in \mathbb{R}^{r+1} \}$ is a compact convex base for $K_0$. Set

$$L^\infty(\Omega, \mathcal{F}, \mathbb{P}) \subseteq \{ G \in L^\infty(\Omega, \mathcal{F}, \mathbb{P}) | E_{P}[G] = 1 \}$$

is a convex base for $L^\infty(\Omega, \mathcal{F}, \mathbb{P})$. We will prove that it is also compact in the weak-* topology. Consider open set

$$U = \{ X \in L^1(\Omega, \mathcal{F}, \mathbb{P}) | \| X \|_1 < 1 \},$$

which contains the origin in space $L^1(\Omega, \mathcal{F}, \mathbb{P})$, by Alaoglu’s theorem,$\| U \| = \{ G \in L^\infty(\Omega, \mathcal{F}, \mathbb{P}) | E_{P}[G] = 1 \} \subseteq U$. Assume that $G \in L^\infty(\Omega, \mathcal{F}, \mathbb{P}) \subseteq U$. Then, $\| G \|_1 = E_{P}[G] = 1$. For any $X \in U,$

$$\| E_{P}[G] \|_1 \leq \| G \|_1 \| X \|_1 < \| G \|_1 \leq 1,$$

implying that $G \in U$. Set $L^\infty(\Omega, \mathcal{F}, \mathbb{P})$ is weak-* closed, because it is the intersection of the weak-* closed sets $L^\infty(\Omega, \mathcal{F}, \mathbb{P})$ and $\{ G \in L^\infty(\Omega, \mathcal{F}, \mathbb{P}) | E_{P}[G] = 1 \}$. Since it is weak-* closed and a subset of the weak-* compact set $U$, it is also weak-* compact. Using Tikhonov’s theorem (Barvinok, 2002, p. 110), we conclude that
cones $\mathbb{R}_+^r \times L_1^e(\Omega, \mathcal{F}, \mathbb{P}) \times L_1^e(\Omega, \mathcal{F}, \mathbb{P})$, and $\mathbb{R}_+^r \times K_2$, have compact convex bases as well.

Consequently, cones $K_2$, $K_q$, $K_n$, $K_{n_1}$, $K_n$, and $K_q$ are weakly-closed (Barvinok, 2002, p. 171). Hence, Problem (3.10) has the same finite optimal value as Problem (3.9).

By reverting the sign of Problem (3.10), we get

$$\begin{align*}
&\text{min} & \eta + t - \varphi_1^TV - \psi_1v_{r+1} \\
&\text{s.t.} & t + v_k + E_0 \left[ \frac{1}{T} \mathbb{E} \left[ Z \right] \right] \leq 0, & k = 1, 2, \ldots, r \\
& & Z + \eta \in X, & \mathbb{P}\text{-a.s.} \\
& & Z \in L^1(\Omega, \mathcal{F}, \mathbb{P}), & \mathbb{P}\text{-a.s.}
\end{align*}$$

which is a dual of Problem (3.4) with no duality gap, and equivalent to Problem (3.7) (see Appendix for the definition of $K_1'$). \(\square\)

Given Theorem 3.1, Problem (3.5) is expressed as

$$\begin{align*}
&\text{min} & \eta + t + \varphi \|V\|_1 + \varphi_1^TV \\
&\text{s.t.} & t + v_k \geq \frac{1}{T} \mathbb{E}_0 \left[ Z \right], & k = 1, 2, \ldots, r \\
& & Z \geq -\eta - \sum_{j=1}^{m} \theta_j X_j, & \mathbb{P}\text{-a.s.} \\
& & \sum_{j=1}^{m} \theta_j = 1 \\
& & \eta, t \in \mathbb{R}, & v \in \mathbb{R}^r, \\
& & Z \in L^1(\Omega, \mathcal{F}, \mathbb{P}), & \theta \in \mathbb{R}^m.
\end{align*}$$

If $X$ is discrete random variables with $n$ states each, we can choose $\Omega = \{1, 2, \ldots, mn\}$, and Problem (3.17) is a convex optimization problem in space $\mathbb{R}^{m+1}-\mathbb{R}$. In addition, if $\|\cdot\|$ is the $l_1$-norm or the $l_\infty$-norm, the problem is linear, and if $\|\cdot\|$ is the $l_2$-norm or any quadratic norm, it becomes a second-order cone programming problem. If $X$ is continuous random variables, then, we have to choose $\Omega = \mathbb{R}^m$, and the problem becomes infinite-dimensional. It can be approximately solved by discretizing the probability distribution of random vector $(X_1, X_2, \ldots, X_m)^T$, which enables the use of a finite $\Omega$. However, the number of discrete states grows exponentially with $m$. This problem is also present in the computation of optimal-CVaR portfolios. To face such a problem the stochastic average approximation method can be used (Shapiro et al., 2009). The computational overhead of Robust CVaR with regard to CVaR consists of variables $v$ and $t$, which add $r + 1$ dimensions to the space of the problem, an increase in the number of linear constraints by $rnn$, and the computation of the dual norm $\|\cdot\|$ in the objective.

4. Robust entropy-based risk measures

Consider the framework in Definition 2.3, assume that $\Omega$ is finite, and $n = |\Omega|$. For $\mathcal{F} = 2^\Omega$, and $P(\omega) = 1/n$, we have that $L^1(\Omega, \mathcal{F}, \mathbb{P}) = L^\infty(\Omega, \mathcal{F}, \mathbb{P}) = \mathbb{R}^n$, and

$$\mathcal{P} = \left\{ P \in \mathbb{R}^n : P(\omega) \geq 0, \omega \in \Omega, \sum_{\omega \in \Omega} P(\omega) = 1 \right\}.$$ 

If $\mu(\omega) \geq 0, \omega \in \Omega$, and

$$\mu^\phi(Q) = \sum_{\omega \in \Omega} \mu(\omega) \ln \left( \frac{Q(\omega)}{P(\omega)} \right) - Q(\omega) + P(\omega), \quad Q, P \in \mathcal{P},$$

we let the family of nominal risk measures be the family of entropy-based risk measures, which are convex risk measures represented (Lüthi and Doege, 2005) by

$$\rho^K_\phi(X) = \max_{Q \in \mathcal{P}} E_Q[\max\{-X\}] - \mu^\phi(Q), \quad X \in \mathbb{R}^n.$$ 

The entropy-based risk of a random position $X \in \mathbb{R}^n$ can be computed through convex optimization methods (Lüthi and Doege, 2005; Ben-Tal and Teboulle, 2007).

If we use the two-stage uncertainty model for the scenario set described in Section 3, the corresponding robust entropy-based risk measure is

$$\rho^K_\phi(X) = \max_{Q, P} \left( E_Q[-X] - \mu^\phi \left( \sum_{t=1}^{m} \psi_t \right) \right), \quad X \in \mathbb{R}^n.$$ 

Function $\mu^\phi(Q)$ is convex in $(Q, P)$ (Boyd and Vandenberghe, 2004, p. 90), which implies that $\mu^\phi \left( \sum_{t=1}^{m} \psi_t \right)$ is convex in $(\xi, \zeta)$. Sets $\Xi$ and $\mathcal{P}$ are also convex. We conclude that Problem (4.3) is a convex optimization problem with twice continuously differentiable objective function and constraints. It is solvable through interior point methods in $O(\log(1/\epsilon))$ time, where $\epsilon$ is the desired accuracy (Boyd and Vandenberghe, 2004).

In the case that $\mu(\omega) = 1$, the nominal risk measures are the expected exponential losses

$$\rho^K_\phi(X) = \max_{Q \in \mathcal{P}} \ln E_Q[e^{-X}], \quad X \in \mathbb{R}^n.$$ 

The corresponding robust risk measure becomes

$$\rho^K_\phi(X) = \max_{Q, P} \left( \ln E_Q \sum_{t=1}^{m} \psi_t e^{-X} \right) = \max_{Q, P} \ln \left( \sum_{t=1}^{m} \psi_t e^{-X} \right) = \max_{Q, P} \ln \left( \sum_{t=1}^{m} \psi_t e^{-X} \right), \quad X \in \mathbb{R}^n.$$ 

Since function $\ln(u), u > 0$, is a monotone increasing function, to compute the risk, it suffices to compute $\max_{Q, P} \sum_{t=1}^{m} \psi_t e^{-X}$. 

5. Empirical performance

To evaluate the properties of the optimal-Robust CVaR portfolios, we collect the weekly stock prices for 20 stocks traded in NYSE and NASDAQ for the period between 2005 and 2010. Table 5.1 shows the stocks which are considered and the sector of the economy to which each stock belongs. We also collect the US Department of Treasury 1-year bond rates for the same time period. We normalize them with reference to a weekly time interval to obtain the weekly risk-free rates for the considered period. We use the weekly risk-free rates to compute the weekly discounted return rates of the stocks during this period. The procedure results in $n = 260$ samples for the discounted return rates of the stocks.

Due to the 2008 financial crisis, the returns depict a high volatility. Fig. 5.1 shows the evolution of the weekly discounted return rate of the S&P500 index for the time period between 2005 and 2010. The market appears to move through various states with different degrees of volatility. The volatility is high especially during the crisis of late 2008. Using these data, the accuracy of the probability measure estimation is debatable. To tackle this inefficiency, we apply the robust risk management principles described in this paper and more specifically, the Robust CVaR.

The set $\Omega = \{1, 2, \ldots, n\}$ is identified with the sample set. The nominal probability measure $P_0$ assigns probability $1/n$ to each $i \in \Omega$. Let $X_i = (x_{i1}, x_{i2}, \ldots, x_{in})^T \in \mathbb{R}^n$ be the random variables representing the returns of the stocks, and $D = (d_1, d_2, \ldots, d_n) \in \mathbb{R}^n$ be the random variable representing the return of the S&P500 index. If $a_0 = (a_{01}, a_{02}, \ldots, a_{0n})^T$, we express $P_0$ in the two-stage format $P_0 = \sum_{t=1}^{m} \psi_t a_t$, where $r = 9$. Such an expression is based on the distribution of the return of the S&P500 index. Probability distributions $P_k, k = 1, 2, \ldots, 9$, are characterized by a low, medium, or high expectation, combined with a low, medium, or high variance. They are determined using 9 normal distributions with low, medium, or high expectation, and low, medium, or high variance. The expectation, as well as the standard deviations, of these 9 normal
The choice of the certain quantiles is intuitive. Weight \( w_i, i \in \Omega, k = 1, 2, \ldots, 9 \), determines the contribution of sample \( i \) to distribution \( P_k \). It is equal to the value of the normal distribution \( N(\mu_k, \sigma_k^2) \) probability density function at \( d_i \). That is, if the probability density of \( N(\mu_k, \sigma_k^2) \) at the \( i \)-th sample of the S&P500 index return is high, then, the \( i \)-th sample contributes much to distribution \( P_k \). According to this principle, the mass that second-stage probability measure \( P_k \) places at \( i \) is given by

\[
P_k(i) = \frac{w_{ik}}{\sum_{j=1}^{n} w_{jk}}.
\]

(5.1)

To ensure that \( P_0 = \sum_{i=1}^{9} \xi_{ik} P_k \), the vector \( \xi \) of nominal probabilities over the scenarios is given by

\[
\xi_{ik} = \frac{\sum_{j=1}^{n} w_{jk}}{n}.
\]

(5.2)

By solving Problem (3.17), we can compute the optimal-\( \gamma \)-CVaR portfolio for various values of \( \phi \). For \( \phi = 0 \), we get the optimal-\( \gamma \)-CVaR portfolio. By solving Problem (3.4) or Problem (3.7), we can compute the Robust CVaR, that is the worst-case CVaR, of some portfolio for various values of \( \phi \). We choose \( \beta = 0.90 \). We compute the optimal-\( \gamma \)-CVaR portfolios with various \( \phi = \gamma \in [0,0.2] \). We evaluate the Robust CVaR of each one of them with various \( \phi = \gamma \in [0,0.2] \). If \( \gamma \neq \delta \), the evaluation distribution is different from the one used in the portfolio design, and thus, we have an out-of-sample test.

In Fig. 5.2, we observe the worst-case CVaR for the optimal-CVaR and the optimal-\( \gamma \)-CVaR portfolio that protects against \( \gamma = \delta \). The worst-case CVaR of the optimal-\( \gamma \)-CVaR portfolio grows at a smaller rate because it protects against the uncertainties in the probability measure that are bounded by \( \gamma \). However, in practice, \( \delta \) is not known, and hence, this portfolio cannot be computed. That is why we also perform out-of-sample tests for the optimal-\( \gamma \)-CVaR portfolios, which are represented by the cases where \( \gamma \neq \delta \). In Fig. 5.3, we observe the worst-case CVaR of the optimal-\( \gamma \)-CVaR portfolios for various values of \( \gamma \). For small values of \( \delta \), the optimal-\( \gamma \)-CVaR portfolio performs better. This happens because the optimal-\( \gamma \)-CVaR portfolios are too conservative and the size of uncertainty is small. As \( \delta \) grows, the worst-case CVaR of all portfolios grows, but the one of the optimal-\( \gamma \)-CVaR portfolio grows at a higher rate than the ones of the optimal-\( \gamma \)-CVaR portfolios. For values of \( \delta \) around 0.05, the optimal-\( \gamma \)-CVaR portfolios are clearly better than the optimal-\( \gamma \)-CVaR portfolio, and for higher values of \( \delta \) a significant difference in their performance appears. This happens because the optimal-\( \gamma \)-CVaR portfolios are able to offer protection against the scenario set, whose size is determined by \( \delta \). The difference in the performance grows up to 11.71%, the improvement that the optimal-\( \gamma \)-CVaR with \( \gamma = 0.15 \) portfolio shows with respect to the optimal-CVaR portfolio at \( \delta = 0.20 \).

---

**Table 5.1** Stocks considered in the portfolio.

<table>
<thead>
<tr>
<th>Company name</th>
<th>Stock symbol</th>
<th>Sector</th>
</tr>
</thead>
<tbody>
<tr>
<td>Google Inc.</td>
<td>NASDAQ:GOOG</td>
<td>Information</td>
</tr>
<tr>
<td>Starbucks Corporation</td>
<td>NASDAQ:SBUX</td>
<td>Consumer</td>
</tr>
<tr>
<td>JPMorgan Chase &amp; Co.</td>
<td>NYSE:JPM</td>
<td>Financials</td>
</tr>
<tr>
<td>Goldman Sachs Group Inc.</td>
<td>NYSE:GS</td>
<td>Financials</td>
</tr>
<tr>
<td>Macy’s Inc.</td>
<td>NYSE:M</td>
<td>Consumer</td>
</tr>
<tr>
<td>McAfee Inc.</td>
<td>NYSE:MFE</td>
<td>Technology</td>
</tr>
<tr>
<td>Merck &amp; Co.</td>
<td>NYSE:MRK</td>
<td>Health Care</td>
</tr>
<tr>
<td>Pfizer Inc.</td>
<td>NYSE:PFN</td>
<td>Health Care</td>
</tr>
<tr>
<td>Bank of America Corporation</td>
<td>NYSE:BAC</td>
<td>Financials</td>
</tr>
<tr>
<td>Biogen Idec Inc.</td>
<td>NASDAQ:BIIB</td>
<td>Health Care</td>
</tr>
<tr>
<td>Boston Scientific Corporation</td>
<td>NASDAQ:BSX</td>
<td>Health Care</td>
</tr>
<tr>
<td>International Business Machines Corporation (IBM)</td>
<td>NYSE:IBM</td>
<td>Information</td>
</tr>
<tr>
<td>Sunoco Inc.</td>
<td>NYSE:SUN</td>
<td>Energy</td>
</tr>
<tr>
<td>Exxon Mobile Corporation</td>
<td>NYSE:XOM</td>
<td>Energy</td>
</tr>
<tr>
<td>Costco Wholesale Corporation</td>
<td>NASDAQ:COST</td>
<td>Consumer Staples</td>
</tr>
<tr>
<td>eBay Inc.</td>
<td>NASDAQ:EBAY</td>
<td>Information</td>
</tr>
<tr>
<td>Apple Inc.</td>
<td>NASDAQ:AAPL</td>
<td>Technology</td>
</tr>
<tr>
<td>AT&amp;T Inc.</td>
<td>NYSE:T</td>
<td>Telecommunications Services</td>
</tr>
<tr>
<td>Oracle Corporation</td>
<td>NASDAQ:ORCL</td>
<td>Information</td>
</tr>
<tr>
<td>Whole Foods Market Inc.</td>
<td>NASDAQ:WFMI</td>
<td>Consumer Staples</td>
</tr>
</tbody>
</table>

---

**Fig. 5.1.** Evolution of the S&P500 discounted return rates in the years 2005–2010.
Fig. 5.2. Robust CVaR of optimal-CVaR and optimal-Robust CVaR \((\gamma = \delta)\) portfolios.

Fig. 5.3. Robust CVaR of optimal-CVaR and optimal-Robust CVaR (fixed \(\gamma\)) portfolios.

Fig. 5.4. Robust expected return of various portfolios.
We also compute the optimal-Robust Expected Return portfolios, that is the portfolios that maximize the worst-case expected return, for various values of $\gamma$. In Fig. 5.4, we observe the worst-case expected return of various portfolios with fixed or varying $\gamma$. The performance of all portfolios drops as $\delta$ increases. The optimal-Robust Expected Return portfolio with $\gamma = \delta$ shows the best performance because the performance metric is the worst-case expected return, and it knows the exact size of the uncertainty. However, we see negligible differences among the optimal-CVaR and the optimal-Robust CVaR portfolios. This implies that adding another layer of robustness by using the Robust CVaR does not affect the expected performance in this case.

6. Conclusions

A way to deal with uncertainty in the probability distribution estimation is to apply robust optimization principles to the risk measures that would be used if the probability distribution was exactly known. In this paper, convex optimization techniques have shown that such a procedure keeps a risk measure convex or coherent, which strengthens the importance of robust optimization in risk management to deal with probability distribution uncertainty. Moreover, convex duality theory in infinite dimensional spaces enabled the use of robust optimization to deal with a special case of two-stage probability distribution uncertainty under a CVaR framework. Our result showed that not only CVaR, but also special robust risk measures defined based on it, can be used in financial decision making. Our result showed that not only CVaR, but also special robust risk measures defined based on it, can be used in financial decision making.

Acknowledgements

The authors thank Professor Alexander Shapiro of Georgia Institute of Technology, and Professor Arkadi Nemirovski of Georgia Institute of Technology and Technion – Israel Institute of Technology for the interesting discussions they had with them on the topic and their useful comments. They would also like to thank an anonymous reviewer for her or his comments.

Appendix A. Mathematical background

In this section, we introduce some useful mathematical notation and background used throughout the paper.

Set $\mathbb{R}$ contains all the extended real numbers, namely $\mathbb{R} = \mathbb{R} \cup (-\infty) \cup (+\infty)$. For $n \in \mathbb{N}, n \geq 1$, cone $\mathbb{R}_n^+$ is the positive orthant of space $\mathbb{R}^n$. If $V$ is a topological vector space equipped with a certain topology, and $A \subseteq V$, functions $c_1(A)$, and $c_2(A)$ return the closure, and the convex hull of $A$, respectively. The indicator function $I_A(a)$ on a set $A \subseteq V$ is defined by

$$I_A(a) = \begin{cases} 0, & \text{if } a \in A, \\ +\infty, & \text{if } a \notin A. \end{cases} \quad (A.1)$$

If $(\Omega, \mathcal{F})$ is a measurable space, and $P_1, P_2$ are two probability measures on it, $P_1 \ll P_2$ denotes that $P_1$ is absolutely continuous with reference to $P_2$. If $P_1 \ll P_2$, then $dP_1/dP_2$ denotes the Radon-Nikodym derivative of $P_1$ with reference to $P_2$. For $n \geq 1$, and $1 \leq p \leq +\infty$, if $x \in \mathbb{R}^n$, then $|x|_p$ is the $l_p$-norm of $x$. For $1 \leq p \leq +\infty$, if $X$ is a random variable in probability space $(\Omega, \mathcal{F}, P)$, and $X \in L^p(\Omega, \mathcal{F}, P)$, then $|X|_p$ is the $L_p$-norm of $X$.

We present the definitions of a proper mapping, a lower semicontinuous mapping, function lsc, and function cv which are going to be used in the representation theorems (Shapiro et al., 2009; Rockafellar, 1974). Note that the definitions depend on the topology that is assumed on $V$.

Definition A.1. Consider a topological vector space $V$. A mapping $\mu : V \to \mathbb{R}$ is called proper, if for any $a \in V$, $\mu(a) > -\infty$, and there exists an $a \in V$ such that $\mu(a) \in \mathbb{R}$.

Definition A.2. Consider a topological vector space $V$. A mapping $\mu : V \to \mathbb{R}$ is called lower semicontinuous, if its epigraph $\text{epi}(\mu) \equiv \{(a,x) \in V \times \mathbb{R}| x \geq \mu(a), a \in V, x \in \mathbb{R}\}$ is closed in the product topology defined on space $V \times \mathbb{R}$.

Definition A.3. Consider a topological vector space $V$, the product topology on space $V \times \mathbb{R}$, a mapping $\mu : V \to \mathbb{R}$, and its epigraph $\text{epi}(\mu)$. Then, $\text{lsc}(\mu) : V \to \mathbb{R}$ is the unique lower semicontinuous mapping whose epigraph is $\text{cl}(\text{epi}(\mu))$.

Definition A.4. Consider a topological vector space $V$, the product topology on space $V \times \mathbb{R}$, a mapping $\mu : V \to \mathbb{R}$, and its epigraph $\text{epi}(\mu)$. Then, $\text{cv}(\mu) : V \to \mathbb{R}$ is the unique convex mapping whose epigraph is $\text{conv}(\text{epi}(\mu))$.

Definition A.5. Consider a pair of dual topological vector spaces $V$, $V^*$, and a mapping $\mu : V \to \mathbb{R}$. If $(\xi, \zeta)$ is the bilinear form which places in duality the two spaces, the conjugate mapping $\mu^* : V^* \to \mathbb{R}$ of $\mu$ is defined by

$$\mu^*(\zeta) = \sup_{x \in V} \langle \zeta, x \rangle - \mu(x), \zeta \in V^*. \quad (A.2)$$

The following proposition is a direct consequence of the Fenchel-Moreau theorem (Shapiro et al., 2009; Ruszczyski and Shapiro, 2006).

Proposition A.1. Consider the framework of Definition A.5. If $\nu : V^* \to \mathbb{R}$ is the conjugate mapping of proper mapping $\mu : V \to \mathbb{R}$, then it is also the conjugate mapping of lsc $(\text{cv}(\mu))$.

Definition A.6. A set $B$ is called a base of cone $C$ if for any $x \in C$ with $x \neq 0$ there exist unique $i > 0$, and $u \in B$, such that $x = iu$.

We define the generalized Lorentz cone, and derive its dual.

Definition A.7. Consider a norm $\| \cdot \|$ in space $\mathbb{R}^n$. The generalized Lorentz cone in space $\mathbb{R}^{n+1}$ with reference to norm $\| \cdot \|$ is

$$K_1 \equiv \{ (x^T, x_{n+1})^T \in \mathbb{R}^{n+1}| x_{n+1} \geq \|x\| \}. \quad (A.3)$$

The following proposition computes the dual of a generalized Lorentz cone.

Proposition A.2. If $\| \cdot \|$ is a norm in $\mathbb{R}^n$, and $\| \cdot \|$ is its dual norm, the dual of the generalized Lorentz cone in $\mathbb{R}^{n+1}$ with reference to norm $\| \cdot \|$ is the generalized Lorentz cone in $\mathbb{R}^{n+1}$ with reference to norm $\| \cdot \|$.

Proof. According to the definition of the dual cone (Barvinok, 2002, p. 162), to prove the proposition, it suffices to prove that for any $(y', y_{n+1})^T \in \mathbb{R}^{n+1}$:

- If for any $(x^T, x_{n+1})^T \in \mathbb{R}^{n+1}$ with $x_{n+1} > \|x\|, x'y + x_{n+1}y_{n+1} > 0$, then, $y_{n+1} \geq \|y\|$.
- If $y_{n+1} \geq \|y\|$ then, for any $(x^T, x_{n+1})^T \in \mathbb{R}^{n+1}$ with $x_{n+1} > \|x\|$, $x'y + x_{n+1}y_{n+1} > 0$.

Fix any $(y', y_{n+1})^T \in \mathbb{R}^{n+1}$. Suppose that for any $(x^T, x_{n+1})^T \in \mathbb{R}^{n+1}$ with $x_{n+1} \geq \|x\|, x'y + x_{n+1}y_{n+1} \geq 0$. Take any $x \in \mathbb{R}^n$ with $\|x\| = 1$. Then, by Hölder’s inequality (Boyd and Vandenberghe, 2004, p. 78).
\[ \langle x, y \rangle \leq |x| \cdot |y|, \]

and

\[ y_{n+1} = 1 + y_{n+1} - \langle x, y \rangle \leq |x| \cdot |y|. \]

Now, we will prove the opposite direction. Assume that

\[ y_{n+1} \leq |y|, \]

and take any \((x', x_{n+1}) \in \mathbb{R}^{n+1} \) with \( x_{n+1} \geq |x| \). Then, by Hölder’s inequality \((\text{Boyd and Vandenberghe, 2004, p. 78})

\[ -\langle x', y \rangle \leq |x| \cdot |y|, \leq x_{n+1}y_{n+1}, \]

which concludes the proof. \( \square \)

References


