Math. Program., Ser. B 108, 339–354 (2006)

Digital Object Identifier (DOI) 10.1007/s10107-006-0714-8

Georg Ch. Pflug

Subdifferential representations of risk measures

Received: August 23, 2004 / Accepted: October 5, 2005
Published online: June 2, 2006 – © Springer-Verlag 2006

Abstract. Measures of risk appear in two categories: Risk capital measures serve to determine the necessary amount of risk capital in order to avoid ruin if the outcomes of an economic activity are uncertain and their negative values may be interpreted as acceptability measures (safety measures). Pure risk measures (risk deviation measures) are natural generalizations of the standard deviation. While pure risk measures are typically convex, acceptability measures are typically concave. In both cases, the convexity (concavity) implies under mild conditions the existence of subgradients (supergradients). The present paper investigates the relation between the subgradient (supergradient) representation and the properties of the corresponding risk measures. In particular, we show how monotonicity properties are reflected by the subgradient representation. Once the subgradient (supergradient) representation has been established, it is extremely easy to derive these monotonicity properties. We give a list of Examples.

Key words. Risk measures – Duality – Stochastic dominance

1. Introduction

In recent years, starting from the seminal paper by Artzner et al. [1], axiomatic approaches to the definition of appropriate measures of risk for random variables and stochastic processes have been in the center of interest of many authors ([5, 8, 7, 9, 15, 17]). It is common sense, that convexity (concavity) plays a key role among the required properties for risk measures. A convex lower semicontinuous function is characterized by the fact that it is the dual of its own dual, hence completely characterized by its dual function. A concave function is characterized by the fact that its negative is convex.

Concavity (convexity) of risk functionals has been recently investigated by Ruszczyński and Shapiro [19]. They show the continuity and super(sub)-differentiability of risk functionals under mild conditions, i.e. the existence of dual representations. Moreover they investigate the dual structure of optimization problems involving super(sub)differentiable risk functionals. For positive homogeneous risk deviation measures (see below), Rockafellar et al. [18] have shown the existence dual representations and characterize the subgradient set, calling it the risk envelope.

In this paper, we show how for super(sub)differentiable risk functionals, the dual representation can be used to derive some properties (in particular monotonicity) of the risk functional in a very simple manner. Moreover we give many examples of dual representations of well known risk functionals.

G. Ch. Pflug: Department of Statistics and Decision Support Systems, Universitätsstrasse 5,
University of Vienna, 1090 Wien-Vienna, Austria.
e-mail: georg.pflug@univie.ac.at

Mathematics Subject Classification (2000): 91B82, 91B28, 90C46

We begin with basic definitions. They are in the spirit of coherence properties introduced in [1] and further developed in [17].

Acceptability Functionals. A mapping $\mathcal{A}$ defined on a set of real valued random variables on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is called *acceptability functional* (safety functional), if it exhibits the following properties:

(A1) **Translation equivariance.**

$$\mathcal{A}(Y + c) = \mathcal{A}(Y) + c$$

for constant c.

(A2) **Strictness.**

$$\mathcal{A}(Y) \leq \mathbb{E}(Y),$$

where the equality sign holds iff Y is a constant.

(A3) **Concavity.**

$$\mathcal{A}(\lambda Y_1 + (1 - \lambda)Y_2) \geq \lambda \mathcal{A}(Y_1) + (1 - \lambda)\mathcal{A}(Y_2),$$

for $0 \leq \lambda \leq 1$.

Besides the basic properties (A1) - (A3), the acceptability functional may also exhibit the following properties

(A4) **Positive homogeneity.**

$$\mathcal{A}(\lambda Y) = \lambda \mathcal{A}(Y),$$

for $\lambda > 0$.

(A5) **monotonicity w.r.t. first order stochastic dominance.**
Y_1 is dominated w.r.t. first stochastic order by Y_2 (in symbol: $Y_1 \prec_{FSD} Y_2$), if $\mathbb{E}(U(Y_1)) \leq \mathbb{E}(U(Y_2))$ for all monotonic integrable utility functions U. The acceptability measure $\mathcal{A}$ is called monotonic w.r.t the first order stochastic dominance, if $Y_1 \prec_{FSD} Y_2$ implies that

$$\mathcal{A}(Y_1) \leq \mathcal{A}(Y_2).$$

(A6) **monotonicity w.r.t. second order stochastic dominance.**
Y_1 is dominated w.r.t. second stochastic order by Y_2 (in symbol: $Y_1 \prec_{SSD} Y_2$), if $\mathbb{E}(U(Y_1)) \leq \mathbb{E}(U(Y_2))$ for all monotonic and concave integrable utility functions U. The acceptability measure $\mathcal{A}$ is called monotonic w.r.t the second order stochastic dominance, if $Y_1 \prec_{SSD} Y_2$ implies that

$$\mathcal{A}(Y_1) \leq \mathcal{A}(Y_2).$$

Properties (A5) and (A6) link the more classical utility approaches to geometric properties of the functionals [2], [11]. Notice that (A6) implies (A5).

Remark. Properties (A5) and (A6) depend only on the distribution. The weaker pointwise versions of (A5) and (A6) read

(A5') **monotonicity w.r.t. pointwise ordering.**
The acceptability measure $\mathcal{A}$ is called monotonic w.r.t pointwise ordering, if $Y_1 \leq Y_2$ implies that

$$\mathcal{A}(Y_1) \leq \mathcal{A}(Y_2).$$

(A6') **monotonicity w.r.t. reverse supermartingale ordering.**
The acceptability measure $\mathcal{A}$ is called monotonic w.r.t. reverse supermartingale ordering, if $\mathbb{E}(Y_1|Y_2) \leq Y_2$ implies that

$$\mathcal{A}(Y_1) \leq \mathcal{A}(Y_2).$$

The name comes from the fact that (Y_1, Y_2) is a reversed supermartingale (see [14] p. 115).

Obviously (A5) implies (A5') and (A6) implies (A6'). A kind of reverse statement is given in Proposition 1 below.

Artzner et al. [1] introduced the notion of a *coherent risk measure*. A risk measure ρ is coherent in their sense, if it is of the form $\rho(Y) = -\mathcal{A}(Y)$, where $\mathcal{A}$ is an acceptability functional, satisfying in addition (A4) and (A5').

An acceptability functional $\mathcal{A}$, which is continuous w.r.t. convergence in probability and satisfies (A4) and (A5') has a representation

$$\mathcal{A}(Y) = \inf\{\mathbb{E}(Y\,Z) : Z \in \mathcal{Z}\} \tag{1}$$

where $\mathcal{Z}$ is a set of probability densities containing the constant density 1, as was shown by Delbaen [5]. Obviously, the converse holds also true, i.e. every functional of the form (1) satisfies (A1)–(A4), (A5').

More generally, any acceptability functional $\mathcal{A}$ which is continuous w.r.t. convergence in probability has a representation as the infimum of linear functions

$$\mathcal{A}(Y) = \inf\{\mathbb{E}(Y\,Z) + \alpha(Z) : Z \in \mathcal{Z}\}, \tag{2}$$

where $\mathcal{Z}$ is a set of (signed) densities. Such representations were introduced by Föllmer and Schied [8]. We call a representation of the form (2) a *superdifferential representation*. The reason is that if $\mathcal{A}(Y^*) = \mathbb{E}(Y^*\,Z^*) + \alpha(Z^*)$, then for all Y

$$\mathcal{A}(Y) \leq \mathcal{A}(Y^*) + \mathbb{E}[Z^*(Y - Y^*)],$$

i.e. Z^* is a superdifferential of $\mathcal{A}$ at Y^*. It is easy to see that conversely every functional of the form (2) satisfies (A1) and (A3). It satisfies (A2) if the constant $\mathbb{1} \in \mathcal{Z}$ and $\alpha(\mathbb{1}) = 0$.

In this paper, we will exclusively consider functionals, which depend on the distribution only, meaning that we will use (A5) and (A6) and not (A5') and (A6'). A characterization of such functionals was given by Kusuoka [12].

Pure Risk Functionals. A mapping $\mathcal{D}$ from set of real valued random variables on $(\Omega, \mathcal{F}, P)$ is called *pure risk functional* or *deviation functional*, if it exhibits the following properties

(D1) **Translation invariance.**

$$\mathcal{D}(Y+c) = \mathcal{D}(Y)$$

for constant c.

(D2) **Strictness.**

$$\mathcal{D}(Y) \geq 0,$$

where the equality sign holds iff Y is a constant.

(D3) **Convexity.**

$$\mathcal{D}(\lambda Y_1 + (1-\lambda)Y_2) \leq \lambda \mathcal{D}(Y_1) + (1-\lambda)\mathcal{D}(Y_2),$$

for $0 \leq \lambda \leq 1$.

Besides the basic properties (D1) - (D3), a pure risk functional may also exhibit the following properties

(D4) **Positive homogeneity.**

$$\mathcal{D}(\lambda Y) = \lambda \mathcal{D}(Y),$$

for $\lambda \geq 0$.

(D5) **monotonicity w.r.t. convex dominance.**
Y_1 is dominated w.r.t. convex dominance (in symbol: $Y_1 \prec_{CXD} Y_2$), if $\mathbb{E}(U(Y_1)) \leq \mathbb{E}(U(Y_2))$ for all convex utility functions U. The pure risk measure $\mathcal{D}$ is called monotonic w.r.t. convex dominance, if $Y_1 \prec_{CXD} Y_2$ implies that

$$\mathcal{D}(Y_1) \leq \mathcal{D}(Y_2).$$

Pure risk functionals serve as the generalization of the variance in risk-return settings [4]. The notion of deviation functionals was introduced by Rockafellar and Uryasev [17]. It is obvious that $\mathcal{D}$ is a deviation functional, if and only if $\mathbb{E}(Y) - \mathcal{D}(Y)$ is an acceptability functional. To each acceptability functional $\mathcal{A}$ we may associate the pertaining pure risk functional $\mathcal{D}$ by $\mathcal{D}(Y) := \mathbb{E}(Y) - \mathcal{A}(Y)$ and conversely, to each pure risk functional $\mathcal{D}$ we may associate the pertaining acceptability functional by $\mathcal{A}(Y) = \mathbb{E}(Y) - \mathcal{D}(Y)$. $\mathcal{A}$ fulfills (Ai) iff the pertaining $\mathcal{D}$ fulfills (Di), where i = 1, 2, 3, 4.

A representation of $\mathcal{D}$ of the form

$$\mathcal{D}(Y) = \sup\{\mathbb{E}(Y\,Z) - \beta(Z) : Z \in \mathcal{Z}\} \tag{3}$$

is called a *subdifferential representation* of $\mathcal{D}$. Subdifferential representations of $\mathcal{D}$ and superdifferential representations of the pertaining $\mathcal{A}$ are just the reverse sides of the same coin: Suppose that $\mathcal{D}$ has the representation (3). Then $\mathcal{A}(Y) = \mathbb{E}Y - \mathcal{D}(Y)$ has the representation

$$\mathcal{A}(Y) = \inf\{\mathbb{E}(Y\,Z) + \beta(1-Z) : 1-Z \in \mathcal{Z}\} \tag{4}$$

as can be easily seen.

The paper is organized as follows: In the next section we formulate the main result and its proof. In section 3, we have collected a list of Examples, for which the dual representations are calculated explicitly. The Appendix contain auxiliary results.

2. The main result

Our main result relates the properties of the acceptability and pure risk functionals to the properties of the subdifferential representation. Our approach is similar to the one used in Ruszczynski and Shapiro [19] however, we do not require a priori that $-\alpha$ is the conjugate dual of $\mathcal{A}$. Also our dual objects Z lie in function spaces and not in spaces of measures.

The Main theorem. *Suppose that $\mathcal{A}$ has the representation (2). Suppose further that the probability space is not atomic. Then*

- *(i) $\mathcal{A}$ depends only on the distribution, if the set $\mathcal{Z}$ is determined by distribution only, i.e. if Z and $\tilde{Z}$ have the same distribution and $Z \in \mathcal{Z}$ then also $\tilde{Z} \in \mathcal{Z}$ and $\alpha(Z) = \alpha(\tilde{Z})$.*
- *(ii) $\mathcal{A}$ is positively homogeneous if $\alpha = 0$.*
- *(iii) $\mathcal{A}$ is monotonic w.r.t. first order stochastic dominance, if $\mathcal{Z}$ contains only nonnegative random variables.*
- *(iv) $\mathcal{A}$ is monotonic w.r.t. second order stochastic dominance, if $\mathcal{Z}$ contains only nonnegative densities, is stable w.r.t. conditional expectations (i.e. $Z \in \mathcal{Z}$ implies that $\mathbb{E}(Z|\mathcal{F}) \in \mathcal{Z}$) and α is monotonic w.r.t. conditional expectations (i.e. $\alpha(\mathbb{E}(Z|\mathcal{F})) \le \alpha(Z)$) for all Z and any σ-algebra $\mathcal{F}$.*

Suppose that $\mathcal{D}$ has the representation (3). Then

- *(v) $\mathcal{D}$ is positively homogeneous if $\beta = 0$.*
- *(vi) $\mathcal{D}$ is monotonic w.r.t. convex dominance, if $\mathcal{Z}$ is stable w.r.t. conditional expectations and β is monotonic w.r.t. conditional expectation.*

To prove the main result, we evoke first the following characterization of first resp. second order stochastic dominance in terms of coupling constructions.

Proposition 1. *(i) The FSD-coupling: If $Y_1 \prec_{FSD} Y_2$, then one may construct a pair $\tilde{Y}_1, \tilde{Y}_2$ of random variables with the same marginal distributions as Y_1, Y_2, such that*

$$\tilde{Y}_1 \le \tilde{Y}_2 \qquad a.s.$$

(ii) The CXD-coupling: If $Y_1 \prec_{CXD} Y_2$, then one may construct a pair $\tilde{Y}_1, \tilde{Y}_2$ of random variables with the same marginal distributions as Y_1, Y_2, such that

$$\tilde{Y}_1 = \mathbb{E}(\tilde{Y}_2|\tilde{Y}_1) \qquad a.s.$$

(iii) The SSD-coupling. If $Y_1 \prec_{SSD} Y_2$, then one may construct a pair $\tilde{Y}_1, \tilde{Y}_2$ of random variables with the same marginal distributions as Y_1, Y_2, such that

$$\mathbb{E}(\tilde{Y}_1|\tilde{Y}_2) \le \tilde{Y}_2 \qquad a.s.$$

Proof. See Strassen [20]. □

Proof of the Main Theorem.

(i) Let Y have the same distribution as $\tilde{Y}$. We have to show that under the given assumptions $\mathcal{A}(Y) = \mathcal{A}(\tilde{Y})$. Let $Z_n \in \mathcal{Z}$ be a sequence satisfying

$$\mathcal{A}(Y) = \lim_n [\mathbb{E}(Y\, Z_n) - \alpha(Z_n)].$$

We show that we may construct for every Z_n a $\tilde{Z}_n$ such that Z_n and $\tilde{Z}_n$ have the same law and

$$\mathbb{E}(Y\, Z_n) = \mathbb{E}(\tilde{Y}\, \tilde{Z}_n). \tag{5}$$

If $F_n(z|y)$ is the conditional distribution of Z_n given Y, then construct (possibly by extending the probability space) a random variable $\tilde{Z}_n$, such that $\tilde{Z}_n$ has the distribution $F_n(\tilde{z}|\tilde{y})$ given $\tilde{Y} = \tilde{y}$. Then Z_n and $\tilde{Z}_n$ have the same distribution, (5) holds and $\tilde{Z}_n \in \mathcal{Z}$. This implies that

$$\mathcal{A}(\tilde{Y}) = \inf\{\mathbb{E}(\tilde{Y}\, Z) + \alpha(Z) : Z \in \mathcal{Z}\} \le \lim_n [\mathbb{E}(\tilde{Y}\, \tilde{Z}_n) - \alpha(\tilde{Z}_n)] = \mathcal{A}(Y)$$

and by symmetricity $\mathcal{A}(\tilde{Y}) = \mathcal{A}(Y)$.

(ii) is obvious.

(iii) Suppose that $\mathcal{A}(Y) = \inf\{\mathbb{E}(Y\, Z) + \alpha(Z) : Z \in \mathcal{Z}\}$ and $\mathcal{A}$ depends only on the distribution of Y. If $Y_1 \prec_{FSD} Y_2$ one may construct by Proposition 1(i) versions $(\tilde{Y}_1, \tilde{Y}_2)$ of the random variables (Y_1, Y_2) that satisfy $\tilde{Y}_1 \le \tilde{Y}_2$ a.s. Equivalently we may assume from the beginning that $Y_1 \le Y_2$ a.s. If $\mathcal{Z}$ contains only nonnegative elements, then $Y_1 \le Y_2$ a.s. implies that $\mathbb{E}(Y_1 Z) + \alpha(Z) \le \mathbb{E}(Y_2 Z) + \alpha(Z)$ for all $Z \in \mathcal{Z}$ and hence $\mathcal{A}(Y_1) = \inf\{\mathbb{E}[Y_1\, Z] + \alpha(Z) : Z \in \mathcal{Z}\} \le \mathcal{A}(Y_2) = \inf\{\mathbb{E}[Y_2\, Z] + \alpha(Z) : Z \in \mathcal{Z}\}$.

(iv) Suppose that $\mathcal{A}(Y) = \inf\{\mathbb{E}(Y\, Z) : Z \in \mathcal{Z}\}$. If $Y_1 \prec_{SSD} Y_2$ one may assume by Proposition 7 (ii) that $\mathbb{E}(Y_1|Y_2) \le Y_2$. If $\mathcal{Z}$ contains only nonnegative elements and is stable w.r.t. conditional expectations, then $\mathbb{E}[Y_1 \mathbb{E}(Z|Y_2)] = \mathbb{E}[\mathbb{E}(Y_1|Y_2)Z] \le \mathbb{E}[Y_2 Z]$. Thus

$$\begin{aligned}
\mathcal{A}(Y_1) &= \inf\{\mathbb{E}[Y_1 Z] + \alpha(Z) : Z \in \mathcal{Z}\} \\
&\le \inf\{\mathbb{E}[Y_1 \mathbb{E}(Z|Y_2)] + \alpha(\mathbb{E}(Z|Y_2)) : Z \in \mathcal{Z}\} \\
&\le \inf\{\mathbb{E}[\mathbb{E}(Y_1|Y_2)Z] + \alpha(Z) : Z \in \mathcal{Z}\} \\
&\le \inf\{\mathbb{E}[Y_2 Z] + \alpha(Z) : Z \in \mathcal{Z}\} = \mathcal{A}(Y_2).
\end{aligned}$$

(v) Suppose that $\mathcal{D}(Y) = \sup\{\mathbb{E}(Y\, Z) - \beta(Z) : Z \in \mathcal{Z}\}$. If $Y_1 \prec_{CXD} Y_2$ one may assume by Proposition 7 (ii) that $Y_1 = \mathbb{E}(Y_2|Y_1)$. Since $\mathbb{E}(Y_1 Z) = \mathbb{E}(\mathbb{E}(Y_2|Y_1)Z) = \mathbb{E}(Y_2 \mathbb{E}(Z|Y_1))$ and since $\mathcal{Z}$ is stable and β is monotonic w.r.t. conditional expectations, we have

$$\begin{aligned}
\mathcal{D}(Y_1) &= \sup\{\mathbb{E}[Y_1 Z] - \beta(Z) : Z \in \mathcal{Z}\} = \sup\{\mathbb{E}[Y_2 \mathbb{E}(Z|Y_1)] - \beta(Z) : Z \in \mathcal{Z}\} \\
&\le \sup\{\mathbb{E}[Y_2 \mathbb{E}(Z|Y_1)] - \beta(\mathbb{E}(Z|Y_1)) : Z \in \mathcal{Z}\} \\
&\le \sup\{\mathbb{E}(Y_2 Z) - \beta(Z) : Z \in \mathcal{Z}\} = \mathcal{D}(Y_2).
\end{aligned}$$

□

The main theorem gives sufficient conditions for monotonicity conditions to hold. Are they also necessary? To answer this question one has to realize that a functional $\mathcal{A}$ of the form (2) does not determine the set $\mathcal{Z}$ and the function α in a unique manner. Notice first that one may get rid of explicitely handling the set $\mathcal{Z}$ by setting $\alpha(Z) = \infty$ for $Z \notin \mathcal{Z}$. With this extension, one may w.l.o.g. assume that α is defined on a full topological linear space of functions. Simply spoken, two functions α_1 and α_2 generate the same $\mathcal{A}$, if both functions have the same l.s.c. (lower semicontinouos) convex minorant.

This result is stated as Theorem 1 in [19]. For convenience, we give a short proof here: Let $\bar{\alpha}(Z)$ be the l.s.c. convex minorant of α. The epigraph of $\bar{\alpha}$ is the closed convex hull of the epigraph of α. We claim that

$$\mathcal{A}(Y) = \inf\{\mathbb{E}(Y\,Z) + \alpha(Z) : Z \in \mathcal{Y}^*\} = \inf\{\mathbb{E}(Y\,Z) + \bar{\alpha}(Z) : Z \in \mathcal{Y}^*\}. \qquad (6)$$

In order to prove (6), let $\bar{\mathcal{A}}(Y) = \inf\{\mathbb{E}(Y\,Z)+\bar{\alpha}(Z) : Z \in \mathcal{Y}^*\}$. Trivially $\bar{\mathcal{A}}(Y) \le \mathcal{A}(Y)$. For the opposite inequality consider the space $\mathcal{Y}^* \times \mathbb{R}$ with dual $\mathcal{Y} \times \mathbb{R}$ and inner product $\langle (Z, a), (Y, b)\rangle = \mathbb{E}(Y\,Z) + a \cdot b$. Let $\mathcal{S} = \{(Z, a) : a \ge \alpha(Z)\}$ and let $\bar{\mathcal{S}}$ its closed convex hull, i.e. $\bar{\mathcal{S}} = \{(Z, a) : a \ge \bar{\alpha}(Z)\}$. Notice that $\langle (Z, 1), (Y, a)\rangle \ge \mathcal{A}(Y)$ for all $(Z, a) \in \mathcal{S}$. Then the same inequality is true for the convex hull of $\mathcal{S}$, i.e. $\mathbb{E}(Y\,Z) + \bar{\alpha}(Z) \ge \mathcal{A}(Y)$ for all $Z \in \mathcal{Y}^*$. This implies that $\bar{\mathcal{A}}(Y) \ge \mathcal{A}(Y)$.

The next proposition gives a necessary conditions under the assumption that α is convex l.s.c.

Proposition 2. *Assume that $\mathcal{A}(Y) = \inf\{\mathbb{E}(Y\,Z) + \alpha(Z)\}$, where α is a proper convex function on $\mathcal{Y}^*$ and $\mathcal{A} < \infty$ is upper semicontinuous . Then*

(i) If $\mathcal{A}(Y)$ is pointwise monotone, then for $Z \in \mathcal{Y}^*$ with $P\{Z < 0\} > 0, \alpha(Z) = \infty$.
(ii) If $\mathcal{A}$ is antitonic w.r.t. conditional expectation, i.e. $\mathcal{A}(\mathbb{E}(Y|\mathcal{F})) \ge \mathcal{A}(Y)$, then α is monotonic w.r.t. conditional expectation, i.e. $\alpha(\mathbb{E}(Z|\mathcal{F})) \le \alpha(Z)$ for all $Z \in \mathcal{Y}^*$.

Proof. By the Fenchel-Moreau theorem (see [3], p. 78), modified for concave functions in obvious manner, we have that for

$$\mathcal{A}^*(Z) = \inf\{\mathbb{E}(Z\,V) - \mathcal{A}(V) : V \in \mathcal{Y}\} \qquad (7)$$

$\mathcal{A}^*(Z) = -\alpha(Z)$ and that the solution set of (7) is nonempty for all Z.

(i) Suppose that $\mathcal{A}(Y)$ is pointwise monotone. If there is a Z with $P\{Z < 0\} > 0$ and $\alpha(Z) < \infty$, then by assumption there is a $V \in \mathcal{Y}$ such that $\mathcal{A}^*(Z) = \mathbb{E}(V\,Z) - \mathcal{A}(V)$. Thus for all $Y \in \mathcal{Y}, \mathbb{E}(V\,Z) - \mathcal{A}(V) \le \mathbb{E}(Y\,Z) - \mathcal{A}(Y)$. Let now $Y = V + \mathbb{1}_{\{Z<0\}}$. Then $Y \ge V$, but $\mathcal{A}(Y)-\mathcal{A}(V) \le EE((Y-V)Z) = -P\{Z < 0\} < 0$, a contradiction.
(ii) Assume that $\mathcal{A}$ is antitonic w.r.t. conditional expectation, i.e. $\mathcal{A}(\mathbb{E}(Y|\mathcal{F})) \ge \mathcal{A}(Y)$. Then, by our main theorem, $\mathcal{A}^*$ given by (7) is also antitonic w.r.t. conditional expectation, thus it is necessary that $\alpha(Z) = -\mathcal{A}^*(Z)$ is monotonic w.r.t. conditional expectation. □

Example: Distortion risk mesures. Consider the following functional

$$\mathcal{A}(Y) = \mathbb{E}[Yk(1 - F_Y(Y))],$$

where k is nonnegative, strictly monotone and continuous function on [0, 1] and $F_Y(u) = P\{Y \le u\}$ is the distribution function of Y. Such functionals were introduced by Denneberg [6] under the name of "distorted probabilities" and studied in detail by De Giorgi [4] under the name of "Choquet integrals". He proves that under the given assumptions on k, $\mathcal{A}$ is concave and monotonic w.r.t.$\prec_{SSD}$. It is easy to see that $\mathcal{A}$ has the representation

$$\mathcal{A}(Y) = \inf\{\mathbb{E}(Y\,Z) : Z = k(U), \text{ where } U \text{ is uniformly [0,1] distributed}\}, \tag{8}$$

i.e. $\mathcal{A} = \inf\{\mathbb{E}(Y\ Z) : Z \in \mathcal{Z}\}$, where $\mathcal{Z} = \{Z : k^{-1}(Z)$ is uniformly distributed in $[0, 1]\}$. (8) follows from a known result by Hoeffding (see Lehmann [13]): If (Y, Z) have joint distribution $H(y, z)$ with marginals $F(y)$ and $G(z)$, then

$$\mathbb{C}\text{ov}(Y, Z) = \int\int H(y, z) - F(y) \cdot G(z)\, dy\, dz.$$

Since $H(y, z) \ge \max(F(y)+G(z)-1, 0)$, we have that for fixed marginals, $\mathbb{C}\text{ov}(Y, Z)$ and hence $\mathbb{E}(Y\,Z)$ is minimized, if H achieves the lower Fréchet bound $H(y, z) = \max(F(y) + G(z) - 1, 0)$, i.e. if Y and Z are antimonotone, i.e. $Y = F^{-1}(U)$, $Z = G^{-1}(1 - U)$, for a uniform [0,1] U. In our case, $G^{-1} = k$ and therefore $Z = k(1 - F_Y(Y))$.

Since $\mathcal{Z}$ consists only of nonnegative variables, $\mathcal{A}$ is monotonic w.r.t. $\prec_{FSD}$ and positively homogeneous. $\mathcal{Z}$ is however not convex. Assume that $k(Z^*)$ has uniform [0,1] distribution and let $\bar{\mathcal{Z}} = \{Z : Z \prec_{CXD} Z^*\}$. Then $\bar{\mathcal{Z}}$ is the closed convex hull of $\mathcal{Z}$. Since $\bar{\mathcal{Z}}$ is stable w.r.t. conditional expectations, $\mathcal{A}$ is monotonic w.r.t. $\prec_{SSD}$, which is in accordance with the findings of De Giorgi [4].

3. Examples

In this section, we list some representations of pure risk (deviation) and acceptability measures and associate monotonicity properties with the representation.

3.1. The variance and similar pure risk measures

For a random variable V on $(\Omega, \mathcal{F}, P)$ let $\|V\|_p = \mathbb{E}^{1/p}[|V|^p]$, if it exists. We consider pure risk measures of the form $\mathcal{D}(Y) = \|Y - \mathbb{E}Y\|_p^p$. Setting $p = 2$ one gets the variance $\mathbb{V}\text{ar}(Y) = \|Y - \mathbb{E}Y\|_2^2$ and therefore these functionals are generalizations of the variance.

Proposition 2.

(i)

$$\mathcal{D}(Y) := \|Y - \mathbb{E}Y\|_p^p = \sup\left\{\mathbb{E}(Y\,Z) - \frac{p^{1-q}}{q} D_q(Z) : \mathbb{E}Z = 0\right\} \tag{9}$$

where $D_q(Z) = \inf\{\|Z - a\|_q^q : a \in \mathbb{R}\}$ *and* $1/p + 1/q = 1$. $\mathcal{D}$ *is convex and monotonic w.r.t. convex dominance (D5).*

(ii)

$$\mathcal{A}(Y) := \mathbb{E}Y - \|Y - \mathbb{E}Y\|_p^p = \inf\left\{\mathbb{E}(Y\,Z) + \frac{p^{1-q}}{q} D_q(Z) : \mathbb{E}Z = 1\right\}. \quad (10)$$

$\mathcal{A}$ is concave, but has none of the properties (A4)–(A6) in general.

Proof. According to Proposition 7(ii) of the Appendix we have that

$$\|Y - \mathbb{E}Y\|_p^p = \sup\left\{\mathbb{E}[(Y - \mathbb{E}Y)\,Z] - \frac{p^{1-q}}{q}\|Z\|_q^q\right\}$$

where $1/p + 1/q = 1$. Since $\mathbb{E}[(Y - \mathbb{E}Y)\,Z] = \mathbb{E}[Y\,(Z - \mathbb{E}Z)]$ and setting $\tilde{Z} = Z - \mathbb{E}Z$ one gets that $\mathbb{E}\tilde{Z} = 0$ and $Z = \tilde{Z} - a$, where $a = -\mathbb{E}Z$. Maximizing $\mathbb{E}(Y\,\tilde{Z}) - \frac{p^{1-q}}{q}\|\tilde{Z} - a\|_q^q$ w.r.t. a leads to minimizing $\|\tilde{Z} - a\|_q^q$ in a and this gives the desired expression (9).

To show (D5) we have, in view of the Main Theorem (vi), to show that for all σ-algebras $\mathcal{F}$

$$D_q(\mathbb{E}(Z|\mathcal{F})) \le D_q(Z) \qquad \text{for all} \qquad Z \in \mathcal{Z}. \quad (11)$$

This property holds by the contraction property of the conditional expectation (see [14], proposition I-1-12, page 11) entailing that $\|\mathbb{E}(Z|\mathcal{F}) - a\|_q = \|\mathbb{E}(Z - a|\mathcal{F})\|_q \le \|Z - a\|_q$ for all a. Hence (11) follows.

To prove (10) one has to apply (4). Notice that $D_q(Z) = D_q(1 - Z)$. This leads immediately to (10). □

In particular, we get the following expression for the variance

$$\mathbb{V}\text{ar}(Y) = \sup\left\{\mathbb{E}(Y\,Z) - \frac{1}{4}\mathbb{V}\text{ar}(Z) : \mathbb{E}Z = 0\right\}. \quad (12)$$

3.2. The standard deviation and similar pure risk measures

We consider deviation measures of the form $\mathcal{D}(Y) = \|Y - \mathbb{E}Y\|_p$. Setting $p = 2$ one gets the standard deviation $\mathbb{S}\text{td}(Y) = \|Y - \mathbb{E}Y\|_2$ and therefore these functionals are generalizations of the standard deviation.

Proposition 3.

(i)

$$\mathcal{D}(Y) := \|Y - \mathbb{E}Y\|_p = \sup\left\{\mathbb{E}(Y\,Z) : \mathbb{E}(Z) = 0,\, D_q(Z) \le 1\right\} \quad (13)$$

where $1/p + 1/q = 1$. $\mathcal{D}$ is convex, homogeneous (D4) and monotonic w.r.t. convex dominance (D5).

(ii)

$$\mathcal{A}(Y) := \mathbb{E}(Y) - \|Y - \mathbb{E}Y\|_p = \inf\left\{\mathbb{E}(Y\,Z) : \mathbb{E}(Z) = 1,\, D_q(Z - 1) \le 1\right\} \quad (14)$$

$\mathcal{A}$ is concave, positively homogeneous (A4), but is not monotonic in general.

(iii) The functional $\mathbb{E}Y - \frac{1}{2}\|Y - \mathbb{E}Y\|_1$ *is monotonic w.r.t. SSD.*

Proof. According to proposition 7(i) of the Appendix we have that

$$\|Y - \mathbb{E}Y\|_p = \sup\left\{\mathbb{E}[(Y - \mathbb{E}Y)\, Z] : \|Z\|_q \le 1\right\}$$

where $1/p + 1/q = 1$. Setting $\tilde{Z} = Z - \mathbb{E}Z$ one gets that $\mathbb{E}\tilde{Z} = 0$ and $Z = \tilde{Z} - a$, where $a = -\mathbb{E}Z$. The condition for $\mathcal{Z}$ is that there is an a such that $\|\tilde{Z} - a\|_q \le 1$, which is equivalent to $D_q(\tilde{Z}) \le 1$. To go from (13) to (14) one has to use (4).

Obviously $\mathcal{A}$ and $\mathcal{D}$ are homogeneous. To show that $\mathcal{D}$ is monotonic w.r.t. convex dominance, we have to show that $\mathcal{Z} = \{Z : \mathbb{E}(Z) = 1, D_q(Z) \le 1\}$ is stable w.r.t. conditional expectations. This follows however from $\mathbb{E}(\mathbb{E}(Z|\mathcal{F})) = \mathbb{E}(Z)$ and from (11).

To prove (iii) notice that

$$\|Y - \mathbb{E}Y\|_1 = \sup\{\mathbb{E}(Y\, Z) : \mathbb{E}Z = 0; \exists a : |a| \le 1, |Z - a| \le 1 \text{ a.s. }\}.$$

and therefore

$$\frac{1}{2}\|Y - \mathbb{E}Y\|_1 = \sup\left\{\mathbb{E}(Y\, Z) : \mathbb{E}Z = 0; \exists a : |a| \le \frac{1}{2}, |Z - a| \le \frac{1}{2} \text{ a.s. }\right\}.$$

By (4)

$$\mathbb{E}Y - \frac{1}{2}\|Y - \mathbb{E}Y\|_1 = \inf\left\{\mathbb{E}(Y\, Z) : \mathbb{E}Z = 1; \exists a : |a| \le \frac{1}{2}, |1 - Z - a| \le \frac{1}{2} \text{ a.s.}\right\}.$$

The inequalities $|a| \le \frac{1}{2}$ and $|1 - Z - a| \le \frac{1}{2}$ imply that $0 \le 1/2 - a \le Z \le 3/2 - a$ for some $|a| \le 1/2$. Thus $Z \ge 0$ and since the conditional expectation respects a.s. bounds, we have shown (iii). □

3.3. The lower semi-variance and similar pure risk measures

The lower semi-variance is $\mathbb{V}\mathrm{ar}^-(Y) = \|[Y - \mathbb{E}Y]^-\|_2^2$, where $[V]^- = -\min(V, 0)$. This case appears as the special case $p = 2$ of the more general case considered below.

Proposition 4.

(i)

$$\begin{aligned}\mathcal{D}(Y) &:= \|[Y - \mathbb{E}Y]^-\|_p^p \\ &= \sup\left\{\mathbb{E}(Y\, Z) - \frac{p^{1-q}}{q}\mathbb{E}[(\mathit{essup}\; Z - Z)^q] : \mathbb{E}(Z) = 0\right\}. \qquad (15)\end{aligned}$$

(ii)

$$\begin{aligned}\mathcal{A}(Y) &:= \mathbb{E}Y - \|[Y - \mathbb{E}Y]^-\|_p^p \\ &= \inf\left\{\mathbb{E}(Y\, Z) + \frac{p^{1-q}}{q}\mathbb{E}[(Z - \mathit{essinf}\; Z)^q] : \mathbb{E}(Z) = 1\right\}. \qquad (16)\end{aligned}$$

$\mathcal{A}$ *is concave, but neither positively homogeneous nor monotonic in general.*

Proof. By Proposition 7(iv) of the Appendix, we have that $\|[Y - \mathbb{E}Y]^-\|_p^p = \sup\left\{\mathbb{E}[Y(Z - \mathbb{E}Z)] - \frac{p^{1-q}}{q}\|Z\|_q : Z \le 0\right\}$. Let $\tilde{Z} = Z - \mathbb{E}Z$. Then $\tilde{Z} \le \mathbb{E}Z$ and $\text{essup}\,\tilde{Z} \le \mathbb{E}Z$. Thus $Z \le \text{essup}\,\tilde{Z} - \tilde{Z}$ and this implies (i).

To get from (i) to (ii) one has to use (4). □

3.4. The lower semi-standard deviation and similar pure risk measures

The lower semi-standard deviation is $\mathbb{S}\text{td}^-(Y) = \|[Y - \mathbb{E}Y]^-\|_2$. This case appears as the special case $p = 2$ of the more general case considered below.

Proposition 5.

(i)

$$\begin{aligned}\mathcal{D}(Y) &:= \|[Y - \mathbb{E}Y]^-\|_p \\ &= \sup\left\{\mathbb{E}(Y\,Z) : \mathbb{E}(Z) = 0,\, Z \le 1;\, \|\,\textit{essup}\; Z - Z\|_q \le 1\right\}. \end{aligned} \tag{17}$$

(ii)

$$\begin{aligned}\mathcal{A}(Y) &:= \mathbb{E}Y - \|[Y - \mathbb{E}Y]^-\|_p \\ &= \inf\left\{\mathbb{E}(Y\,Z) : Z \ge 0,\, \mathbb{E}(Z) = 1,\, \|Z - \textit{essinf}\; Z\|_q \le 1\right\}. \end{aligned} \tag{18}$$

Therefore $\mathcal{A}$ is positively homogeneous and monotonic w.r.t. SSD.

Proof. By Proposition 7(iii) of the Appendix, we have that

$$\|[Y - \mathbb{E}Y]^-\|_p = \sup\left\{\mathbb{E}[Y(Z - \mathbb{E}Z)] : Z \le 0,\, \mathbb{E}(Z) \ge -1,\, \|Z\|_q \le 1\right\}.$$

Setting $\tilde{Z} = Z - \mathbb{E}Z$ and $a = -\mathbb{E}Z \ge 0$ one gets

$$\|[Y - \mathbb{E}Y]^-\|_p = \sup\left\{\mathbb{E}(Y\tilde{Z}) : \exists a,\, 0 \le a \le 1,\, \|a - \tilde{Z}\|_q \le 1;\, \mathbb{E}(\tilde{Z}) = 0;\, \tilde{Z} \le a\right\}. \tag{19}$$

Obviously one may set $a = \text{essup}\,\tilde{Z}$ to get the assertion (17). To prove (ii) notice that $\mathcal{A}(Y)$ equals

$$\begin{aligned}&\inf\left\{\mathbb{E}(Y\,Z) : \mathbb{E}(1 - Z) = 0,\, \|\,\text{essup}\,(1 - Z) - (1 - Z)\|_q \le 1,\, 1 - Z \le 1\right\} \\ &\quad = \inf\left\{\mathbb{E}(Y\,Z) : \mathbb{E}(Z) = 1,\, Z \ge 0,\, \|Z - \text{essinf} Z\|_q \le 1\right\}.\end{aligned}$$

The set $\mathcal{Z} = \{Z : \mathbb{E}(Z) = 1,\, Z \ge 0,\, \|Z - \text{essinf} Z\|_q \le 1\}$ consists of nonnegative densities and is stable w.r.t. conditional expectations, since $Z \ge 0$ implies that $\mathbb{E}(Z|\mathcal{F}) \ge 0$ and

$$\|\mathbb{E}(Z|\mathcal{F}) - \text{essinf}\,\mathbb{E}(Z|\mathcal{F})\|_q \le \|\mathbb{E}(Z|\mathcal{F}) - \text{essinf}\; Z\|_q \le \|Z - \text{essinf}\; Z\|_q$$

Here we have again used the contraction property of the conditional expectation as in the proof of Proposition 2. □

Remark. For $p = q = 2$, there is another representation

$$\begin{aligned}\mathcal{D}(Y) &= \|[Y - \mathbb{E}Y]^-\|_2 \\ &= \sup\left\{\mathbb{E}[YZ] : \mathbb{E}(Z) = 0, Z \leq \sqrt{1 - \|Z\|_2^2}, \|Z\|_2 \leq 1\right\}\end{aligned} \tag{20}$$

This representation follows from (19) by using

$$\|a - \tilde{Z}\|_2 = \sqrt{\|\tilde{Z}\|_2^2 - 2\mathbb{E}(\tilde{Z}) + a^2} = \sqrt{\|\tilde{Z}\|_2^2 + a^2} \leq 1.$$

Combining this with the condition $\tilde{Z} \leq a$ one gets the conditions

$$\tilde{Z} \leq \sqrt{1 - \|\tilde{Z}\|_2^2}, \mathbb{E}(\tilde{Z}) = 0, \|\tilde{Z}\|_2^2 \leq 1.$$

This leads to (20).

3.5. Minimal prediction error risk measures

Suppose that the random variable Y has to be predicted by a point estimate a. The cost of deviation between the estimate and the true value is given by a convex function h, satisfying $h(0) = 0$ and $h(u) > 0$ for $u \neq 0$. Then the *minimal prediction error* is $\mathcal{D}(Y) = \inf\{\mathbb{E}h(Y - a) : a \in \mathbb{R}\}$, where a is the *prediction estimate*. $\mathcal{D}(Y)$ may be viewed as a pure risk estimate, since it exhibits properties (D1) - (D3) as it is shown below.

Proposition 6. *Let*

$$\mathcal{D}(Y) := \inf\{\mathbb{E}[h(Y - a)] : a \in \mathbb{R}\}.$$

(i) Then $\mathcal{D}$ fulfills (D1)–(D3) and has the representation

$$\mathcal{D}(Y) := \sup\{\mathbb{E}(YZ) - \mathbb{E}h^*(Z); \mathbb{E}Z = 0\} \tag{21}$$

where h^* is the Fenchel dual of h,

$$h^*(v) = \sup\{uv - h(u) : u \in \mathbb{R}\}.$$

(ii)

$$\mathcal{A}(Y) := \mathbb{E}Y - \mathcal{D}(Y) = \inf\left\{\mathbb{E}(Y\,Z) + \mathbb{E}[h^*(1 - Z)] : \mathbb{E}(Z) = 1\right\}.$$

Proof. Properties (D1) and (D2) are obvious due to the assumption on h. (D3) will follow, if we show (21). To this end, suppose first that h is piecewise linear, i.e.

$$h(y) = \sup\{yx_i - h^*(x_i) : i = 1, \ldots, I\}$$

where x_i are the subgradients and h^* is the dual function. Let $(A_i)_{1\le i\le I}$ be a partition of the measure space Ω. Then using the saddlepoint property of convex-concave functions, we get

$$\inf_a \mathbb{E}[h(Y-a)] = \inf_a \sup_{A_i} \mathbb{E}\left[\sum_i \mathbb{1}_{A_i}(Y-a)x_i - \sum_i \mathbb{1}_{A_i} h^*(x_i)\right]$$
$$= \sup_{A_i} \inf_a \left[\mathbb{E}\sum_i \mathbb{1}_{A_i} Y x_i\right] - a\mathbb{E}\left[\sum_i \mathbb{1}_{A_i} x_i\right] - \mathbb{E}\left[\sum_i \mathbb{1}_{A_i} h^*(x_i)\right] \tag{22}$$

Denote the random variable $\sum_i \mathbb{1}_{A_i} x_i$ by Z. The supremum (22) will be $-\infty$, unless $\mathbb{E}\left[\sum_i \mathbb{1}_{A_i} x_i\right] = 0$. The conditions on Z are therefore that $\mathbb{E}(Z) = 0$ and that $\mathbb{E}[h^*(Z)] < \infty$. If $\mathbb{E}(Z) = 0$ the value of the supremum (22) is $\sup\{\mathbb{E}(YZ) - \mathbb{E}(h^*(Z)) : \mathbb{E}Z = 0\}$. In general, h may be approximated from below by a piecewise linear convex function and a limiting argument proves the assertion. □

Example 1. If we take $h(u) = u^+ + \frac{1-\alpha}{\alpha}u^-$, then we get the $\mathcal{D}(Y) = E(Y) - \mathbb{A}\mathrm{V@R}(Y)$, resp. $\mathcal{A}(Y) = \mathbb{A}\mathrm{V@R}_\alpha(Y)$, where $\mathbb{A}\mathrm{V@R}_\alpha$ is the average value-at-risk, given by

$$\mathbb{A}\mathrm{V@R}_\alpha(Y) = \max\left\{a - \frac{1}{\alpha}\mathbb{E}([Y-a]^-) : a \in \mathbb{R}\right\}.$$

$\mathbb{A}\mathrm{V@R}_\alpha$ was introduced under the name of CVaR (conditional value-at-risk) by Rockafellar and Uryasev (see [16]).

Since the dual of h is

$$h^*(v) = \begin{cases} 0 & \frac{\alpha-1}{\alpha} \le v \le 1 \\ \infty & \text{otherwise} \end{cases}$$

the dual representation of $\mathbb{A}\mathrm{V@R}$ is

$$\mathbb{A}\mathrm{V@R}_\alpha(Y) = \min\{\mathbb{E}(Y\,Z) : \mathbb{E}(Z) = 1, 0 \le Z \le 1/\alpha\}.$$

One arrives at the well known fact that $\mathbb{A}\mathrm{V@R}_\alpha$ is positively homogeneous and monotonic w.r.t. SSD.

Example 2. If we take $h(u) = u^2$, then we get $\mathcal{D}(Y) = \mathbb{V}\mathrm{ar}(Y)$. Since the dual of $h(u) = u^2$ is

$$h^*(v) = \frac{1}{4}v^2,$$

we arrive at the formula (12).

Example 3. Let us take $h(u) = u^2\mathbb{1}_{u\le 0} + cu\mathbb{1}_{u>0}$. This asymmetric pure risk measure penalizes large negative deviation much higher than large positive deviations. We have that

$$h^*(v) = \begin{cases} \infty & \text{if } v > c \\ 0 & \text{if } 0 \le v \le c \\ \frac{v^2}{4} & \text{if } v > 0 \end{cases}$$

Thus the pertaining pure risk measure is

$$\mathcal{D}(Y) = \sup\left\{\mathbb{E}(YZ) - \frac{1}{4}\mathbb{E}[(Z^-)^2] : \mathbb{E}Z = 0; Z \leq c\right\}.$$

4. Conclusions

We have presented sufficient criteria for the dual representation to entail some properties of the risk functionals. It is conjectured that most criteria are also necessary. While for all of the concrete risk functionals it is easy to construct examples to show that they lack of the corresponding property, whenever the sufficient condition is not satisfied, a more general statement seems difficult since one has to assume that super(sub)gradient set $\mathcal{Z}$ is in a way minimal. Adding some irrelevant functions to the super(sub)gradient set may not change the risk functional but destroy the structure. Further research is needed to clarify.

Acknowledgements. The author is indebted to three anonymous referees for various suggestions and improvements. In particular, the representation (20) was found by one of them. Also a stronger assertion of the Main Theorem was suggested.

Appendix: Some basic facts

Let $\|V\|_p = [\mathbb{E}|V|^p]^{1/p}$ and $1/p + 1/q = 1$.

Proposition 7.

(i) $\|V\|_p = \sup\{\mathbb{E}(VZ) : \|Z\|_q \leq 1\}$

(ii) $\|V\|_p^p = \sup\left\{\mathbb{E}(VZ) - \frac{p^{1-q}}{q}\|Z\|_q^q\right\}$

(iii) $\|[V]^-\|_p = \sup\left\{\mathbb{E}(VZ) : Z \leq 0, -1 \leq \mathbb{E}(Z) \leq 0, \|Z\|_q^q \leq 1\right\}$

(iv) $\|[V]^-\|_p^p = \sup\left\{\mathbb{E}(VZ) - \frac{p^{1-q}}{q}\|Z\|_q^q : Z \leq 0\right\}$

Here $[V]^- = -\min(V, 0)$.

Proof. Although (i) is well known (see [10]), let us give the argument here. By Hölder's inequality $\mathbb{E}(VZ) \leq \|V\|_p \cdot \|Z\|_q$ and therefore

$$\|V\|_p \geq \sup\{\mathbb{E}(VZ) : \|Z\|_q \leq 1\}.$$

Setting $Z = \text{ sgn }(V)|V|^{p-1}/\|V\|_p^{p-1}$ one sees that the inequality is in fact an equality. To prove (ii) start from the inequality

$$vw \leq \frac{|v|^p}{p} + \frac{|w|^q}{q}$$

and set $z = pw$ to get

$$vz \leq |v|^p + \frac{p}{q}p^{-q}|z|^q = |v|^p + \frac{p^{1-q}}{q}|z|^q.$$

Inserting the random variables V and Z and taking the expectations we get $\mathbb{E}(VZ) \leq \|V\|_p^p + \frac{p^{1-q}}{q}\|Z\|_q^q$. Therefore

$$\|V\|_p^p \geq \sup\left\{\mathbb{E}(VZ) - \frac{p^{1-q}}{q}\|Z\|_q^q\right\}.$$

Setting $Z = \operatorname{sgn}(V)p|V|^{p/q}$ one sees that the inequality is in fact an equality. To prove (iii) notice that

$$\begin{aligned}\|[V]^-\|_p &= \sup\left\{\mathbb{E}(VZ) : \|Z\|_q \leq 1, \{Z \neq 0\} \subset \{V < 0\}\right\} \\ &\geq \sup\left\{\mathbb{E}(VZ) : \|Z\|_q \leq 1, Z \leq 0; \{Z \neq 0\right\} \subset \{V < 0\}\} \\ &= \sup\left\{\mathbb{E}(VZ) : \|Z\|_q \leq 1, Z \leq 0\right\}\end{aligned}$$

Setting $Z = -\mathbb{1}_{\{V<0\}}|V|^{p/q}\|[V]^-\|_p^{-p/q}$ one sees that equality holds. Moreover, notice that

$$\mathbb{E}(-Z) = \mathbb{E}(|[V]^-|^{p-1})\|[V]^-\|_p^{-p/q} \leq (\|[V]^-\|_p^p)^{p-1}\|[V]^-\|_p^{-p/q} = 1$$

and this shows that the additional condition $\mathbb{E}(Z) \geq -1$ does not harm. A very similar argument, using (ii) leads to (iv). □

References

1. Artzner, P., Delbaen, F., Eber, J.M., Heath, D.: Coherent measures of risk. Mathematical Finance **9**, 203–228 (1999)
2. Basset, G., Koenker, R., Kordas, G.: Pessimistic portfolio allocation and Choquet expected utility. Journal of Financial Economics **2**, 477–492 (2004)
3. Bonnans, F., Shapiro, A.: Perturbation Analysis of Optimization Problems. Springer Verlag, New York, 2000
4. De Giorgi, E.: Reward-risk portfolio selection and stochastic dominance. Journal of Banking & Finance **29** (4), 895–926 (2005)
5. Delbaen, F.: Coherent measures of risk on general probability spaces. In: K. Sandmann, P. Schönbucher (eds.), Advances in Finance and Stochastics. Essays in Honour of Dieter Sondermann. Springer Verlag, New York, 2002, pp. 1–38
6. Denneberg, D.: Distorted probabilities and insurance premiums. In: Proceedings of the 14th SOR, Ulm. Athenäum, Frankfurt, 1989
7. Föllmer, H., Schied, A.: Convex measures of risk and trading constraints. Stochastics and Finance **6**, 429–447 (2002)
8. Föllmer, H., Schied, A.: Robust preferences and convex measures of risk. In: K. Sandmann, P. Schönbucher (eds.), Advances in Finance and Stochastics. Springer Verlag, Berlin, Heidelberg, New York, 2002, pp. 39–56
9. Frittelli, M., Rosazza Gianin, E.: Dynamic convex risk measures. In: G. Szegö (ed.), Risk measures for the 21st century, chap. 12. J. Wiley and Sons, New York, 2004, pp. 227–248
10. Hewitt, E., Stromberg, K.: Real and Abstract Analysis. Springer Verlag, Berlin, 1969
11. Jaschke, S., Küchler, U.: Coherent risk measures, valuation bounds, and $(\mu; \sigma)$- portfolio optimization. Finance and Stochastics **5**, 181–200 (2001)
12. Kusuoka, S.: On law invariant risk measures. Advances In Mathematical Economics **3**, 83–95 (2001)
13. Lehmann, E.: Some concepts of dependence. Ann. Math. Stat. **37**, 1137–1153 (1966)
14. Neveu, J.: Discrete Parameter Martingals. North Holland, Amsterdam, 1975
15. Ogryczak, W., Ruszczynski, A.: Dual stochastic dominance and related mean risk models. SIAM Journal on Optimization **13**, 60–78 (2002)
16. Rockafellar, R., Uryasev, S.: Optimization of conditional value-at-risk. Journal of Risk **2**, 21–41 (2000)
17. Rockafellar, R., Uryasev, S., Zabarankin, M.: Deviation measures in risk analysis. Tech. rep., ISE Dept.,Univerity of Florida (2002). Research report, http://www.ise.ufl.edu/uryasev

18. Rockafellar, R., Uryasev, S., Zabarankin, M.: Generalized deviations in risk analysis. Tech. rep., ISE Dept.,Univerity of Florida (2004). Research report, http://www.ise.ufl.edu/uryasev
19. Ruszczynski, A., Shapiro, A.: Optimization of convex risk functions. In: G. Calafiore, F. Dabbene (eds.), Probabilistic and Randomized Methods for Design Under Uncertainty. Springer Verlag, 2006
20. Strassen, V.: The existence of probability measures with given marginals. Annals of Mathematical Statistics **36**, 423–439 (1965)