Math. Program., Ser. B (2010) 123:85–100
DOI 10.1007/s10107-009-0321-6

FULL LENGTH PAPER

Robust stochastic dominance and its application to risk-averse optimization

Darinka Dentcheva · Andrzej Ruszczyński

Received: 3 March 2008 / Accepted: 6 March 2009 / Published online: 12 November 2009

Abstract We introduce a new preference relation in the space of random variables, which we call robust stochastic dominance. We consider stochastic optimization problems where risk-aversion is expressed by a robust stochastic dominance constraint. These are composite semi-infinite optimization problems with constraints on compositions of measures of risk and utility functions. We develop necessary and sufficient conditions of optimality for such optimization problems in the convex case. In the nonconvex case, we derive necessary conditions of optimality under additional smoothness assumptions of some mappings involved in the problem.

Keywords Robust preferences · Stochastic order · Stochastic dominance constraints · Risk constraints · Semi-infinite optimization

Mathematics Subject Classification (2000) 90C15 · 90C46 · 90C48 · 46N10 · 60E15

1 Introduction

The relation of *stochastic dominance* is a fundamental concept of probability, decision theory, and economics. A random variable $X \in \mathcal{L}_1(\Omega, \mathcal{F}, P_0)$ *dominates* another

This research was supported by the NSF awards DMS-0603728 and DMS-0604060.

D. Dentcheva (✉)
Department of Mathematical Sciences, Stevens Institute of Technology, Hoboken, NJ 07030, USA
e-mail: darinka.dentcheva@stevens.edu

A. Ruszczyński
Department of Management Science and Information Systems,
Rutgers University, Piscataway, NJ 08854, USA
e-mail: rusz@business.rutgers.edu

random variable $Y \in \mathcal{L}_1(\Omega, \mathcal{F}, P_0)$ in the second order, which we write $X \succeq_{(2)} Y$, if

$$\mathbb{E}[u(X)] \geq \mathbb{E}[u(Y)] \tag{1}$$

for every concave nondecreasing function $u(\cdot)$, for which these expected values are finite. We refer the reader to the monographs [27,34] for a modern view on the stochastic dominance relation and other comparison methods for random outcomes.

In stochastic systems, random outcomes are functions of our decisions z in some space $\mathcal{Z}$. Therefore, it is natural to consider $X = G(z)$, where G is a mapping from our decision space to the space of random variables.

In our earlier publications [5,6], we have introduced and analyzed the following optimization model with stochastic dominance constraints:

$$\begin{aligned} &\text{maximize} \quad \mathbb{E}\,[\mathrm{H(z)}] \\ &\text{subject to} \quad G(z) \succeq_{(2)} Y, \\ &\qquad\qquad\quad z \in Z_0. \end{aligned} \tag{2}$$

In this model, we assume that Z_0 is a convex closed subset of a Banach space $\mathcal{Z}$, and G and H are continuous operators from $\mathcal{Z}$ to the space of integrable random variables $\mathcal{L}_1(\Omega, \mathcal{F}, P_0)$.

The random variable Y plays the role of a benchmark outcome. For example, one may consider $Y = G(\bar{z})$, where $\bar{z} \in Z_0$ is some reasonable value of the decision vector, which is currently employed in the system.

In view of (1), the dominance constraint in (2) can also be expressed as infinitely many constraints of form

$$\mathbb{E}\,[u(G(z))] \geq \mathbb{E}\,[u(Y)]\,,$$

for all nondecreasing concave functions $u(\cdot)$, such that $\mathbb{E}[u(G(z))] < \infty$.

Model (2) is a convenient way to express risk-aversion in a stochastic optimization problem. It has recently been applied to financial optimization in [7] and to electricity market models in [19]. We have also analyzed model (2) for other dominance relations $\succeq_{(k)}$, where $k = 1, 2, 3, \ldots$ (see [5,8]), in the multivariate and dynamic cases (see [10,13]), and in inverse formulations, related to Lorenz curves (see [9]). Relations to the modern theory of measures of risk are investigated in [12].

In this paper we address the issue of uncertainty of the probability measure P_0. We consider the case when the "true" probability measure P is an element of a certain set $\mathcal{Q}$ of probability measures.

Let us introduce some notation used throughout the paper. The expected value operator is denoted by $\mathbb{E}$. The standard symbol $\mathcal{L}_p(\Omega, \mathcal{F}, P_0)$ denotes the space of all integrable mappings X from Ω to $\mathbb{R}$, such that $\mathbb{E}[|X|^p] < \infty$, where $p \in [1, \infty)$. The space of continuous functions on a compact set D is denoted by $\mathcal{C}(D)$. The space of regular countably additive measures on a compact set D having finite variation is denoted by $\mathcal{M}(D)$. Its subset of nonnegative measures is denoted by $\mathcal{M}_+(D)$.

For a measurable function $f : \mathbb{R} \to \mathbb{R}$ and an element $Y \in \mathcal{L}_1(\Omega, \mathcal{F}, P_0)$, we understand $f(Y)$ as a real random variable V with realizations $V(\omega) = f(Y(\omega))$, $\omega \in \Omega$.

For a convex set Z_0 in $\mathcal{Z}$ and a point $z \in Z_0$ we denote by $N_{Z_0}(z)$ the normal cone to Z_0 at z: $N_{Z_0}(z) = \{d \in \mathcal{Z}^* : \langle d, y - z \rangle \leq 0 \text{ for all } y \in Z_0\}$. Here $\mathcal{Z}^*$ is the topological dual space to $\mathcal{Z}$, and $\langle \cdot, \cdot \rangle$ refers to the dual pairing.

2 Robust stochastic dominance

The notion of stochastic ordering for scalar random variables (or *stochastic dominance of first order*) has been introduced in statistics in [21,24] and further applied and developed in economics [15,20,31]. It is defined as follows. For a random variable X we consider its distribution function, $F(X; \eta) = P[X \leq \eta]$, $\eta \in \mathbb{R}$. We say that a random variable X *dominates in the first order* a random variable Y if

$$F(X; \eta) \leq F(Y; \eta) \quad \forall\, \eta \in \mathbb{R}. \tag{3}$$

We denote this relation $X \succeq_{(1)} Y$. We refer the reader to [26,27,34,36] for a modern perspective on stochastic ordering.

Consider a scalar random variable $X \in \mathcal{L}_1(\Omega, \mathcal{F}, P_0)$ and define the function $F_2(X; \cdot)$ as

$$F_2(X; \eta) = \int_{-\infty}^{\eta} F(X; \alpha)\, \mathrm{d}\alpha, \quad \eta \in \mathbb{R}. \tag{4}$$

As an integral of a nondecreasing function, it is a convex function of η.

Definition 1 We say that a random variable $X \in \mathcal{L}_1(\Omega, \mathcal{F}, P_0)$ *dominates* in the second order another random variable $Y \in \mathcal{L}_1(\Omega, \mathcal{F}, P_0)$ if

$$F_2(X; \eta) \leq F_2(Y; \eta) \quad \forall \eta \in \mathbb{R}. \tag{5}$$

We denote relation (5) by $X \succeq_{(2)} Y$. In a similar way we can define higher order dominance relations (see [27]).

Changing the order of integration in (4) we get (see, e.g., [28])

$$F_2(X; \eta) = \mathbb{E}[(\eta - X)_+]. \tag{6}$$

Therefore, an equivalent representation of the second order stochastic dominance relation is:

$$\mathbb{E}[(\eta - X)_+] \leq \mathbb{E}[(\eta - Y)_+] \quad \forall\, \eta \in \mathbb{R}. \tag{7}$$

Let us consider the set $\mathcal{U}$ of concave nondecreasing functions $u : \mathbb{R} \to \mathbb{R}$ satisfying the following linear growth condition:

$$\lim_{t \to -\infty} u(t)/t < \infty. \tag{8}$$

For every random variable $X \in \mathcal{L}_1(\Omega, \mathcal{F}, P_0)$ and for every $u \in \mathcal{U}$ the quantity

$$\mathbb{E}[u(X)] = \int u(X(\omega))\, \mathrm{d}P(\omega)$$

is well-defined and finite. It is well-known in the theory of stochastic dominance that the set $\mathcal{U}$ is the generator of this order.

Proposition 1 *For each $X, Y \in \mathcal{L}_1(\Omega, \mathcal{F}, P_0)$ the relation $X \succeq_{(2)} Y$ is equivalent to*

$$\mathbb{E}[u(X)] \geq \mathbb{E}[u(Y)] \quad \forall\, u \in \mathcal{U}. \tag{9}$$

Let us consider random variables in the space $\mathcal{L}_p(\Omega, \mathcal{F}, P_0)$, where $p \in [1, \infty)$. Its topological dual space is $\mathcal{L}_q(\Omega, \mathcal{F}, P_0)$, where $1/p + 1/q = 1$. The set of measures Q which are absolutely continuous with respect to P_0 and have densities (Radon–Nikodym derivatives) dQ/dP_0 in the space $\mathcal{L}_q(\Omega, \mathcal{F}, P_0)$ will be identified with $\mathcal{L}_q(\Omega, \mathcal{F}, P_0)$. For $X \in \mathcal{L}_p(\Omega, \mathcal{F}, P_0)$ and $Q \in \mathcal{L}_q(\Omega, \mathcal{F}, P_0)$ we shall write

$$\mathbb{E}_Q[X] = \langle Q, X \rangle = \int_\Omega X(\omega)\, Q(\mathrm{d}\omega) = \int_\Omega X(\omega) \frac{\mathrm{d}Q}{\mathrm{d}P_0}(\omega)\, P_0(\mathrm{d}\omega).$$

Observe that the integral above is always well-defined.

Suppose we are given a set of probability measures $\mathcal{Q} \subset \mathcal{L}_q(\Omega, \mathcal{F}, P_0)$. We assume that $\mathcal{Q}$ is convex, closed and bounded in $\mathcal{L}_q(\Omega, \mathcal{F}, P_0)$, that is, the constant

$$B = \sup_{P \in \mathcal{Q}} \left\| \frac{\mathrm{d}P}{\mathrm{d}P_0} \right\|_q \tag{10}$$

is finite.

Definition 2 A random variable $X \in \mathcal{L}_p(\Omega, \mathcal{F}, P_0)$ *dominates robustly* a random variable $Y \in \mathcal{L}_p(\Omega, \mathcal{F}, P_0)$ in the second order over a set of probability measures $\mathcal{Q} \subset \mathcal{L}_q(\Omega, \mathcal{F}, P_0)$ if

$$\mathbb{E}_P[u(X)] \geq \mathbb{E}_P[u(Y)] \quad \forall\, u \in \mathcal{U},\ \forall\, P \in \mathcal{Q}.$$

We denote the robust dominance relation by $X \succeq_{(2)}^{\mathcal{Q}} Y$.

Our definition is related to the theory of robust preferences developed in [17] and further extended in [23]. In this theory, one derives from a set of axioms the existence of a utility function $\bar{u} : \mathbb{R} \to \mathbb{R}$ and a set of measures $\mathcal{Q}$ such that X is preferred over Y if and only if

$$\inf_{P \in \mathcal{Q}} \mathbb{E}_P[\bar{u}(X)] \geq \inf_{P \in \mathcal{Q}} \mathbb{E}_P[\bar{u}(Y)].$$

In general, the robust dominance relation does not entail a numerical representation of the preferences, but rather allows pairwise comparison.

It follows from (7) that an equivalent representation of the robust dominance relation is the following:

$$\mathbb{E}_P[(\eta - X)_+] \leq \mathbb{E}_P[(\eta - Y)_+], \quad \forall\, \eta \in \mathbb{R},\ \forall\, P \in \mathcal{Q}. \tag{11}$$

Equivalently,

$$\sup_{P \in \mathcal{Q}} \mathbb{E}_P\left[(\eta - X)_+ - (\eta - Y)_+\right] \leq 0, \quad \forall\, \eta \in \mathbb{R}.$$

Consider the functional $\sigma : \mathcal{L}_p(\Omega, \mathcal{F}, P_0) \to \overline{\mathbb{R}}$ defined as

$$\sigma(V) = \sup_{P \in \mathcal{Q}} \mathbb{E}_P[V].$$

We can view $\sigma(\cdot)$ as the support functional of the set $\mathcal{Q}$. The following facts are well-known in convex analysis [30, p. 41].

Proposition 2 *If the set $\mathcal{Q}$ is convex, closed and bounded, then $\sigma(\cdot)$ is a convex functional which is subdifferentiable everywhere, and*

$$\partial\sigma(V) = \{P \in \mathcal{Q} : \mathbb{E}_P[V] = \sigma(V)\}, \quad \forall\, V \in \mathcal{L}_p(\Omega, \mathcal{F}, P_0).$$

Moreover, $\sigma(\cdot)$ is Lipschitz continuous on $\mathcal{L}_p(\Omega, \mathcal{F}, P_0)$ with modulus B given by (10).

We can represent relation (11) as follows:

$$\sigma\left[(\eta - X)_+ - (\eta - Y)_+\right] \leq 0, \quad \forall\, \eta \in \mathbb{R}.$$

We see that the robust dominance relation can be interpreted as a continuum of inequalities for the functional $\sigma(\cdot)$ evaluated at the difference of shortfall random variables.

For a fixed Y, we define the functionals

$$\rho_\eta(X) = \sigma\left[(\eta - X)_+ - (\eta - Y)_+\right], \quad \eta \in \mathbb{R}. \tag{12}$$

Proposition 3 *For every $\eta \in \mathbb{R}$ the functional $\rho_\eta(\cdot)$ defined by* (12) *has the following properties*:

(i) *it is convex*;
(ii) *it is nonincreasing in the sense that*

$$(X_1(\omega) \geq X_2(\omega),\ \forall\, \omega \in \Omega) \Rightarrow \big(\rho_\eta(X_1) \leq \rho_\eta(X_2)\big)\,;$$

(iii) *it satisfies the subtranslation property: for all* $X \in \mathcal{L}_p(\Omega, \mathcal{F}, P_0)$ *and all* $a > 0$

$$\rho_\eta(X + a\mathbb{1}) \geq \rho_\eta(X) - a.$$

(iv) *it is Lipschitz continuous with modulus* B.

Proof By Proposition 2, the functional $V \mapsto \sup_{P\in\mathcal{Q}} E_P(V)$ is convex and Lipschitz continuous with modulus B. Since $\mathcal{Q}$ contains only probability measures, this functional is nondecreasing:

$$(V_1(\omega) \geq V_2(\omega),\ \forall\, \omega \in \Omega) \Rightarrow \left(\sup_{P\in\mathcal{Q}} E_P(V_1) \geq \sup_{P\in\mathcal{Q}} E_P(V_2)\right).$$

The function $X \mapsto (\eta - X)_+$ is convex and Lipschitz continuous with modulus 1. It is nonincreasing for every $\omega \in \Omega$. We conclude that $\rho_\eta(\cdot)$, as a composition of these two mappings, is convex, nonincreasing in the sense of (ii), and Lipschitz continuous with modulus B.

If $a > 0$, then for every $\omega \in \Omega$ the convexity and positive homogeneity of the function $x \mapsto (x)_+$ imply that

$$(\eta - X(\omega))_+ \leq a + (\eta - X(\omega) - a)_+.$$

Therefore, for every $P \in \mathcal{Q}$,

$$\begin{aligned}\mathbb{E}_P\left[(\eta - X - a)_+ - (\eta - Y)_+\right] &\geq \mathbb{E}_P\left[(\eta - X)_+ - a\mathbb{1} - (\eta - Y)_+\right]\\ &= \mathbb{E}_P\left[(\eta - X)_+ - (\eta - Y)_+\right] - a.\end{aligned}$$

Taking the supremum of both sides over $P \in \mathcal{Q}$ we obtain (iii). □

Properties (i) and (ii) established in Proposition 3 are identical to the first two axioms of coherent measures of risk (see [1,16,33]). The subtranslation property (iii) is weaker than the *translation property* of [1,16] requiring the equation $\rho(X + a\mathbb{1}) = \rho(X) - a$ for a convex measure of risk $\rho(\cdot)$.

The robust dominance relation can be rewritten as an infinite system of inequalities

$$\rho_\eta(X) \leq 0, \quad \forall\, \eta \in \mathbb{R}. \tag{13}$$

Let us calculate the subdifferential of the functional $\rho_\eta(\cdot)$. To this end we define for $\eta \in [a, b]$ the family of multifunctions $D_\eta : \mathcal{L}_p(\Omega, \mathcal{F}, P_0) \times \Omega \rightrightarrows \mathbb{R}$ as follows:

$$D_\eta(X, \omega) = \begin{cases} \{-1\} & \text{if } X(\omega) < \eta,\\ [-1, 0] & \text{if } X(\omega) = \eta,\\ \{0\} & \text{if } X(\omega) > \eta.\end{cases}$$

We also define the sets

$$\mathcal{A}_\eta(X) = \partial\sigma\left[(\eta - X)_+ - (\eta - Y)_+\right], \quad X \in \mathcal{L}_p(\Omega, \mathcal{F}, P_0).$$

Lemma 1 *For every $\eta \in \mathbb{R}$ the functional $\rho_\eta(\cdot)$ is continuous and subdifferentiable everywhere, and its subdifferenial at a point $X \in \mathcal{L}_p(\Omega, \mathcal{F}, P_0)$ is given by the formula*

$$\begin{aligned}\partial\rho_\eta(X) &= D_\eta(X, \cdot) \circ \mathcal{A}(X)\\ &= \Big\{ Q \in \mathcal{L}_q(\Omega, \mathcal{F}, P_0) : \exists \left(P_\eta \in \mathcal{A}_\eta(X)\right) \; \forall\, (\omega \in \Omega)\\ &\qquad \frac{dQ}{dP_\eta}(\omega) \in D_\eta(X, \omega) \Big\}. \end{aligned} \tag{14}$$

Proof Observe that $\rho_\eta(\cdot)$ is a composite function

$$\rho_\eta(X) = \sigma\left(\Psi_\eta(X)\right)$$

where $\Psi_\eta : \mathcal{L}_p(\Omega, \mathcal{F}, P_0) \to \mathcal{L}_p(\Omega, \mathcal{F}, P_0)$ is given by

$$\Psi_\eta(X) = (\eta - X)_+ - (\eta - Y)_+.$$

Writing $[\Psi_\eta(X)](\omega) = \psi(X, \omega)$, we notice that

$$\partial\psi(X, \omega) = D_\eta(X, \omega).$$

Applying [33, Prop.3.3], we conclude that

$$\begin{aligned}\partial\rho_\eta(X) = \operatorname{cl}\Big\{ & Q \in \mathcal{L}_q(\Omega, \mathcal{F}, P_0) : \exists \left(P_\eta \in \mathcal{A}_\eta(X)\right) \; \forall\, (\omega \in \Omega)\\ & \frac{\mathrm{d}Q}{\mathrm{d}P_\eta}(\omega) \in D_\eta(X, \omega) \Big\}.\end{aligned}$$

It remains to prove that the set on the right hand side of (14) is closed and the closure operation in the last equation may be skipped. Consider a sequence of measures

$$Q^k = \lambda^k P^k, \quad \lambda^k(\cdot) \in D_\eta(X, \cdot), \quad P^k \in \mathcal{A}_\eta(X),$$

and assume that $Q^k \to \hat{Q}$ in the weak* topology. By Proposition 2 the set $\mathcal{A}_\eta(X)$ is convex, weakly* closed, and bounded. Therefore, it is weakly* compact. Consequently, we can extract from the sequence $\{P^k\}$ a weakly* convergent subsequence $\{P^k\}_{k\in\mathcal{K}}$. Its weak* limit $\hat{P}$ is an element of $\mathcal{A}_\eta(X)$. For any event $A \subset \{\omega : X(\omega) < \eta\}$ we have $Q^k(A) = -P^k(A)$, and, thus, $\hat{Q}(A) = \hat{P}(A)$. For any event $B \subset \{\omega : X(\omega) > \eta\}$ we have $Q^k(B) = 0$ and $\hat{Q}(B) = 0$. Finally, if $C \subset \{\omega : X(\omega) = \eta\}$ we get

$$0 \geq Q^k(C) \geq -P^k(C),$$

which entails, after passing to the limit for $k \to \infty$, $k \in \mathcal{K}$, the inequalities

$$0 \geq \hat{Q}(C) \geq -\hat{P}(C).$$

It follows that all events of $\hat{P}$-measure zero have also $\hat{Q}$-measure zero, and, thus, the measure $\hat{Q}$ is absolutely continuous with respect to $\hat{P}$. By virtue of Radon–Nikodym theorem (see, e.g., [14, Thm. 5.5.4]) $\hat{Q}$ has a density (Radon–Nikodym derivative) with respect to $\hat{P}$. From our considerations it follows that the density satisfies the relation

$$\frac{\mathrm{d}\hat{Q}}{\mathrm{d}\hat{P}}(\omega) \in D_\eta(X, \omega), \quad \omega \in \Omega.$$

Therefore, the limit point $\hat{Q}$ is an element of the set (14). □

3 The risk optimization problem with robust dominance constraints

Let the operators $H : \mathcal{Z} \to \mathcal{L}_p(\Omega, \mathcal{F}, P_0)$, $G : \mathcal{Z} \to \mathcal{L}_p(\Omega, \mathcal{F}, P_0)$, and the functional $\varphi : \mathcal{L}_p(\Omega, \mathcal{F}, P_0) \to \mathbb{R}$, as well as a set Z_0 in $\mathcal{Z}$ be given. We assume that a benchmark random variable $Y \in \mathcal{L}_p(\Omega, \mathcal{F}, P_0)$ is given as well. In what follows we focus on optimization problems with robust dominance constraints, formulated as follow:

$$\begin{aligned} &\text{minimize} && \varphi(H(z)) \\ &\text{subject to} && G(z) \succeq_{(2)}^{\mathcal{Q}} Y, \\ &&& z \in Z_0. \end{aligned} \tag{15}$$

In view of (13), the robust dominance constraint can be equivalently formulated as follows:

$$\rho_\eta(G(z)) \leq 0, \quad \eta \in \mathbb{R}.$$

We relax problem (15) by enforcing the dominance relation on a compact interval $[a, b]$, rather than of the entire real line. We analyze the problem

$$\text{minimize} \quad \varphi(H(z)) \tag{16}$$

$$\text{subject to} \quad \rho_\eta(G(z)) \leq 0, \quad \eta \in [a, b], \tag{17}$$

$$z \in Z_0. \tag{18}$$

The reason for this relaxation is the need to satisfy the constraint qualification condition of Definition 3. In general, problems (15) and (16)–(18) are not equivalent, because the second one has a slightly larger feasible set. However, if we restricted the set of utility functions $\mathcal{U}$ to contain only concave nondecreasing functions which are affine outside of $[a, b]$, then both problems would be equivalent. Our objective is to develop optimality conditions for problem (16)–(18) in two cases: when both G and H are concave, and when they are continuously Fréchet differentiable, but not necessarily concave.

Problem (16)–(18) is related to two well-established structures in optimization theory: semi-infinite optimization and composite optimization. In the analysis of semi-infinite problems, it is usually assumed that the semi-infinite constraints (17) are defined by a linear or continuously differentiable function on the left hand side of the inequality, and the space $\mathcal{Z}$ is finite dimensional (see, *inter alia*, [2],[3, Sec. 5.4], [4,18]). In our case, these assumptions are not satisfied. The research on composite optimization focuses on the composite structure of the objective functional, as in (16) (see, *inter alia*, [3, Sec. 3.4.1], [29,35] and [37]).

In our setting, the main difficulty is associated with the infinite system of inequalities (17) of composite structure, which has only been investigated in [11].

4 Optimality conditions: convex case

Consider problem (16)–(18) under the assumption that G and H are concave continuous mappings from $\mathcal{Z}$ to $\mathcal{L}_p(\Omega, \mathcal{F}, P_0)$, and $\varphi(\cdot)$ is convex and nonincreasing functional. We also assume that the set Z_0 is convex.

Under these conditions, problem (16)–(18) becomes a convex optimization problem.

Definition 3 Problem (16)–(18) satisfies the *uniform robust dominance condition* if there exists a point $\tilde{z} \in Z_0$ such that

$$\max_{P\in\mathcal{Q}} \max_{a\le\eta\le b} \mathbb{E}_P\left[(\eta - G(\tilde{z}))_+ - (\eta - Y)_+\right] < 0.$$

Theorem 1 *Assume that the uniform robust dominance condition is satisfied. If $\hat{z}$ is an optimal solution of (16)–(18) then there exist measures $\hat{S} \in \partial\varphi(H(\hat{z}))$, $P_\eta \in \mathcal{A}_\eta(G(\hat{z}))$, a measurable selection $\lambda_\eta(\omega) \in D_\eta(G(\hat{z}), \omega)$, $\eta \in [a, b]$, $\omega \in \Omega$, and a measure $\hat{\nu} \in \mathcal{M}_+([a, b])$, such that $\hat{z}$ is an optimal solution of the problem*

$$\underset{z\in Z_0}{\text{minimize}} \left\{ \int_\Omega H(z)\,\hat{S}(\mathrm{d}\omega) + \int_a^b \int_\Omega \lambda_\eta(\omega) G(z)\, P_\eta(\mathrm{d}\omega)\,\hat{\nu}(\mathrm{d}\eta) \right\}, \tag{19}$$

and the following complementarity condition is satisfied

$$\int_a^b \mathbb{E}_{P_\eta}\left[(\eta - G(\hat{z}))_+\right]\,\hat{\nu}(\mathrm{d}\eta) = \int_a^b \mathbb{E}_{P_\eta}\left[(\eta - Y)_+\right]\,\hat{\nu}(\mathrm{d}\eta). \tag{20}$$

Conversely, if for some $\hat{S} \in \partial\varphi(H(\hat{z})$, $P_\eta \in \mathcal{A}_\eta(G(\hat{z}))$, $\lambda_\eta(\omega) \in D_\eta(G(\hat{z}), \omega)$, and $\hat{\nu} \in \mathcal{M}_+([a, b])$ the optimal solution of (19) satisfies (17) and (20), then $\hat{z}$ is an optimal solution of (16)–(18).

Proof We define the set

$$E = \{(z, X, V) \in Z_0 \times \mathcal{L}_p(\Omega, \mathcal{F}, P_0) \times \mathcal{L}_p(\Omega, \mathcal{F}, P_0) : G(z) \geq X,\ H(z) \geq V\}.$$

Consider the problem:

$$\begin{aligned} &\text{minimize} && \varphi(V) \\ &\text{subject to} && \rho_\eta(X) \leq 0, \quad \eta \in [a, b], \\ &&& (z, X, V) \in E. \end{aligned} \tag{21}$$

If $\hat{z}$ is an optimal solution of (16)–(18), then the triple $(\hat{z}, G(\hat{z}), H(\hat{z}))$ is an optimal solution of (21). Conversely, for every optimal solution $(\hat{z}, \hat{X}, \hat{V})$ of problem (21), due to the monotonicity of $\varphi(\cdot)$ and $\rho_\eta(\cdot)$ the triple $(\hat{z}, G(\hat{z}), H(\hat{z}))$ is also an optimal solution of (21). Then $\hat{z}$ is an optimal solution of (16)–(18). We can, therefore, focus on optimality conditions for problem (21).

The Lagrangian of problem (21) has the form:

$$L(X, V, \mu) = \varphi(V) + \int_a^b \rho_\eta(X)\, \mu(\mathrm{d}\eta). \tag{22}$$

Let us observe that the uniform dominance condition is a form of a generalized Slater condition: the function $\eta \mapsto \rho_\eta(G(\tilde{z}))$ is in the interior of the cone of nonpositive functions in $\mathcal{C}([a, b])$. If the point $(\hat{z}, \hat{X}, \hat{V})$ is a a solution of problem (21), we apply the necessary conditions of optimality in Banach spaces (see, e.g., [3, Thm. 3.4]). We conclude that there exists a nonnegative measure $\hat{\nu} \in \mathcal{M}([a, b])$ such that

$$L(G(\hat{z}), H(\hat{z}), \hat{\nu}) = \min_{(z,X,V)\in E} L(X, V, \hat{\nu}) \tag{23}$$

and the following complementarity condition holds true:

$$\int_a^b \rho_\eta(G(\hat{z}))\, \hat{\nu}(\mathrm{d}\eta) = 0. \tag{24}$$

As the Lagrangian (22) is convex, continuous and everywhere subdifferentiable with respect to (X, V), condition (23) can be equivalently stated as follows: there exists a subgradient

$$(\hat{Q}, \hat{S}) \in \partial_{(X,V)} L(G(\hat{z}), H(\hat{z}), \hat{\nu})$$

such that for all $(z, X, V) \in E$

$$\langle \hat{S}, V - H(\hat{z})\rangle + \langle \hat{Q}, X - G(\hat{z})\rangle \geq 0. \tag{25}$$

In order to calculate the subgradient $\hat{Q}$, consider the integral in (22). Since the space $\mathcal{L}_p(\Omega, \mathcal{F}, P_0)$, for $p \in [1, \infty)$ is separable, we can apply [22, Thm. 1.1]. We obtain

$$\partial \int_a^b \rho_\eta(X)\, \hat{\nu}(\mathrm{d}\eta) = \int\!\!\!\!\!\!*_a^b \partial \rho_\eta(X)\, \hat{\nu}(\mathrm{d}\eta),$$

where the last integral is understood as the collection of weak* integrals of all weakly* measurable selections of the multifunction $\eta \mapsto \partial\rho_\eta(X)$. Using the representation given in (14), we conclude that

$$\partial \int_a^b \rho_\eta(X)\, \hat{\nu}(\mathrm{d}\eta) = \left\{ \int\!\!\!\!\!\!*_a^b Q_\eta\, \hat{\nu}(\mathrm{d}\eta) : \frac{\mathrm{d}Q_\eta}{\mathrm{d}P_\eta}(\omega) \in D_\eta(X, \omega),\ \omega \in \Omega,\ P_\eta \in \mathcal{A}_\eta(X) \right\}. \tag{26}$$

Again, the integral on the right hand side is a weak* integral, i.e.,

$$\left\langle \int\!\!\!\!\!\!*_a^b Q_\eta\, \hat{\nu}(\mathrm{d}\eta), \bar{X} \right\rangle = \int_a^b \langle Q_\eta, \bar{X}\rangle\, \hat{\nu}(\mathrm{d}\eta), \quad \forall\, \bar{X} \in \mathcal{L}_p(\Omega, \mathcal{F}, P_0),$$

with Q_η satisfying the conditions in (26). As each P_η is a probability measure and all selections of $D_\eta(X, \cdot)$ are non-positive, every element Q of the subdifferential (26) is a non-positive measure. Since $\varphi(\cdot)$ is nonincreasing, its subgradients (and in particular $\hat{S}$) are non-positive measures as well. Therefore, condition (25) is equivalent to the inequality:

$$\langle \hat{S}, H(z) - H(\hat{z})\rangle + \langle \hat{Q}, G(z) - G(\hat{z})\rangle \geq 0, \quad \forall\, z \in Z_0. \tag{27}$$

Using representation (26) we conclude that there exist measures $P_\eta \in \mathcal{A}_\eta(G(\hat{z}))$ and measurable selections $\lambda_\eta(\omega) \in D_\eta(G(\hat{z}), \omega)$ such that

$$\begin{aligned} 0 &\leq \langle \hat{S}, H(z) - H(\hat{z})\rangle + \int_a^b \langle Q_\eta, G(z) - G(\hat{z})\rangle\, \hat{\nu}(\mathrm{d}\eta) \\ &= \int_\Omega \big[H(z) - H(\hat{z})\big]\, \hat{S}(\mathrm{d}\omega) + \int_a^b \int_\Omega \lambda_\eta(\omega) \big[G(z) - G(\hat{z})\big]\, P_\eta(\mathrm{d}\omega)\, \hat{\nu}(\mathrm{d}\eta). \end{aligned}$$

It follows that $\hat{z}$ is a solution of (19). Furthermore, using the selection of measures $P_\eta \in \mathcal{A}_\eta(G(\hat{z}))$, we transform condition (24) as follows

$$\int_a^b \rho_\eta(G(\hat{z}))\, \hat{\nu}(\mathrm{d}\eta) = \int_a^b \sup_{P \in \mathcal{Q}} \left\{ \mathbb{E}_P\left[(\eta - G(\hat{z}))_+\right] - \mathbb{E}_P\left[(\eta - Y)_+\right]\right\} \hat{\nu}(\mathrm{d}\eta)$$
$$= \int_a^b \left\{ \mathbb{E}_{P_\eta}\left[(\eta - G(\hat{z}))_+\right] - \mathbb{E}_{P_\eta}\left[(\eta - Y)_+\right]\right\} \hat{\nu}(\mathrm{d}\eta).$$

We conclude that condition (24) can be written as (20). □

If $D_\eta(G(\hat{z}), \cdot)$ is almost always single-valued, i.e., $P_0\left(X(\omega) = \eta\right) = 0$, we do not need the selections $\lambda_\eta(\cdot)$.

Corollary 1 *If $G(\hat{z})$ has a continuous distribution, problem (19) is equivalent to*

$$\underset{z \in Z_0}{\text{minimize}} \left\{ \int_\Omega H(z)\, \hat{S}(\mathrm{d}\omega) - \int_a^b \int_{G(\hat{z}) \le \eta} G(z)\, P_\eta(\mathrm{d}\omega)\, \hat{\nu}(\mathrm{d}\eta) \right\}.$$

5 Optimality conditions: nonconvex case

Now we consider problem (16)–(18) under a different set of assumptions. We assume that the mappings H and G are continuously Fréchet differentiable; we denote their Fréchet derivatives by H' and G', respectively. As in the previous section, we assume that $\varphi(\cdot)$ is a convex and nonincreasing functional, and that the set Z_0 is convex.

Define the set of feasible solutions:

$$Z = \left\{ z \in Z_0 : \rho_\eta(G(z)) \le 0,\ \forall\, \eta \in [a, b] \right\}.$$

For every feasible point $z \in Z$, we identify the set of active constraints:

$$I_0(z) = \left\{ \eta \in [a, b] : \rho_\eta(G(z)) = 0 \right\}.$$

Definition 4 The set Z satisfies the *differential robust dominance condition* at the point $z_0 \in Z$ if there exists a point $z_S \in Z_0$ and a constant $\delta > 0$ such that for all $\eta \in I_0(z_0)$

$$\sup_{Q \in \partial \rho_\eta(G(z))} \int_\Omega G'(z_0)(z_S - z_0)\, Q(\mathrm{d}\omega) \le -\delta. \tag{28}$$

The above condition may be interpreted as a generalization to the nonsmooth composite case of Robinson's constraint qualification condition introduced in [32].

We can characterize the Bouligand tangent cone $T_Z(z_0)$. We have proved that the mapping $\rho_\eta(\cdot)$ is Lipschitz continuous with constant B, for all η. The differential robust dominance condition implies the differential constraint qualification condition of [11, Thm.3], and, thus all assumptions of this theorem are satisfied. The following result follows by virtue of [11, Thm.3].

Theorem 2 *Assume that the differential robust dominance condition is satisfied at the point $z_0 \in Z$. Then $T_Z(z_0)$ is the set of vectors $d \in T_{Z_0}(z_0)$ satisfying the inequalities*

$$\int_\Omega G'(x_0) d Q(\mathrm{d}\omega) \le 0, \tag{29}$$

for all $Q \in \partial\rho_\eta(G(z_0))$ and all $\eta \in I_0(z_0)$.

Observe that our characterization of the tangent cone involves only the derivative of the mapping G. This allows us to work with a convex approximation $G^{\mathrm{c}}(z_0)$ of the feasible set Z, obtained by replacing $G(x)$ with its linearization $G(z_0) + G'(x_0)(z - z_0)$. It is defined as follows:

$$Z^{\mathrm{c}}(z_0) = \left\{ z \in Z_0 : \rho_\eta(G(z_0) + G'(z_0)(z - z_0))) \le 0, \ \forall\, \eta \in [a, b] \right\}.$$

For a linear operator $B : \mathcal{Z} \to \mathcal{L}_p(\Omega, \mathcal{F}, P_0)$ we use B^* to denote its adjoint operator, $B^* : \mathcal{L}_q(\Omega, \mathcal{F}, P_0) \to \mathcal{Z}^*$. Specifically, for a measure $Q \in \mathcal{L}_q(\Omega, \mathcal{F}, P_0)$ its value $B^* Q$ is the element of $\mathcal{Z}^*$ such that for all $z \in \mathcal{Z}$

$$\langle B^* Q, z \rangle = \int_\Omega [Bz](\omega)\, Q(\mathrm{d}\omega).$$

In particular, if Q is defined as a weak* integral, $Q = \displaystyle{*}\!\!\int_a^b P_\eta\, \nu(\mathrm{d}\eta)$, we have

$$\langle B^* Q, z \rangle = \int_a^b \int_\Omega [Bz](\omega)\, P_\eta(\mathrm{d}\omega)\, \nu(\mathrm{d}\eta).$$

Following the techniques of [11], we obtain the following result.

Theorem 3 *Assume that the point $\hat{z}$ is a local minimum of (16)–(18) and the differential robust dominance condition is satisfied at $\hat{x}$. Then there exist measures $\hat{S} \in \partial\varphi(H(\hat{z}))$, $P_\eta \in \mathcal{A}_\eta(G(\hat{z}))$, a measurable selection $\lambda_\eta(\omega) \in D_\eta(G(\hat{z}), \omega)$, $\eta \in [a, b]$, $\omega \in \Omega$, and a measure $\hat{\nu} \in \mathcal{M}_+([a, b])$, such that $\hat{z}$ is the solution of the problem*

$$\underset{z \in Z_0}{\text{minimize}} \left\{ \int_\Omega H'(\hat{z}) z \,\mathrm{d}\hat{S} + \int_a^b \int_\Omega \lambda_\eta(\omega) G'(\hat{z}) z\ P_\eta(\mathrm{d}\omega)\, \hat{\nu}(\mathrm{d}\eta) \right\}. \tag{30}$$

and

$$\int_a^b \mathbb{E}_{P_\eta}\left[(\eta - G(\hat{z}))_+\right]\ \hat{\nu}(\mathrm{d}\eta) = \int_a^b \mathbb{E}_{P_\eta}\left[(\eta - Y)_+\right]\ \hat{\nu}(\mathrm{d}\eta). \tag{31}$$

Proof Applying Theorem 4 from [11], we conclude that there exist a measure $\hat{\nu} \in \mathcal{M}_+([a,b])$ such that

$$0 \in [H'(\hat{z})]^* \partial\varphi(H(\hat{z})) + [G'(\hat{z})]^* \mathop{{\int}^{*}}_a^b \partial\rho_\eta(G(\hat{z}))\,\hat{\nu}(\mathrm{d}\eta) + N_{Z_0}(\hat{z}), \tag{32}$$

$$\int_a^b \rho_\eta(G(\hat{z}))\,\hat{\nu}(\mathrm{d}\eta) = 0. \tag{33}$$

We can transform these conditions by substituting the explicit forms of $\rho_\eta(\cdot)$ and of its subdifferential. Using the representation of the subdifferential given in (26), we conclude that there exist measures $P_\eta \in \mathcal{A}_\eta(G(\hat{z}))$ and $\hat{S} \in \partial\varphi(H(\hat{z}))$, as well as a measurable selection $\lambda_\eta(\omega) \in D_\eta(G(\hat{z}), \omega)$ such that

$$0 \in [H'(\hat{z})]^* \hat{S} + [G'(\hat{z})]^* \mathop{{\int}^{*}}_a^b \lambda_\eta(\omega) P_\eta\,\hat{\nu}(\mathrm{d}\eta) + N_{Z_0}(\hat{z}).$$

As Z_0 is a convex set, the last inclusion can be expressed also as

$$\Big\langle [H'(\hat{z})]^* \hat{S}, z - \hat{z} \Big\rangle + \left\langle G'(\hat{z})]^* \mathop{{\int}^{*}}_a^b \lambda_\eta(\omega) P_\eta\,\hat{\nu}(\mathrm{d}\eta), z - \hat{z} \right\rangle \ge 0, \quad \forall\, z \in Z_0.$$

By the definition of the weak* integral, the last relation is equivalent to

$$\int_\Omega H'(\hat{z})[z - \hat{z}]\,\mathrm{d}\hat{S} + \int_a^b \int_\Omega \lambda_\eta(\omega) G'(\hat{z})[z - \hat{z}]\,P_\eta(\mathrm{d}\omega)\,\hat{\nu}(\mathrm{d}\eta) \ge 0, \quad \forall\, z \in Z_0.$$

This means that $\hat{z}$ is the solution of the problem (30).

Furthermore, using the selection of measures $P_\eta \in \mathcal{A}_\eta(G(\hat{z}))$ in the definition of ρ_η, we transform condition (33) to the form (31) as in the proof of Theorem 1. □

If G and H are linear, the conditions (19) and (30) become the same. As in the convex case, if $D_\eta(G(\hat{z}), \cdot)$ is almost always single-valued, we do not need the selections $\lambda_\eta(\cdot)$.

Corollary 2 *If $G(\hat{z})$ has a continuous distribution, problem* (30) *can be written as follows*

$$\operatorname*{minimize}_{z\in Z_0} \left\{ \int_{\Omega} H'(\hat{z})z\, \hat{S}(\mathrm{d}\omega) - \int_a^b \int_{G(\hat{z})\le\eta} G'(\hat{z})z\, P_\eta(\mathrm{d}\omega)\, \hat{\nu}(\mathrm{d}\eta) \right\}.$$

Let us introduce the Lagrangian functional for the nonconvex problem (16)–(18), $\Lambda : \mathcal{Z} \times \mathcal{M}_+([a,b]) \to \mathbb{R}$:

$$\Lambda(z,\nu) = \varphi(H(z)) + \int_a^b \rho_\eta(G(z))\, \nu(\mathrm{d}\eta).$$

The functional $\Lambda(\cdot,\nu)$ is Lipschitz continuous and, therefore, we can utilize Mordukhovich calculus. Using [25, Prop. 1.112], under the additional assumption that the derivatives $H'(\hat{z})$ and $G'(\hat{z})$ are surjective, we can evaluate the Mordukhovich subdifferential of $\Lambda(\cdot,\hat{\nu})$ at $\hat{z}$:

$$\hat{\partial}\Lambda(\hat{z},\hat{\nu}) = [H'(\hat{z})]^*\hat{S} + [G'(\hat{z})]^* \int_a^b \lambda_\eta(\omega) P_\eta\, \hat{\nu}(\mathrm{d}\eta).$$

Therefore, the condition (32) can be equivalently stated as

$$0 \in \hat{\partial}\Lambda(\hat{z},\hat{\nu}) + N_{Z_0}(\hat{z}).$$

Theorem 3, however, cannot be derived by applying the Mordukhovich calculus to problem (16)–(18), because the structure of this problem is outside of the scope covered in [25].

If G and H are linear, $L(\cdot,\hat{\nu})$ is convex and achieves its minimum at $\hat{z}$ over the set Z_0.

References

1. Artzner, P., Delbaen, F., Eber, J.-M., Heath, D.: Coherent measures of risk. Math. Finance **9**, 203–228 (1999)
2. Bonnans, J.F., Cominetti, R.: Perturbed optimization in Banach spaces. III. Semi-infinite optimization. SIAM J. Control Optim. **34**, 1555–1567 (1996)
3. Bonnans, J.F., Shapiro, A.: Perturbation Analysis of Optimization Problems. Springer, New York (2000)
4. Canovas, M.J., Dontchev, A.L., Lopez, M.A., Parra, J.: Metric regularity of semi-infinite constraint systems. Math. Program. **104**, 329–346 (2005)
5. Dentcheva, D., Ruszczyński, A.: Optimization with stochastic dominance constraints. SIAM J. Optim. **14**, 548–566 (2003)
6. Dentcheva, D., Ruszczyński, A.: Optimality and duality theory for stochastic optimization problems with nonlinear dominance constraints. Math. Program. **99**, 329–350 (2004)

7. Dentcheva, D., Ruszczyński, A.: Portfolio optimization with stochastic dominance constraints. J. Banking Finance **30**(2), 433–451 (2006)
8. Dentcheva, D., Ruszczyński, A.: Semi-infinite probabilistic optimization: first order stochastic dominance constraints. Optimization **53**, 583–601 (2004)
9. Dentcheva, D., Ruszczyński, A.: Inverse stochastic dominance constraints and rank dependent expected utility theory. Math. Program. **108**, 297–311 (2006)
10. Dentcheva, D., Ruszczyński, A.: Optimization with multivariate stochastic dominance constraints. Math. Program. **117**, 111–127 (2009)
11. Dentcheva, D., Ruszczyński, A.: Composite semi-infinite optimization. Control Cybern **36**, 633–646 (2007)
12. Dentcheva, D., Ruszczyński, A.: Duality between coherent risk measures and stochastic dominance constraints in risk-averse optimization. Pac. J. Optim. **4**, 433–446 (2008)
13. Dentcheva, D., Ruszczyński, A.: Stochastic dominance for sequences and implied utility in dynamic optimization. C. R. Acad. Bulgare Sci. **57**(1), 15–22 (2008)
14. Dudley, R.M.: Real Analysis and Probability. Cambridge University Press, Cambridge (2002)
15. Fishburn, P.C.: Utility Theory for Decision Making. Wiley, New York (1970)
16. Föllmer, H., Schied, A.: Convex measures of risk and trading constraints. Finance Stoch. **6**, 429–447 (2002)
17. Gilboa, I., Schmeidler, D.: Maxmin expected utility with non-unique prior. J. Math. Econ. **18**, 141–153 (1989)
18. Goberna, M.A., Lopez, M.A.: Linear Semi-Infinite Optimization. Wiley, Chichester (1998)
19. Gollmer, R., Neise, F., Schultz, R.: Stochastic programs with first-order dominance constraints induced by mixed-integer linear recourse. SIAM J. Optim. **19**, 552–571 (2008)
20. Hadar, J., Russell, W.: Rules for ordering uncertain prospects. Am. Econ. Rev. **59**, 25–34 (1969)
21. Lehmann, E.: Ordered families of distributions. Ann. Math. Stat. **26**, 399–419 (1955)
22. Levin, V.L.: Convex Analysis in Spaces of Measurable Functions and its Applications in Economics, Nauka, Moscow (1985, in Russian)
23. Maccheroni, F., Rustichini, A., Marinacci, M.: Ambiguity aversion, robustness, and the variational representation of preferences. Econometrica **74**, 1447–1498 (2006)
24. Mann, H.B., Whitney, D.R.: On a test of whether one of two random variables is stochastically larger than the other. Ann. Math. Stat. **18**, 50–60 (1947)
25. Mordukhovich, B.: Variational Analysis and Generalized Differentiation. Springer, Berlin (2006)
26. Mosler, K., Scarsini, M. (eds.): Stochastic Orders and Decision Under Risk. Institute of Mathematical Statistics, Hayward (1991)
27. Müller, A., Stoyan, D.: Comparison Methods for Stochastic Models and Risks. Wiley, Chichester (2002)
28. Ogryczak, W., Ruszczyński, A.: From stochastic dominance to mean—risk models: semideviations as risk measures. Eur. J. Oper. Res. **116**, 33–50 (1999)
29. Penot, J.P.: Optimality conditions in mathematical programming and composite optimization. Math. Program. **67**, 225–246 (1994)
30. Phelps, R.R.: Convex Functions, Monotone Operators and Differentiability, Lecture Notes in Mathematics 1364. Springer, Berlin (1989)
31. Quirk, J.P., Saposnik, R.: Admissibility and measurable utility functions. Rev. Econ. Stud. **29**, 140–146 (1962)
32. Robinson, S.M.: Stability theory for systems of inequalities. II: differentiable nonlinear systems. SIAM J. Numer. Anal. **13**, 497–513 (1976)
33. Ruszczyński, A., Shapiro, A.: Optimization of convex risk functions. Math. Oper. Res. **31**, 433–452 (2006)
34. Shaked, M., Shanthikumar, J.G.: Stochastic Orders and Their Applications. Academic Press, Boston (1994)
35. Studniarski, M., Jeyakumar, V.: A generalized mean-value theorem and optimality conditions in composite nonsmooth minimization. Nonlinear Anal. **24**, 883–894 (1995)
36. Szekli, R.: Stochastic Ordering and Dependence in Applied Probability. Springer, New York (1995)
37. Yang, X.Q.: Second-order global optimality conditions for convex composite optimization. Math. Prog. **81**, 327–347 (1998)