Minimizing Vector Risk Measures

Alejandro Balbás, Beatriz Balbás, and Raquel Balbás

Abstract The minimization of risk functions is becoming very important due to its interesting applications in Mathematical Finance and Actuarial Mathematics. This paper addresses this issue in a general framework. Vector optimization problems involving many types of risk functions are studied. The "balance space approach" of multiobjective optimization and a general representation theorem of risk functions is used in order to transform the initial minimization problem in an equivalent one that is convex and usually linear. This new problem permits us to characterize optimality by saddle point properties that easily apply in practice. Applications in finance and insurance are presented.

1 Introduction

General risk functions are becoming very important in finance and insurance. Since the seminal paper of Artzner et al. (1999) introduced the axioms and properties of their "Coherent Measures of Risk", many authors have extended the discussion. The recent development of new markets (insurance or weather linked derivatives, commodity derivatives, energy/electricity markets, etc.) and products (inflation-linked bonds, equity indexes annuities or unit-links, hedge funds, etc.), the necessity of managing new types of risk (credit risk, operational risk, etc.) and the (often legal) obligation of providing initial capital requirements have made it rather convenient to overcome the variance as the most important risk measure and to introduce more general risk functions allowing us to address far more complex problems.

A. Balbás (✉)
University Carlos III of Madrid, CL. Madrid 126, 28903 Getafe, Madrid, Spain
e-mail: alejandro.balbas@uc3m.es

[1] It has been proved that the variance is not compatible with the Second Order Stochastic Dominance if asymmetries and/or heavy tails are involved. See Ogryczak and Ruszczynski (2002) for a very interesting analysis on the compatibility of more complex risk functions.

D. Jones et al. (eds.), *New Developments in Multiple Objective and Goal Programming*, Lecture Notes in Economics and Mathematical Systems 638,
DOI 10.1007/978-3-642-10354-4_4,

Hence, it is not surprising that the recent literature presents many interesting contributions focusing on new methods for measuring risk levels. Among others, Föllmer and Schied (2002) have defined the Convex Risk Measures, Goovaerts et al. (2004) have introduced the Consistent Risk Measures, and Rockafellar et al. (2006a) have defined the General Deviations and the Expectation Bounded Risk Measures.

Many classical actuarial and financial problems lead to optimization problems and have been revisited by using new risk functions. So, dealing with Portfolio Choice Problems, Mansini et al. (2007) use the Conditional Value at Risk ($CVaR$) and other complex risk measures in a discrete probability space, Alexander et al. (2006) compare the minimization of the Value at Risk (VaR) and the $CVaR$ for a portfolio of derivatives, Calafiore (2007) studies "robust" efficient portfolios in discrete probability spaces if risk levels are given by standard deviations or absolute deviations, and Schied (2007) deals with Optimal Investment with Convex Risk Measures.

Pricing and hedging issues in incomplete markets have also been studied (Föllmer and Schied 2002; Nakano 2004; Staum 2004; etc.) as well as Optimal Reinsurance Problems involving the $CVaR$ and stop loss reinsurance contracts (Cai and Tan 2007), and other practical problems.

Risk functions are almost never differentiable, which makes it rather difficult to provide general optimality conditions. This provokes that many authors must look for concrete properties of the special problem they are dealing with in order to find its solutions. Recent approaches by Rockafellar et al. (2006b) and Ruszczynski and Shapiro (2006) use the convexity of many risk functions so as to give general optimality conditions based on the sub-differential of the risk measure and the Fenchel Duality Theory (Luenberger 1969). The present article follows the ideas of the interesting papers above, in the sense that it strongly depends on Classical Duality Theory, but we attempt to use more properties of many risk functions that will enable us to yield new and alternative necessary and sufficient optimality conditions. Furthermore, since there is not any consensus with respect to "the best risk measure" to use in many practical applications, and the final result of many problems may critically depend on the risk measurement methodology we draw on, a second important difference between our approach and the previous literature is that we will deal with the simultaneous minimization of several risk functions, i.e., we will consider multiobjective problems. Bearing in mind the important topics of Mathematical Finance and Actuarial Mathematics that involve the minimization of risk measures, the discovery of new simple and practical rules seems to be a major objective.

The article's outline is as follows. Section 2 will present the general properties of the vector risk measure $\rho = (\rho_1, \rho_2, \ldots, \rho_r)$ and the optimization problem we are going to deal with. Since ρ is not differentiable in general, the optimization problem is not differentiable either, and Sect. 3 will be devoted to overcome this caveat. We will use the Balance Space Approach of multiobjective optimization (Balbás et al. 2005) and the Representation Theorem of Risk Measures so as to transform the initial optimization problem in an equivalent one that is differentiable and often linear. This goal is achieved by following and extending an original idea of Balbás et al. (2009).[2] However, the new problem involves new infinite dimensional Banach spaces of σ-additive measures, which provokes high degree of complexity when dealing with duality and optimality conditions. Therefore, the Mean Value Theorem (Lemma 3) is one of the most important results in this section and in the whole paper, since it will absolutely simplify the dual problem. As a consequence, Theorem 4 characterizes the optimal solutions by saddle points of a bilinear function of the feasible set and the sub-gradients of the risk measures to be simultaneously optimized. This seems to be profound finding whose proof is based on major results in Functional Analysis. Besides, the provided necessary and sufficient optimality conditions are quite different if one compares with those of previous literature. They are very general and easily apply in practical situations.

Section 4 presents two classical examples of Actuarial and Financial Mathematics that may be studied by minimizing risks. They are the Optimal Reinsurance Problem and the Portfolio Selection Problem. The novelty is given by the form of the problems, the level of generality of the analysis and the high weakness of the assumptions. The two examples are very important in practice, but this is not an exhaustive list of the real-world issues related to the optimization of risk functions. Another very interesting topics, like credit or operational risk, may be considered.

The last section of the paper points out the most important conclusions.

2 Dealing with Vector Risk Functions

Consider a probability space $(\Omega, \mathcal{F}, \mu)$ composed of the set of states of the word Ω, the σ-algebra $\mathcal{F}$ indicating the information available at a future date T, and the probability measure μ. Consider also $p \in [1, \infty)$ and $q \in (1, \infty]$ such that $1/p + 1/q = 1$, and the corresponding Banach spaces L^p and L^q. It is known that L^q is the dual space of L^p (Luenberger 1969). We will deal with a vector

$$\rho = (\rho_1, \rho_2, \ldots, \rho_r)$$

of risk functions

$$\rho_j : L^p \longrightarrow \mathbb{R}$$

such that the following condition holds:[3]

[2] Balbás and Romera (2007) also dealt with an infinite-dimensional linear optimization problem that allows us to hedge the interest rate risk, and Balbás et al. (2009) used Risk Measures Representation Theorems so as to extend the discussion and involve more general and complex sorts of risk.

[3] Hereafter $\mathbb{E}(x)$ will denote the mathematical expectation of the random variable x.

Assumption I. There exists $\varkappa_j \in \mathbb{R}$, $j = 1, 2, \dots, r$, such that

$$\Delta^q_{(\rho_j, \varkappa_j)} = \left\{ z \in L^q; -\mathbb{E}(yz) - \varkappa_j \leq \rho_j(y) \;\; \forall y \in L^p \right\} \tag{1}$$

is $\sigma(L^q, L^p)$-compact.[4]

Proposition 1. *Fix* $j \in \{1, 2, \dots, r\}$.

(a) *The sets* $\Delta^q_{(\rho_j, \varkappa_j)}$,

$$\Delta_{(\rho_j, \varkappa_j)} = \left\{ (z, k) \in L^q \times \left(-\infty, \varkappa_j\right]; \; -\mathbb{E}(yz) - k \leq \rho(y) \;\; \forall y \in L^p \right\} \tag{2}$$

and

$$\Delta^{\mathbb{R}}_{(\rho_j, \varkappa_j)} = \left\{ k \in \mathbb{R}; \;\; (z, k) \in \Delta_{(\rho_j, \varkappa_j)} \; for \; some \; z \in L^q \right\}$$

are convex. Moreover, $\Delta^q_{(\rho_j, \varkappa_j)}$ *is the natural projection of* $\Delta_{(\rho_j, \varkappa_j)}$ *on* L^q, *whereas* $\Delta^{\mathbb{R}}_{(\rho_j, \varkappa_j)}$ *is its natural projection on* $\mathbb{R}$.

(b) *Under Assumption I the set* $\Delta_{(\rho_j, \varkappa_j)}$ *is compact when endowed with the topology* $\tilde{\sigma}$, *product topology of* σ^* *and the usual one of the real line. Furthermore,* $\Delta^{\mathbb{R}}_{(\rho_j, \varkappa_j)}$ *is also compact and* $\Delta_{(\rho_j, \varkappa_j)}$ *is included in the* $\tilde{\sigma}$*-compact set* $\Delta^q_{(\rho_j, \varkappa_j)} \times \Delta^{\mathbb{R}}_{(\rho_j, \varkappa_j)}$.

Proof. (a) is trivial, so let us prove (b). Since the inclusion $\Delta_{(\rho_j, \varkappa_j)} \subset \Delta^q_{(\rho_j, \varkappa_j)} \times \Delta^{\mathbb{R}}_{(\rho_j, \varkappa_j)}$ is obvious it is sufficient to show that $\Delta^{\mathbb{R}}_{(\rho_j, \varkappa_j)}$ is compact and $\Delta_{(\rho_j, \varkappa_j)}$ is closed.

To see that $\Delta^{\mathbb{R}}_{(\rho_j, \varkappa_j)}$ is compact let us prove that it is closed and bounded. To see that it is closed let as assume that $(k_n)_{n \in \mathbb{N}}$ is a sequence in $\Delta^{\mathbb{R}}_{(\rho_j, \varkappa_j)}$ that converges to $k \in \mathbb{R}$. Take a sequence $(z_n, k_n)_{n \in \mathbb{N}} \subset \Delta_{(\rho_j, \varkappa_j)}$. Since $\Delta^q_{(\rho_j, \varkappa_j)}$ is compact take an agglomeration point z of $(z_n)_{n \in \mathbb{N}}$. Then it is easy to see that (z, k) is an agglomeration point of $(z_n, k_n)_{n \in \mathbb{N}}$. Thus,

$$-\mathbb{E}(y z_n) - k_n \leq \rho_j(y)$$

for every $n \in \mathbb{N}$ and every $y \in L^p$ leads to

$$-\mathbb{E}(yz) - k \leq \rho_j(y)$$

for every $y \in L^p$, and $(z, k) \in \Delta_{(\rho_j, \varkappa_j)}$, i.e., $k \in \Delta^{\mathbb{R}}_{(\rho_j, \varkappa_j)}$.

[4] In order to simplify the notation, henceforth the $\sigma(L^q, L^p)$ topology will be denoted by σ^*.

To see that $\Delta^{\mathbb{R}}_{(\rho_j,\varkappa_j)}$ is bounded it is sufficient to prove that it is bounded from below, since $\varkappa_j$ is an obvious upper bound. Expression (2) leads to

$$-\mathbb{E}(0) - k \le \rho_j(0),$$

for every $k \in \Delta^{\mathbb{R}}_{(\rho_j,\varkappa_j)}$, and $\mathbb{E}(0) = 0$ implies that $k \ge -\rho_j(0)$ for every $k \in \Delta^{\mathbb{R}}_{(\rho_j,\varkappa_j)}$.

To see that $\Delta_{(\rho_j,\varkappa_j)}$ is closed consider the net $(z_i, k_i)_{i \in I} \subset \Delta_{(\rho_j,\varkappa_j)}$ and its limit (z, k). Then,

$$-\mathbb{E}(yz_i) - k_i \le \rho_j(y)$$

for every $i \in I$ and every $y \in L^p$ leads to

$$-\mathbb{E}(yz) - k \le \rho_j(y)$$

for every $y \in L^p$, so $(z, k) \in \Delta_{(\rho_j,\varkappa_j)}$. □

Remark 1. As a consequence of the latter result and its proof Assumption I implies that $\Delta^{\mathbb{R}}_{(\rho_j,\varkappa_j)}$ is a bounded closed interval

$$\Delta^{\mathbb{R}}_{(\rho,\varkappa)} = \left[\varkappa_{0,j}, \varkappa_j\right] \subset \left[-\rho_j(0), \varkappa_j\right]. \tag{3}$$

Furthermore, as shown in the proof above, $\varkappa_{0,j} \ge -\rho_j(0)$.

We will also impose the following assumption:

Assumption II. The equality

$$\rho_j(y) = \text{Max}\left\{-\mathbb{E}(yz) - k;\ \ (z,k) \in \Delta_{(\rho_j,\varkappa_j)}\right\} \tag{4}$$

holds for every $y \in L^p$ and every $j = 1, 2, \ldots, r$.[5]

Next let us provide a proposition with a trivial (and therefore omitted) proof.

Proposition 2. *Under Assumptions I and II ρ_j is a convex function for $j = 1, 2, \ldots, r$.*

---

[5] Assumptions I and II frequently hold. For instance, they are always fulfilled if ρ_j is expectation bounded or a general deviation, in the sense of Rockafellar et al. (2006a) (in which case $\varkappa_{0,j} = \kappa_j = 0$), and often fulfilled if ρ_j is coherent (Artzner et al. 1999) or consistent Goovaerts et al. (2004). Furthermore, many convex risk measures (Föllmer and Schied 2002) also satisfy these assumptions.

Particular examples are the Absolute Deviation, the Standard Deviation, Down Side Semi-Demiations, the $CVaR$, the Wang Measure and the Dual Power Transform (Wang 2000, see also Cherny 2006).

Consider now a convex subset X included in an arbitrary vector space and a function

$$f : X \longrightarrow L^p$$

such that

$$\rho \circ f : X \longrightarrow \mathbb{R}^r$$

is convex. Possible examples arise when f is concave and ρ is decreasing (for instance, if every ρ_j is a coherent measure of risk) or if f is an affine function, i.e.,

$$f(tx_1 + (1-t)x_2) = tf(x_1) + (1-t)f(x_2)$$

holds for every $t \in [0, 1]$ and every $x_1, x_2 \in X$. We will deal with the multiobjective optimization problem

$$\begin{cases} \text{Min } \rho \circ f(x) \\ x \in X \end{cases}. \tag{5}$$

3 Saddle Point Optimality Conditions

Since (5) is convex, for every Pareto solution $x_0 \in X$ there exists $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_r) \geq (0, 0, \ldots, 0)$ such that $\sum_{j=1}^{r} \alpha_j = 1$ and $x_0 \in X$ solves

$$\begin{cases} \text{Min } \sum_{j=1}^{r} \alpha_j \rho_j \circ f(x) \\ x \in X \end{cases}. \tag{6}$$

The very well-known scalarization method consists in choosing an "adequate α" and then solving the problem (6) above. "Adequate α" means that α must be selected according to the decision maker preferences.

However, in this paper we will follow an alternative approach based on the notion of "Balance Point" (Galperin and Wiecek 1999 or Balbás et al. 2005, among others), since it will allow us to provide saddle point necessary and sufficient optimality conditions for (5).

So, consider that $d = (d_1, d_2, \ldots, d_r)$ is composed of strictly positive numbers and plays the role of "direction of preferential deviations" (Galperin and Wiecek 1999). Let us suppose the existence of a Pareto solution of (5) in the direction of d. According to Galperin and Wiecek (1999) d can be chosen by the decision maker depending on her/his preferences, and it indicates the marginal worsening of a given objective with respect to the improvement of an alternative one. If we assume the existence of "an ideal point" $\Upsilon \in \mathbb{R}^r$ whose coordinates are the optimal values of (5) when ρ_j substitutes ρ,[6] Balbás et al. (2005) have shown that if (x^*, θ^*) is a

[6] This assumption may be significantly relaxed (see Balbás et al. 2005), but it simplifies the exposition.

solution of the scalar problem

$$\begin{cases} \text{Min } \theta \\ \theta d + \Upsilon \geq \rho \circ f(x) \\ \theta \in \mathbb{R},\ x \in X \end{cases} \tag{7}$$

then x^* is a Pareto solution of (5) that satisfies

$$\rho \circ f\left(x^*\right) = \Upsilon + d\theta^*.$$

Conversely, for every Pareto solution x^* of (5) such that

$$\rho_j \circ f\left(x^*\right) > \Upsilon_j,$$

$j = 1, 2, \ldots, r$, there exist $d_1, d_2, \ldots, d_r > 0$ and $\theta^* > 0$ such that (x^*, θ^*) solves (7).[7]

Equation (4) clearly implies the equivalence between (7) and

$$\begin{cases} \text{Min } \theta \\ d_j\theta + \mathbb{E}\left(f(x) z_j\right) + k_j + \Upsilon_j \geq 0,\ \forall \left(z_j, k_j\right) \in \Delta_{\left(\rho_j, \varkappa_j\right)},\ j = 1, 2, \ldots, r. \\ \theta \in \mathbb{R},\ x \in X \end{cases} \tag{8}$$

The solutions of (8) will be characterized by a saddle point condition. In order to reach this result we need some additional notations and a crucial instrumental lemma. Hereafter $\mathcal{C}\left(\Delta_{\left(\rho_j, \varkappa_j\right)}\right)$, $j = 1, 2, \ldots, r$, will represent the Banach space composed of the real-valued σ^*-continuous functions on the σ^*-compact space $\Delta_{\left(\rho_j, \varkappa_j\right)}$. Similarly, $\mathcal{M}\left(\Delta_{\left(\rho_j, \varkappa_j\right)}\right)$ will denote the Banach space of σ-additive inner regular measures on the Borel σ-algebra of $\Delta_{\left(\rho_j, \varkappa_j\right)}$ (Horvàth 1966 or Luenberger 1969), and $\mathcal{P}\left(\Delta_{\left(\rho_j, \varkappa_j\right)}\right) \subset \mathcal{M}\left(\Delta_{\left(\rho_j, \varkappa_j\right)}\right)$ will be the set of inner regular probabilities. Recall that $\mathcal{M}\left(\Delta_{\left(\rho_j, \varkappa_j\right)}\right)$ is the dual space of $\mathcal{C}\left(\Delta_{\left(\rho_j, \varkappa_j\right)}\right)$.

Lemma 1 (Mean Value Theorem). *Fix $j \in \{1, 2, \ldots, r\}$. If $\nu \in \mathcal{P}\left(\Delta_{\left(\rho_j, \varkappa_j\right)}\right)$ then there exist $z_\nu \in \Delta^q_{\left(\rho_j, \varkappa_j\right)}$ and $k_\nu \in \left[\varkappa_{0,j}, \varkappa_j\right]$ such that $(z_\nu, k_\nu) \in \Delta_{\left(\rho_j, \varkappa_j\right)}$,*

$$\int_{\Delta^q_{\left(\rho_j, \varkappa_j\right)}} \mathbb{E}(yz)\, d\nu_q(z) = E(yz_\nu) \tag{9}$$

holds for every $y \in L^p$ and

[7] Moreover, this converse implication would also hold even if (5) were not a convex problem.

$$\int_{\varkappa_{0,j}}^{\varkappa_j} k d\nu_{\mathrm{R}}(k) = k_\nu. \tag{10}$$

Proof. Consider the natural projections $\nu_q \in \mathcal{P}\left(\Delta^q_{(\rho_j,\varkappa_j)}\right)$ and $\nu_{\mathrm{R}} \in \mathcal{P}\left[\varkappa_{0,j}, \varkappa_j\right]$ of ν, and the function

$$L^p \ni y \longrightarrow \psi(y) = \int_{\Delta^q_{(\rho_j,\varkappa_j)}} \mathbb{E}(yz)\, d\nu_q(z) \in \mathbb{R}.$$

It is obvious that ψ is linear so let us prove that it is also continuous. If $\Delta^q_{(\rho_j,\varkappa_j)}$ were bounded then there would exist $M \in \mathbb{R}$ such that $\|z\|_q \le M$ for every $z \in \Delta^q_{(\rho_j,\varkappa_j)}$. Then the Hölder inequality (Luenberger 1969) would lead to

$$|\mathbb{E}(yz)| \le \|y\|_p \|z\|_q \le \|y\|_p M$$

for every $y \in L^p$ and every $z \in \Delta^q_{(\rho_j,\varkappa_j)}$, and

$$|\psi(y)| \le \int M \|y\|_p \, d\nu_q(z) = M \|y\|_p$$

for every $y \in L^p$. Whence ψ would be continuous (Horvàth 1966 or Luenberger 1969). Let us see now that $\Delta^q_{(\rho_j,\varkappa_j)}$ is bounded. Since it is σ^*-compact the set $\left\{\mathbb{E}(yz);\, z \in \Delta^q_{(\rho_j,\varkappa_j)}\right\} \subset \mathbb{R}$ is bounded for every $y \in L^p$ because

$$L^q \ni z \longrightarrow \mathbb{E}(yz) \in \mathbb{R}$$

is σ^*-continuous. Then the Banach–Steinhaus Theorem (Horvàth 1966) shows that $\Delta^q_{(\rho_j,\varkappa_j)}$ is bounded.

Since ψ is continuous the Riesz Representation Theorem (Horvàth 1966) shows the existence of $z_\nu \in L^q$ such that (9) holds.

Besides, the inequalities

$$\varkappa_{0,j} \le \int_{\varkappa_{0,j}}^{\varkappa_j} k d\nu_{\mathrm{R}}(k) \le \varkappa_j$$

are obvious, so the existence of $k_\nu \in \left[\varkappa_{0,j}, \varkappa_j\right]$ satisfying (10) is obvious too.

It only remains to show that $(z_\nu, k_\nu) \in \Delta_{(\rho_j,\varkappa_j)}$. Indeed, (9) and (10) imply that

$$\begin{aligned}-\mathbb{E}\left(yz_{\nu}\right)-k_{\nu} &= -\int_{\Delta^{q}_{\left(\rho_j,\varkappa_j\right)}}\mathbb{E}\left(yz\right)d\nu_{q}\left(z\right)-\int_{\varkappa_{0,j}}^{\varkappa_j}kd\nu_{\mathbb{R}}\left(k\right)\\ &= \int_{\Delta_{\left(\rho_j,\varkappa_j\right)}}\left(-\mathbb{E}\left(yz\right)-k\right)d\nu\left(z,k\right)\\ &\leq \int_{\Delta_{\left(\rho_j,\varkappa_j\right)}}\rho\left(y\right)d\nu\left(z,k\right)\\ &= \rho_j\left(y\right)\end{aligned}$$

for every $y \in L^p$. □

Theorem 1 (Saddle Point Theorem). *Take $x^* \in X$ and $\theta^* \in \mathbb{R}$. (x^*, θ^*) solves (8) if and only if there exist $\left(z_j^*, k_j^*\right) \in \Delta_{\left(\rho_j,\varkappa_j\right)}$, $j = 1, 2, \ldots, r$ and*

$$\lambda^* \in \left\{\lambda = \left(\lambda_1, \lambda_2, \ldots, \lambda_r\right);\ \sum_{j=1}^{r} d_j\lambda_j = 1,\ \lambda_j \geq 0,\ j = 1, 2, \ldots, r\right\}$$

such that

$$\lambda_j^*\left(d_j\theta^* + \Upsilon_j + \mathbb{E}\left(f\left(x^*\right)z_j^*\right) + k_j^*\right) = 0,$$

$j = 1, 2, \ldots, r$, and

$$\sum_{j=1}^{r}\lambda_j^*\left(\mathbb{E}\left(f\left(x^*\right)z_j^*\right) + k_j^*\right) \geq \sum_{j=1}^{r}\lambda_j^*\left(\mathbb{E}\left(f\left(x\right)z_j^*\right) + k_j^*\right) \qquad (11)$$

for every $x \in X$. If so,

$$\rho_j \circ f\left(x^*\right) = -\left(k_j^* + \mathbb{E}\left(f\left(x^*\right)z_j^*\right)\right)$$

holds for every $j = 1, 2, \ldots, r$, and

$$\lambda_j^*\left(k_j^* + \mathbb{E}\left(f\left(x^*\right)z_j^*\right)\right) \leq \lambda_j^*\left(k_j + \mathbb{E}\left(f\left(x^*\right)z_j\right)\right) \qquad (12)$$

holds for every $j = 1, 2, \ldots, r$ and every $\left(z_j, k_j\right) \in \Delta_{\left(\rho_j,\varkappa_j\right)}$,[8]

---

[8] Notice that (11) and (12) show that

$$\left(x^*, \left(z_j^*, k_j^*\right)\right)$$

is a Saddle Point of the function

$$X \times \Pi_{j=1}^{r}\Delta_{\left(\rho_j,\varkappa_j\right)} \ni \left(x, \left(z_j, k_j\right)\right) \longrightarrow \sum_{j=1}^{r}\lambda_j^*\left(\mathbb{E}\left(f\left(x\right)z_j\right) + k_j\right) \in \mathbb{R}.$$

Proof. The constraints of (8) are valued on the Banach space $\mathcal{C}\left(\Delta_{\left(\rho_j,\varkappa_j\right)}\right)$, $j = 1,2,\ldots,r$. Accordingly, the Lagrangian function

$$\mathcal{L}: X \times \mathbb{R} \times \prod_{j=1}^{r} \mathcal{M}\left(\Delta_{\left(\rho_j,\varkappa_j\right)}\right) \longrightarrow \mathbb{R}$$

of (8) becomes (Luenberger 1969)

$$\begin{aligned}\mathcal{L}\left(x,\theta,\left(\nu_j\right)_{j=1}^{r}\right) = \theta&\left(1-\sum_{j=1}^{r} d_j \int_{\Delta_{\left(\rho_j,\varkappa_j\right)}} d\nu_j\left(z_j,k_j\right)\right)\\ &-\sum_{j=1}^{r}\int_{\Delta_{\left(\rho_j,\varkappa_j\right)}} \mathbb{E}\left(f\left(x\right)z_j\right) d\nu_j\left(z_j,k_j\right)\\ &-\sum_{j=1}^{r}\int_{\Delta_{\left(\rho_j,\varkappa_j\right)}} k_j d\nu_j\left(z_j,k_j\right)\\ &-\sum_{j=1}^{r}\Upsilon_j\int_{\Delta_{\left(\rho_j,\varkappa_j\right)}} d\nu_j\left(z_j,k_j\right),\end{aligned}$$

that may simplify to

$$\begin{aligned}\mathcal{L}\left(x,\theta,\left(\nu_j\right)_{j=1}^{r}\right) = \theta&\left(1-\sum_{j=1}^{r} d_j\lambda_j\right)\\ &-\sum_{j=1}^{r}\int_{\Delta_{\left(\rho_j,\varkappa_j\right)}} \mathbb{E}\left(f\left(x\right)z_j\right) d\nu_j\left(z_j,k_j\right)\\ &-\sum_{j=1}^{r}\int_{\Delta_{\left(\rho_j,\varkappa_j\right)}} k_j d\nu_j\left(z_j,k_j\right)\\ &-\sum_{j=1}^{r}\Upsilon_j\lambda_j,\end{aligned}$$

if $\lambda_j = \int_{\Delta_{(\rho,\varkappa)}} d\nu_j\left(z_j,k_j\right) \geq 0$ for $j = 1,2,\ldots,r$. It is clear that the infimum

$$Inf\ \left\{\mathcal{L}\left(x,\theta,\left(\nu_j\right)_{j=1}^{r}\right):\ \theta\in\mathbb{R},\ x\in X\right\} \tag{13}$$

can only be finite if $\sum_{j=1}^{r} d_j\lambda_j = 1$. Thus, the dual problem of (8), given by (13), becomes (Luenberger 1969)

$$\begin{cases} \text{Max } -\sum_{j=1}^{r} \Upsilon_j \lambda_j \\ \quad + \left(Inf_{x\in X} \left\{ -\sum_{j=1}^{r} \int_{\Delta_{(\rho_j,\varkappa_j)}} \left(\mathbb{E}\left(f(x) z_j\right) + k_j\right) d\nu_j \left(z_j, k_j\right) \right\} \right) \\ \lambda_j = \int_{\Delta_{(\rho_j,\varkappa_j)}} d\nu_j \left(z_j, k_j\right), \ j = 1, 2, \ldots, r \\ \sum_{j=1}^{r} d_j \lambda_j = 1 \\ \nu_j \geq 0, \ j = 1, 2, \ldots, r \end{cases} . \tag{14}$$

$d_j > 0$, $j = 1, 2, \ldots, r$ implies that (8) satisfies the Slater Qualification,[9] so, if (8) (or (7)) is bounded, then the dual problem above is solvable and there is no duality gap (the optimal values of (8) and (14) coincide) (Luenberger 1969). If (7) were unbounded then taking a feasible solution (x, θ) with $\theta < 0$ we would have $\Upsilon \geq \rho \circ f(x) - \theta d > \rho \circ f(x)$, against the election of Υ.

Take $\nu^* = \left(\nu_j^*\right)_{j=1}^{r}$ solving (14) and $\lambda_j^* = \nu_j^*\left(\Delta_{(\rho_j,\varkappa_j)}\right)$, $j = 1, 2, \ldots, r$. Take $\left(z_j^*, k_j^*\right) \in \Delta_{(\rho_j,\varkappa_j)}$, $j = 1, 2, \ldots, r$ satisfying the conditions of the Mean Value Theorem (previous lemma) for

$$\tilde{\nu}_j^* = \frac{\nu_j^*}{\lambda_j^*}$$

if $\lambda_j^* > 0$, and $\left(z_j^*, k_j^*\right) \in \Delta_{(\rho_j,\varkappa_j)}$, if $\lambda_j^* = 0$. According to Luenberger (1969), a (8)-feasible element (x^*, θ^*) solves (8) if and only if

$$-\sum_{j=1}^{r} \lambda_j^* \left(\mathbb{E}\left(f(x^*) z_j^*\right) + k_j^*\right) \leq -\sum_{j=1}^{r} \lambda_j^* \left(\mathbb{E}\left(f(x) z_j^*\right) + k_j^*\right)$$

for $j = 1, 2, \ldots, r$ and every $x \in X$, and

$$\lambda_j^* \left(d_j \theta^* + \mathbb{E}\left(f(x^*) z_j^*\right) + k_j^* + \Upsilon_j\right) = 0, \ j = 1, 2, \ldots, r.$$

Then, if $\lambda_j^* \neq 0$, bearing in mind the constraint of (7) we have

$$\rho_j \circ f\left(x^*\right) \leq \theta^* d_j + \Upsilon_j = -\mathbb{E}\left(f\left(x^*\right) z_j^*\right) - k_j^*,$$

so

$$-\left(k_j^* + \mathbb{E}\left(f\left(x^*\right) z_j^*\right)\right) \geq \rho_j \circ f\left(x^*\right) \geq -\left(k_j + \mathbb{E}\left(f\left(x^*\right) z_j\right)\right)$$

for every $\left(z_j, k_j\right) \in \Delta_{(\rho_j,\varkappa_j)}$, holds from the definition of $\Delta_{(\rho_j,\varkappa_j)}$. □

[9] That is, there is a least one feasible solution of (8) satisfying all the constraints in terms of strict inequalities. Indeed, $d_j > 0$, $j = 1, 2, \ldots, r$ implies that one only have to take a value of θ large enough.

4 Applications

This section is devoted to present two practical applications. The first one may be considered as "classical" in Financial Mathematics, while the second one is "classical" in Actuarial Mathematics. Both lead to optimization problems that perfectly fit on (5), so the theory above absolutely applies. The two examples are very important in practice, but this is far of being an exhaustive list of the real-world issues related to the optimization of risk functions. Another very interesting topics, like pricing and hedging issues, credit risk or operational risk, etc., may be considered.

4.1 Portfolio Choice

The optimal portfolio selection is probably the most famous multiobjective optimization problem in finance. Let us assume that

$$y_1, y_2, \ldots, y_n \in L^p$$

represent the random returns of n available assets,[10] and denote by $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ the portfolio composed of the percentages invested in these assets. If ρ is the ($\mathbb{R}^r$-valued) vector risk function used by the investor then he/she will select that strategy solving

$$\begin{cases} \text{Min } \rho\left(\sum_{i=1}^n x_i y_i\right) \\ \sum_{i=1}^n x_i = 1 \\ \sum_{i=1}^n x_i \mathbb{E}(y_i) \geq r_0 \end{cases} \tag{15}$$

$r_0 \in \mathbb{R}$ denoting the minimum required expected return. If some short-sale restrictions must be imposed then constraints such as $x_i \geq 0$ for some (or all) subscripts must be added. Similarly, additional equality or inequality constraints reflecting several market-linked or agent-linked restrictions may arise. It is obvious that (15) is a particular case of (5).

4.2 Optimal Reinsurance

The "Optimal Reinsurance Problem" is classical in Actuarial Mathematics. Many authors have dealt with it by using different "Premium Principles", and a quite general approach may be found in Kaluszka (2005), where the author uses even some

[10] That is, y_i will be the final pay-off received at a future date $t = T$ if one invests one dollar in the ith-security at the initial date $t = 0$.

coherent measures of risk to price the insurance. However, the minimized risk functions are usually classical deviations (standard deviation or absolute deviation) or classical down side semi-deviations. More recently Cai and Tan (2007) minimize the *Value at Risk* and the *Conditional Tail Expectation* (Artzner et al. 1999) for a very particular case, since they only deal with the *Expected Value Principle* and, more importantly, *stop–loss* reinsurance contracts. We will show below that the general approach of this paper may apply to minimize general risk functions in the optimal reinsurance problem and we do not need to be constrained by any kind of reinsurance contract.

Consider that an insurance company receives the fixed amount S_0 (premium) and will have to pay the random variable $y_0 \in L^p$ within a given period $[0, T]$ (claims). Suppose also that a reinsurance contract is signed in such a way that the company will only pay $x \in L^p$ whereas the reinsurer will pay $y_0 - x$. If the reinsurer premium is given by the convex function,[11]

$$\pi : L^p \longrightarrow \mathbb{R}$$

and π_1 is the highest amount that the insurer would like to pay for the contract, then the insurance company will chose x (optimal retention) so as to solve

$$\begin{cases} \text{Min } \rho\left(S_0 - x - \pi\left(y_0 - x\right)\right) \\ \pi\left(y_0 - x\right) \leq \pi_1 \\ 0 \leq x \leq y_0 \end{cases} \tag{16}$$

ρ being a vector risk function. Notice that

$$x \longrightarrow S - x - \pi\left(y_0 - x\right)$$

is a concave function, so (16) is included in (5) and the developed theory obviously applies.

5 Conclusions

The minimization of risk functions is becoming very important in Mathematical Programming, Mathematical Finance and Actuarial Mathematics, which provokes a growing interest in this topic that is becoming the focus of many researchers.

Since risk functions are not differentiable there are significant difficulties when they are involved in minimization problems. Convex programming and duality

[11] Insurance premiums are usually given by convex functions. See for instance Deprez and Gerber (1985).

methods have been proposed. This paper has also followed this line of research, though there are two major differences. On the one hand, we deal with multiobjective problems, which is far more realistic due to the lack of consensus with respect to the risk function to be used in many applications. Secondly, the provided necessary and sufficient optimality conditions are quite different if one compares with previous literature. Indeed, they are related to saddle point properties of a bilinear function of the feasible set and the sub-gradient of the risk measures to be optimized. This seems to be profound finding whose proof is based on the $weak^*$-compactness of the sub-gradient of the risk measure, the duality theory in general Banach spaces and a given Mean Value Theorem for risk measures. The yielded optimality conditions easily apply in practice. Interesting applications in finance and insurance have been given.

Acknowledgements Research partially developed during the sabbatical visit to Concordia University (Montreal, Quebec, Canada). The authors thank the Department of Mathematics and Statistics for its great hospitality.

Research partially supported by "Welzia Management SGIIC SA", "RD Sistemas SA", "Comunidad Autónoma de Madrid" (Spain), Grant s-0505/tic/000230, and "MEyC" (Spain), Grant SEJ2006-15401-C04. The usual caveat applies.

References

Alexander S, Coleman TF, Li Y (2006) Minimizing $CVaR$ and VaR for a portfolio of derivatives. J Bank Finance 30:538–605

Artzner P, Delbaen F, Eber JM, Heath D (1999) Coherent measures of risk. Math Finance 9:203–228

Balbás A, Balbás R, Mayoral S (2009) Portfolio choice problems and optimal hedging with general risk functions: a simplex-like algorithm. Eur J Oper Res 192(2):603–620

Balbás A, Galperin E, Guerra PJ (2005) Sensitivity of Pareto solutions in multiobjective optimization. J Optim Theor Appl 126(2):247–264

Balbás A, Romera R (2007) Hedging bond portfolios by optimization in Banach spaces. J Optim Theor Appl 132(1):175–191

Cai J, Tan KT (2007) Optimal retention for a stop loss reinsurance under the VaR and CTE risk measures. ASTIN Bull 37(1):93–112

Calafiore GC (2007) Ambiguous risk measures and optimal robust portfolios. SIAM J Optim 18(3):853–877

Cherny AS (2006) Weighted $V@R$ and its properties. Finance Stochastics 10:367–393

Deprez O, Gerber U (1985) On convex principles of premium calculation. Insur Math Econ 4:179–189

Föllmer H, Schied A (2002) Convex measures of risk and trading constraints. Finance Stochastics 6:429–447

Galperin EA, Wiecek MM (1999) Retrieval and use of the balance set in multiobjective global optimization. Comput Math Appl 37:111–123

Goovaerts M, Kaas R, Dhaene J, Tang Q (2004) A new classes of consistent risk measures. Insur Math Econ 34:505–516

Horvàth J (1966) Topological vector spaces and distributions, vol I. Addison Wesley, Reading, MA

Kaluszka M (2005) Optimal reinsurance under convex principles of premium calculation. Insur Math Econ 36:375–398

Luenberger DG (1969) Optimization by vector spaces methods. Wiley, New York

Mansini R, Ogryczak W, Speranza MG (2007) Conditional value at risk and related linear programming models for portfolio optimization. Ann Oper Res 152:227–256

Nakano Y (2004) Efficient hedging with coherent risk measure. J Math Anal Appl 293:345–354

Ogryczak W, Ruszczynski A (2002) Dual stochastic dominance and related mean risk models. SIAM J Optim 13:60–78

Rockafellar RT, Uryasev S, Zabarankin M (2006a) Generalized deviations in risk analysis. Finance Stochastics 10:51–74

Rockafellar RT, Uryasev S, Zabarankin M (2006b) Optimality conditions in portfolio analysis with general deviations measures. Math Program B 108:515–540

Ruszczynski A, Shapiro A (2006) Optimization of convex risk functions. Math Oper Res 31(3):433–452

Schied A (2007) Optimal investments for risk- and ambiguity-averse preferences: a duality approach. Finance Stochastics 11:107–129

Staum J (2004) Fundamental theorems of asset pricing for good deal bounds. Math Finance 14:141–161

Wang SS (2000) A class of distortion operators for pricing financial and insurance risks. J Risk Insur 67:15–36