# Minimizing measures of risk by saddle point conditions

Alejandro Balbás [a,*], Beatriz Balbás [b], Raquel Balbás [c]

[a] *University Carlos III of Madrid, C/ Madrid, 126, 28903 Getafe, Madrid, Spain*
[b] *University Rey Juan Carlos, Paseo Artilleros s/n, 28032 Madrid, Spain*
[c] *Complutense University of Madrid, Department of Actuarial and Financial Economics, 28223 Pozuelo de Alarcón, Madrid, Spain*

ARTICLE INFO

*Article history:*
Received 27 September 2008
Received in revised form 6 April 2010

*MSC:*
90C25
90C46
90C48
91B28
91B30

*Keywords:*
Risk minimization
Saddle point condition
Actuarial and financial applications

ABSTRACT

The minimization of risk functions is becoming a very important topic due to its interesting applications in Mathematical Finance and Actuarial Mathematics. This paper addresses this issue in a general framework. Many types of risk function may be involved. A general representation theorem of risk functions is used in order to transform the initial optimization problem into an equivalent one that overcomes several mathematical caveats of risk functions. This new problem involves Banach spaces but a mean value theorem for risk measures is stated, and this simplifies the dual problem. Then, optimality is characterized by saddle point properties of a bilinear expression involving the primal and the dual variable. This characterization is significantly different if one compares it with previous literature. Furthermore, the saddle point condition very easily applies in practice. Four applications in finance and insurance are presented.

## 1. Introduction

General risk functions are becoming very important in finance and insurance. The study of risk measures beyond the variance has a long history. For example, [1] studied the upper expectations. The paper by [2] about their "Coherent Measures of Risk" gave a new impulse to this topic, and, since then, many authors have extended the discussion. The recent development of new markets and products (insurance or weather-linked derivatives, commodity or energy/electricity derivatives, inflation-linked bonds, equity indexes annuities, hedge funds, etc.), the necessity of managing new types of risk (credit risk, operational risk, etc.), the presence of asymmetries and fat tails, and the (often legal) obligation of providing initial capital requirements have significantly increased the importance of finding proper measures of risk and risk management techniques.[1] Accordingly, the recent literature presents many interesting contributions focusing on new methods for measuring risk. Among others, we have the convex risk measures [4], the consistent risk measures [5], and the general deviations and the expectation bounded risk measures [6].

Many classical actuarial and financial problems lead to optimization problems and have been revisited by using new risk functions. For example, [7,8] analyze the capital allocation problem using general risk functions. Pricing and hedging issues in incomplete markets have also been studied [4,9], as well as optimal reinsurance problems involving the *CVaR* and stop loss reinsurance contracts [10].

Risk functions are almost never differentiable, which makes it rather difficult to provide general optimality conditions. This has caused many authors to look for concrete properties of the special problem they are dealing with in order to find its

* Corresponding author. Tel.: +34 91 624 9636; fax: +34 91 624 9607.
*E-mail addresses:* alejandro.balbas@uc3m.es (A. Balbás), beatriz.balbas@urjc.es (B. Balbás), raquel.balbas@ccee.ucm.es (R. Balbás).

1 The variance is not compatible with the second-order stochastic dominance if asymmetries and/or heavy tails are involved [3].

solutions. Recent approaches by [11,12] 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 [13]. 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. 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 risk measure $\rho$ and the optimization problem we will deal with. Since $\rho$ is not differentiable in general, the optimization problem is not differentiable either, and Section 3 will be devoted to overcome this caveat. We will use the representation theorem of risk measures so as to transform the initial optimization problem into an equivalent one that is differentiable and often linear. This goal is achieved by following and extending an original idea of [14]. However, the new problem involves new infinite-dimensional Banach spaces of $\sigma$-additive measures, which provokes a high degree of complexity when dealing with duality and optimality conditions. Therefore, the mean value theorem (Lemma 4) is one of the most important results of the paper, since it will absolutely simplify the dual problem.[2] As a consequence, Theorem 5 characterizes the optimal solutions by saddle points of a bilinear function of the feasible set and the sub-gradient of the risk measure to be optimized. This seems to be profound finding whose proof is based on major results in functional analysis. Besides, the provided optimality conditions are quite different if one compares them with those of previous literature. They are very general and easily apply in practice.

Section 4 presents four problems of Actuarial and Financial Mathematics that may be studied by minimizing risks. They are optimal portfolio selection, pricing and hedging in incomplete markets, the loaded rate of equity-linked annuities and the optimal reinsurance problem. The novelty is given by the form of some problems, the level of generality of the analysis and the high weakness of the assumptions. The four 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 interesting topic (credit or operational risk, etc.) may be considered. The last section points out the most important conclusions.

## 2. Preliminaries

Consider a probability space $(\Omega, \mathcal{F}, \mu)$ composed of the set $\Omega$ of states of the word, the $\sigma$-algebra $\mathcal{F}$ indicating the information available at a future date $T$, and the probability measure $\mu$. 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 $L^p$. We will deal with a risk function $\rho : L^p \longrightarrow \mathbb{R}$ such that the following condition holds.[3]

*Assumption* I. There exists $\varkappa \in \mathbb{R}$ such that

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

is $\sigma(L^q, L^p)$-compact.[4] □

**Proposition 1.** (a) *The sets* $\Delta^q_{(\rho,\varkappa)}$, *and* $\Delta_{(\rho,\varkappa)}$ *and* $\Delta^{\mathbb{R}}_{(\rho,\varkappa)}$ *given by*

$$\begin{aligned} \Delta_{(\rho,\varkappa)} &= \left\{ (z,k) \in L^q \times (-\infty, \varkappa] ; -\mathbb{E}(yz) - k \leq \rho(y), \forall y \in L^p \right\} \\ \Delta^{\mathbb{R}}_{(\rho,\varkappa)} &= \left\{ k \in \mathbb{R}; (z,k) \in \Delta_{(\rho,\varkappa)} \quad \text{for some } z \in L^q \right\} \end{aligned} \tag{2}$$

*are convex. Moreover,* $\Delta^q_{(\rho,\varkappa)}$ *and* $\Delta^{\mathbb{R}}_{(\rho,\varkappa)}$ *are the projections of* $\Delta_{(\rho,\varkappa)}$ *on* $L^q$ *and* $\mathbb{R}$.

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

**Proof.** (a) is trivial, so let us prove (b). Since the inclusion $\Delta_{(\rho,\varkappa)} \subset \Delta^q_{(\rho,\varkappa)} \times \Delta^{\mathbb{R}}_{(\rho,\varkappa)}$ is obvious, it is sufficient to show that $\Delta^{\mathbb{R}}_{(\rho,\varkappa)}$ is closed and bounded and that $\Delta_{(\rho,\varkappa)}$ is closed. To see that $\Delta^{\mathbb{R}}_{(\rho,\varkappa)}$ is closed, let as assume that $(k_n)_{n\in\mathbb{N}}$ is a sequence in $\Delta^{\mathbb{R}}_{(\rho,\varkappa)}$ that converges to $k \in \mathbb{R}$. Take a sequence $(z_n, k_n)_{n\in\mathbb{N}} \subset \Delta_{(\rho,\varkappa)}$. Since $\Delta^q_{(\rho,\varkappa)}$ 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, $-\mathbb{E}(yz_n) - k_n \leq \rho(y)$ for every $n \in \mathbb{N}$ and every $y \in L^p$ leads to $-\mathbb{E}(yz) - k \leq \rho(y)$ for every $y \in L^p$ and $(z,k) \in \Delta_{(\rho,\varkappa)}$, i.e., $k \in \Delta^{\mathbb{R}}_{(\rho,\varkappa)}$.

To see that the set $\Delta^{\mathbb{R}}_{(\rho,\varkappa)}$ is bounded, it is sufficient to prove that it is bounded from below, since $\varkappa$ is an obvious upper bound. Expression (2) leads to $-\mathbb{E}(0) - k \leq \rho(0)$ for every $k \in \Delta^{\mathbb{R}}_{(\rho,\varkappa)}$, and $\mathbb{E}(0) = 0$ implies that $k \geq -\rho(0)$ for every $k \in \Delta^{\mathbb{R}}_{(\rho,\varkappa)}$.

To see that the set $\Delta_{(\rho,\varkappa)}$ is closed, consider the net $(z_i, k_i)_{i\in I} \subset \Delta_{(\rho,\varkappa)}$ and its limit $(z,k)$. Then, $-\mathbb{E}(yz_i) - k_i \leq \rho(y)$ for every $i \in I$ and every $y \in L^p$ leads to $-\mathbb{E}(yz) - k \leq \rho(y)$ for every $y \in L^p$, so $(z,k) \in \Delta_{(\rho,\varkappa)}$. □

---

2 Actually, Lemma 4 extends a previous mean value theorem given in [15]. The extension permits us to deal with very general risk measures.

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

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

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

$$\Delta^{\mathbb{R}}_{(\rho,\varkappa)} = [\varkappa_0, \varkappa] \subset [-\rho(0), \varkappa]. \tag{3}$$

*Assumption* II. The equality

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

holds for every $y \in L^p$.[5] □

**Proposition 2.** *Under Assumptions* I *and* II, *$\rho$ is a convex function.*

**Proof.** Take $y_1, y_2 \in L^p$ and $t \in [0,1]$. Take $(z,k) \in \Delta_{(\rho,\varkappa)}$ with $\rho((1-t)y_1 + ty_2) = -\mathbb{E}(((1-t)y_1 + ty_2)z) - k$. We have that $\rho((1-t)y_1 + ty_2) = (1-t)(-\mathbb{E}(y_1 z) - k) + t(-\mathbb{E}(y_2 z) - k) \le (1-t)\rho(y_1) + t\rho(y_2)$. □

Consider now a convex subset $X$ included in an arbitrary vector space and a function $f : X \longrightarrow L^p$ such that $\rho \circ f : X \longrightarrow \mathbb{R}$ is convex. Possible examples arise when $f$ is concave and $\rho$ is decreasing (for instance, if $\rho$ is coherent) 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 optimization problem

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

## 3. Saddle point optimality conditions

In general, neither $\rho$ nor $f$ in (5) is differentiable. However, due to Assumptions I and II, and following [14], it is easy to see that (5) is equivalent to

$$\begin{cases} Min\ \theta \\ \theta + \mathbb{E}(f(x)z) + k \ge 0, \quad \forall (z,k) \in \Delta_{(\rho,\varkappa)} \\ \theta \in \mathbb{R}, \quad x \in X \end{cases} \tag{6}$$

in the sense that $x \in X$ solves (5) if and only if there exists $\theta \in \mathbb{R}$ such that $(x,\theta)$ solves (6), in which case $\theta = \rho \circ f(x)$ holds. Notice that $(x,\theta)$ is the decision variable of (6).

The first constraint of (6) is valued in $\mathcal{C}\left(\Delta_{(\rho,\varkappa)}\right)$, the Banach space of the real-valued and continuous functions on the compact space $\Delta_{(\rho,\varkappa)}$. Besides, $\rho \circ f$ being convex, one has that (6) is a convex problem, so the dual variable must belong to the dual space of $\mathcal{C}\left(\Delta_{(\rho,\varkappa)}\right)$ [13], that is, the space $\mathcal{M}\left(\Delta_{(\rho,\varkappa)}\right)$ of $\sigma$-additive and inner regular measures on the Borel $\sigma$-algebra of $\Delta_{(\rho,\varkappa)}$ [17]. Furthermore, the Lagrangian function

$$\mathcal{L} : X \times \mathbb{R} \times \mathcal{M}\left(\Delta_{(\rho,\varkappa)}\right) \longrightarrow \mathbb{R}$$

will be given by

$$\mathcal{L}(x,\theta,\nu) = \theta\left(1 - \int_{\Delta_{(\rho,\varkappa)}} \mathrm{d}\nu(z,k)\right) - \int_{\Delta_{(\rho,\varkappa)}} \mathbb{E}(f(x)z)\,\mathrm{d}\nu(z,k) - \int_{\Delta_{(\rho,\varkappa)}} k\mathrm{d}\nu(z,k). \tag{7}$$

According to [13], $\nu \in \mathcal{M}\left(\Delta_{(\rho,\varkappa)}\right)$ is dual feasible if and only if it is non-negative and $Inf_{(x,\theta)\in X\times\mathbb{R}}\{\mathcal{L}(x,\theta,\nu)\} > -\infty$, in which case this infimum is the dual objective. Thus, if $\nu_q$ and $\nu_{\mathbb{R}}$ are the projections (or marginal probabilities) of $\nu$ on $\Delta^q_{(\rho,\varkappa)}$ and $[\varkappa_0, \varkappa]$, respectively, the dual problem becomes

$$\begin{cases} Max \underset{(x,\theta)\in X\times\mathbb{R}}{Inf} \left\{-\int_{\varkappa_0}^{\varkappa} k\mathrm{d}\nu_{\mathbb{R}}(k) - \int_{\Delta^q_{(\rho,\varkappa)}} \mathbb{E}(f(x)z)\,\mathrm{d}\nu_q(z)\right\} \\ \qquad \nu_{\mathbb{R}} = \pi_{\mathbb{R}}(\nu) \\ \qquad \nu_q = \pi_q(\nu) \\ \qquad \nu \in \mathcal{P}\left(\Delta_{(\rho,\varkappa)}\right) \end{cases} \tag{8}$$

with $\mathcal{P}\left(\Delta_{(\rho,\varkappa)}\right)$ denoting the set of inner regular probability measures on the Borel $\sigma$-algebra of $\Delta_{(\rho,\varkappa)}$ and $\pi_q$ and $\pi_{\mathbb{R}}$ denoting the mentioned projections. Problems (6) and (8) might lead to the so-called "duality gap", since we are facing

[5] Assumptions I and II frequently hold. For instance, they are always fulfilled if $\rho$ is expectation bounded or a general deviation, in the sense of [6] (in which case $\varkappa_0 = \kappa = 0$), and often fulfilled if $\rho$ is coherent [2] or consistent [5]. Furthermore, many convex risk measures [4] also satisfy these assumptions. Particular examples are the absolute deviation, the standard deviation, down side semi-deviations, the *CVaR*, the Wang measure and the dual power transform [16].

infinite-dimensional Banach spaces [13]. Let us see that this caveat does not apply here, along with the necessary and sufficient optimality conditions for both problems.

**Lemma 3.** *Suppose that Problem (6) is bounded.* [6] *The following assertions are fulfilled.*

(a) *The infimum of (6) equals the maximum of (8), and (8) attains its optimal value.*

(b) *If $(x^*, \theta^*)$ is (6)-feasible, then it solves (6) if and only if there exists a (8)-feasible $\nu^*$ with*

$$\begin{cases} \int_{\varkappa_0}^{\varkappa} k \mathrm{d}\nu_{\mathbb{R}}^* (k) + \int_{\Delta^q_{(\rho,\varkappa)}} \mathbb{E}(f(x) z) \mathrm{d}\nu_q^* (z) \\ \quad \leq \int_{\varkappa_0}^{\varkappa} k \mathrm{d}\nu_{\mathbb{R}}^* (k) + \int_{\Delta^q_{(\rho,\varkappa)}} \mathbb{E}\left(f\left(x^*\right) z\right) \mathrm{d}\nu_q^* (z), \quad \forall x \in X \\ \theta^* + \int_{\varkappa_0}^{\varkappa} k \mathrm{d}\nu_{\mathbb{R}}^* (k) + \int_{\Delta^q_{(\rho,\varkappa)}} \mathbb{E}\left(f\left(x^*\right) z\right) \mathrm{d}\nu_q^* (z) = 0. \end{cases} \tag{9}$$

*In such a case $\nu^*$ solves (8).*

**Proof.** (a) According to [13], (a) will hold if the Slater qualification holds, i.e., if there exists a (6)-feasible solution $(x, \theta)$ such that $\theta + \mathbb{E}(f(x) z) + k > 0$ for every $(z, k) \in \Delta_{(\rho,\varkappa)}$. Fix $x \in X$. Since $L^q \times \mathbb{R} \ni (z, k) \longrightarrow \mathbb{E}(f(x) z) + k \in \mathbb{R}$ is a $\tilde{\sigma}$-continuous function and $\Delta_{(\rho,\varkappa)}$ is $\tilde{\sigma}$-compact, the Weierstrass theorem shows that it is sufficient to take $\theta > Max\left\{-\mathbb{E}(f(x) z) - k; (z, k) \in \Delta_{(\rho,\varkappa)}\right\}$.

(b) According to [13], $(x^*, \theta^*)$ solves (6) and $\nu^*$ solves (8) if and only if

$$\begin{cases} \mathcal{L}\left(x^*, \theta^*, \nu^*\right) \leq \mathcal{L}\left(x, \theta, \nu^*\right), & \forall (x, \theta) \in X \times \mathbb{R} \\ \int_{\Delta_{(\rho,\varkappa)}} \left(\theta^* + \mathbb{E}\left(f\left(x^*\right) z\right) + k\right) \mathrm{d}\nu^* (z, k) = 0. \end{cases}$$

Thus, (9) immediately holds due to (7) and $\nu^* \in \mathcal{P}\left(\Delta_{(\rho,\varkappa)}\right)$. □

Problem (8) may be complex in practice since so is the dual variable. However, we will present a "mean value theorem" that will allow us to simplify this dual.

**Lemma 4** (*Mean Value Theorem*)**.** *If $\nu \in \mathcal{P}\left(\Delta_{(\rho,\varkappa)}\right)$, then there exist $z_\nu \in \Delta^q_{(\rho,\varkappa)}$ and $k_\nu \in [\varkappa_0, \varkappa]$ such that $(z_\nu, k_\nu) \in \Delta_{(\rho,\varkappa)}$,*

$$\int_{\Delta^q_{(\rho,\varkappa)}} \mathbb{E}(yz) \mathrm{d}\nu_q (z) = E(yz_\nu) \tag{10}$$

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

$$\int_{\varkappa_0}^{\varkappa} k \mathrm{d}\nu_{\mathbb{R}} (k) = k_\nu. \tag{11}$$

**Proof.** Consider the function $L^p \ni y \longrightarrow \psi(y) = \int_{\Delta^q_{(\rho,\varkappa)}} \mathbb{E}(yz) \mathrm{d}\nu_q(z) \in \mathbb{R}$. It is obvious that $\psi$ is linear, so let us prove that it is also continuous. If $\Delta^q_{(\rho,\varkappa)}$ were bounded, then there would exist $M \in \mathbb{R}$ such that $\|z\|_q \leq M$ for every $z \in \Delta^q_{(\rho,\varkappa)}$. Then the Hölder inequality [13] would lead to $|\mathbb{E}(yz)| \leq \|y\|_p \|z\|_q \leq \|y\|_p M$ for every $y \in L^p$ and every $z \in \Delta^q_{(\rho,\varkappa)}$, and $|\psi(y)| \leq \int_{\Delta^q_{(\rho,\varkappa)}} M \|y\|_p \mathrm{d}\nu_q(z) = M \|y\|_p$ for every $y \in L^p$. Whence $\psi$ would be continuous [13]. Let us see now that $\Delta^q_{(\rho,\varkappa)}$ is bounded. Since it is $\sigma^*$-compact, the set $\left\{\mathbb{E}(yz); z \in \Delta^q_{(\rho,\varkappa)}\right\} \subset \mathbb{R}$ is bounded for every $y \in L^p$ because $L^q \ni z \longrightarrow \mathbb{E}(yz) \in \mathbb{R}$ is $\sigma^*$- continuous. Then the Banach–Steinhaus theorem [17] shows that $\Delta^q_{(\rho,\varkappa)}$ is bounded.

Since $\psi$ is continuous, the Riesz representation theorem [17] shows the existence of $z_\nu \in L^q$ such that (10) holds.

Besides, the inequalities $\varkappa_0 \leq \int_{\varkappa_0}^{\varkappa} k \mathrm{d}\nu_{\mathbb{R}}(k) \leq \varkappa$ are obvious, so the existence of $k_\nu \in [\varkappa_0, \varkappa]$ satisfying (11) is obvious too.

It only remains to show that $(z_\nu, k_\nu) \in \Delta_{(\rho,\varkappa)}$. Indeed, (10) and (11) imply that

$$\begin{aligned} -\mathbb{E}(yz_\nu) - k_\nu &= -\int_{\Delta^q_{(\rho,\varkappa)}} \mathbb{E}(yz) \mathrm{d}\nu_q(z) - \int_{\varkappa_0}^{\varkappa} k \mathrm{d}\nu_{\mathbb{R}}(k) \\ &= \int_{\Delta_{(\rho,\varkappa)}} (-\mathbb{E}(yz) - k) \mathrm{d}\nu(z, k) \leq \int_{\Delta_{(\rho,\varkappa)}} \rho(y) \mathrm{d}\nu(z, k) = \rho(y) \end{aligned}$$

for every $y \in L^p$. □

6 (6) is bounded if and only if (5) is bounded, since the infimum of both problems coincide.

The latter lemma enables us to significantly simplify (8). Consider Problem

$$\begin{cases} Max\ \underset{x\in X}{Inf} \left\{-k - \mathbb{E}\left(f\left(x\right)z\right)\right\} \\ \left(z, k\right) \in \Delta_{(\rho,\varkappa)}. \end{cases} \tag{12}$$

It is obvious that (8) and (12) have the same optimal value and that every solution of (12) also solves (8).[7] Therefore, as stated in the saddle point theorem below, we can consider that (12) is the dual problem of (6).

**Theorem 5** (*Saddle Point Theorem*). *Suppose that Problem (5) is bounded. Then:*
(a) *The infimum of (5) equals the maximum of (12), and (12) is solvable.*
(b) *Take $x^* \in X$. $x^*$ solves (5) if and only if there exists $(z^*, k^*) \in \Delta_{(\rho,\varkappa)}$ such that*

$$\mathbb{E}\left(f\left(x\right)z^*\right) + k^* \leq \mathbb{E}\left(f\left(x^*\right)z^*\right) + k^* \leq \mathbb{E}\left(f\left(x^*\right)z\right) + k \tag{13}$$

*hold for every $x \in X$ and every $(z, k) \in \Delta_{(\rho,\varkappa)}$. In the affirmative case, $(z^*, k^*)$ solves (12) and the optimal value of both problems equals $-\mathbb{E}\left(f\left(x^*\right)z^*\right) - k^*$.*

**Proof.** (a) Lemma 3 implies that the maximum of (8) is achieved and equals the infimum of (5). Take $\nu^*$ solving (8) and $(z^*, k^*) \in \Delta_{(\rho,\varkappa)}$ satisfying the conditions (10) and (11) of Lemma 4. Then, it is obvious that the Dirac delta $\delta_{(z^*,k^*)}$ solves (8) and therefore $(z^*, k^*)$ solves (12). Furthermore, the objective value of (12) on $(z^*, k^*)$ equals the objective value of (8) on $\nu^*$ and the infimum value of (5).

(b) If $x^* \in X$ and $(z^*, k^*) \in \Delta_{(\rho,\varkappa)}$ are the solutions, then the absence of a duality gap leads to

$$\rho\left(f\left(x^*\right)\right) = \underset{x\in X}{Inf}\left\{-k^* - \mathbb{E}\left(f\left(x\right)z^*\right)\right\} \leq -k^* - \mathbb{E}\left(f\left(x^*\right)z^*\right).$$

On the other hand, (4) implies that $\rho\left(f\left(x^*\right)\right) \geq -k - \mathbb{E}\left(f\left(x^*\right)z\right)$ for every $(z, k) \in \Delta_{(\rho,\varkappa)}$. Hence, $k + \mathbb{E}\left(f\left(x^*\right)z\right) \geq k^* + \mathbb{E}\left(f\left(x^*\right)z^*\right)$. Moreover, as shown in the proof of Statement (a), $\delta_{(z^*,k^*)}$ solves (8). If $x \in X$, the first inequality in (9) applied to $\nu^* = \delta_{(z^*,k^*)}$ clearly implies that $k^* + \mathbb{E}\left(f\left(x^*\right)z^*\right) \geq \mathbb{E}\left(f\left(x\right)z^*\right) + k^*$.

Conversely, suppose that $x^* \in X$, $(z^*, k^*) \in \Delta_{(\rho,\varkappa)}$ and (13) holds. Then,

$$-\mathbb{E}\left(f\left(x\right)z^*\right) - k^* \geq -\mathbb{E}\left(f\left(x^*\right)z^*\right) - k^* \tag{14}$$

implies that the (12)-objective on $(z^*, k^*)$ equals $-\mathbb{E}\left(f\left(x^*\right)z^*\right) - k^*$, whereas

$$-\mathbb{E}\left(f\left(x^*\right)z^*\right) - k^* \geq -\mathbb{E}\left(f\left(x^*\right)z\right) - k \tag{15}$$

implies that the (12)-objective on every $(z, k) \in \Delta_{(\rho,\varkappa)}$ is lower and $(z^*, k^*)$ solves (12).

Finally, (14) is equivalent to the first condition in (9) for $\nu^* = \delta_{(z^*,k^*)}$, so $x^*$ will solve (5) if we find $\theta^* \in \mathbb{R}$ that satisfies the second equation in (9). It is sufficient to take $\theta^* = -k^* - \mathbb{E}\left(f\left(x^*\right)z^*\right)$ because $(x^*, \theta^*)$ is (6)-feasible. Indeed, (15) leads to $\theta^* + \mathbb{E}\left(f\left(x^*\right)z\right) + k \geq 0$ for every $(z, k) \in \Delta_{(\rho,\varkappa)}$. □

## 4. Applications

This section is devoted to presenting four applications. The first two may be considered as "classical" in Financial Mathematics, whereas the last two are "classical" in Actuarial Mathematics. All of them lead to optimization problems that perfectly fit on (5). The four 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 topic, such as credit risk or operational risk, may be considered.

### 4.1. Portfolio choice

Optimal portfolio selection is probably the most famous 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, 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 $\rho$ is the risk function used by the investor then he/she will select the strategy solving

$$\begin{cases} Min\ \rho\left(\sum_{i=1}^{n} x_i y_i\right) \\ \sum_{i=1}^{n} x_i = 1 \\ \sum_{i=1}^{n} x_i \mathbb{E}\left(y_i\right) \geq r_0 \end{cases} \tag{16}$$

7 That is, if $(z, k)$ solves (12) then the Dirac delta $\delta_{(z,k)}$ concentrating the whole mass on $(z, k)$ solves (8).

with $r_0 \in \mathbb{R}$ 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. Obviously (16) is a particular case of (5).

### 4.2. Pricing and hedging

The valuation of new securities in an incomplete (and maybe imperfect) financial market is a major issue in Mathematical Finance. Recent literature has used coherent measures of risk to deal with this topic ([9,18], etc.). We will address this problem by drawing on the optimization of risk functions.

Consider a time interval $[0, T]$. Suppose that $Y \subset L^p$ is a convex cone of final (at $T$) reachable pay-offs of a given financial market and that $\pi : Y \longrightarrow \mathbb{R}$ is a pricing rule providing the current price $\pi(y)$ of every $y \in Y$. For instance, $\pi(y)$ may be the infimum initial price of those self-financing portfolios replicating the final pay-off $y$.[8] Assume that $\pi$ is convex.[9] Assume finally that a trader is interested in selling a new asset $g \notin Y$ and simultaneously he/she attempts to hedge this sale. Then, if $r_f \geq 0$ denotes the riskless rate, the trader will choose the hedging strategy $y \in Y$ so as to solve the optimal hedging problem

$$\begin{cases} Min\ \rho(y - g)\, e^{-r_f T} + A \\ \pi(y) \leq A \\ y \in Y, \quad A \in \mathbb{R} \end{cases} \tag{17}$$

with $x = (y, A)$ being the decision variable and $\rho$ being a risk function. If $\rho$ may be understood as an initial capital requirement that prevents possible negative evolutions of the market, and $(y_0, A_0)$ solves (17), then $y_0$ will be the optimal hedging strategy and

$$\rho(y_0 - g)\, e^{-r_f T} + A_0 \tag{18}$$

the ask price of $g$, composed of the ask price of $y_0$ plus the required reserves $\rho(y_0 - g)\, e^{-r_f T}$; i.e., (18) reflects the capital needed by the seller of $g$. (17) is a particular case of (5).

### 4.3. Equity-indexed annuities

Equity-indexed annuities are very important in the insurance industry of Canada and the USA, and they are becoming interesting in Europe too. Their loaded rate is a very important topic closely related to the risk level of the issuer of the product (see, for instance, [20]). We will show that this problem is also related to the minimization of risk functions studied in this article.

Suppose for instance that $T = 1$ year is an initial horizon,[10] and consider $n$ clients of the insurer. The $i$th client will pay the premium $P_i$ at $t = 0$ and will receive the pay-off

$$g_i = \begin{cases} H_i, & \text{if he/she is not alive at } T \\ \alpha P_i g & \text{otherwise,} \end{cases}$$

where $0 < \alpha \leq 1$, and $g$ denotes the (annual) realized return of a chosen index. The insurer is obviously facing a risky investment since $g$ is random and the survival of the clients also reflects uncertainty. It may be worth pointing out that the pay-off $g_i$ above is just a possible one, but many other expressions for $g_i$ are also usual in practice. For instance, there may be guarantees, i.e., $g_i$ may be substituted by $g_i^* = Min\left\{g_i, \tilde{g}_i\right\}$, $\tilde{g}_i$ being a fixed amount known at $t = 0$. Anyway, the general form of our optimization problem will not depend on the expression of $g_i$.

In order to determine the optimal loaded rate notice that this problem may be a particular case of the "optimal hedging problem" studied in the subsection above. We will suppose that we are pricing in an incomplete market. Thus, if $Y \subset L^p$ is the convex cone of those final pay-offs that are attainable at $T$, and $\pi : Y \longrightarrow \mathbb{R}$ is the convex pricing rule that permits us to price those securities in $Y$, then the insurer will solve

$$\begin{cases} Min\ \rho\left(y - \sum_{i=1}^{n} g_i\right) e^{-r_f T} + A \\ \pi(y) \leq A \\ y \in Y, \quad A \in \mathbb{R} \end{cases} \tag{19}$$

8 If there are no imperfections then the absence of arbitrage implies that every self-financing strategy replicating $y$ will have the same initial price.

9 According to [19], amongst others, $\pi$ must usually satisfy more restrictive assumptions, since it often has to be sub-additive and positively homogeneous.

10 We are assuming that the loaded rate is computed once a year.

with $x = (y, A)$ being the decision variable, $r_f \geq 0$ denoting the risk-free annual rate and $\rho$ being a risk function. It is obvious that (19) is a particular case of (17) and, consequently, a particular case of (5). If $(y_0, A_0)$ solves (19), then $y_0$ will be the optimal hedging strategy and $\rho\left(y_0 - \sum_{i=1}^{n} g_i\right) e^{-r_f T} + A_0$ the price of the global portfolio of annuities, which is composed of the price of $y_0$ plus the capital requirement $\rho\left(y_0 - \sum_{i=1}^{n} g_i\right) e^{-r_f T}$. The global loaded rate

$$R = \rho\left(y_0 - \sum_{i=1}^{n} g_i\right) e^{-r_f T} + A_0 - \sum_{i=1}^{n} P_i.$$

may be decomposed into individual loaded rates $R = \sum_{i=1}^{n} R_i$ by using the mortality table and other actuarial methods. Notice that computing first a global loaded rate will significantly reduce the value of the individual ones.

### 4.4. 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 [21], where the author uses even some 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, [10] minimize the *Value at Risk* and the *Conditional Tail Expectation* [2] 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 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] $\pi : L^p \longrightarrow \mathbb{R}$, $\rho$ is the risk function, and $S_1$ is the highest amount to pay for the contract, then the insurer will chose $x$ (optimal retention) so as to solve

$$\begin{cases} Min\ \rho\left(S_0 - x - \pi\left(y_0 - x\right)\right) \\ \pi\left(y_0 - x\right) \leq S_1 \\ 0 \leq x \leq y_0. \end{cases} \quad (20)$$

$x \longrightarrow S_0 - x - \pi\left(y_0 - x\right)$ is a concave function, so (20) is included in (5).

## 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 methods have been proposed, and this paper has also followed this line of research, though the provided necessary and sufficient optimality conditions are quite different if one compares them with previous literature. Indeed, they are related to the saddle point properties of a bilinear function of the feasible set and the sub-gradient of the risk-involved measure. This seems to be profound finding, whose proof is based on the *weak**-compactness of the sub-gradient, the duality theory in general Banach spaces, and a given mean value theorem for risk measures that permits us to simplify many expressions. The yielded optimality conditions easily apply in practice. Four applications in finance and insurance have been given.

## Acknowledgements

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

The authors also thank the anonymous reviewer for her/his valuable comments.

This research was partially supported by "*Welzia Management SGIIC SA, RD_Sistemas SA*", "*Comunidad Autónoma de Madrid*" (Spain), Grant S2009/ESP-1594, and "*MEyC*" (Spain), Grant ECO2009-14457-C04.

The usual caveat applies.

## References

[1] P. Huber, Robust Statistics, Wiley, New York, 1981.
[2] P. Artzner, F. Delbaen, J.M. Eber, D. Heath, Coherent measures of risk, Mathematical Finance 9 (1999) 203–228.
[3] W. Ogryczak, A. Ruszczynski, Dual stochastic dominance and related mean risk models, SIAM Journal of Optimization 13 (2002) 60–78.

[11] Insurance premiums are often given by convex functions [22].

[4] H. Föllmer, A. Schied, Convex measures of risk and trading constraints, Finance & Stochastics 6 (2002) 429–447.
[5] M. Goovaerts, R. Kaas, J. Dhaene, Q. Tang, A new classes of consistent risk measures, Insurance: Mathematics and Economics 34 (2004) 505–516.
[6] R.T. Rockafellar, S. Uryasev, M. Zabarankin, Generalized deviations in risk analysis, Finance & Stochastics 10 (2006) 51–74.
[7] R. Laeven, M. Goovaerts, An optimization approach to the dynamic allocation of economic capital, Insurance: Mathematics and Economics 35 (2004) 299–319.
[8] J. Dhaene, R. Laeven, S. Vanduffel, G. Darkiewicz, M. Goovaerts, Can a coherent risk measure be too subadditive? Journal of Risk and Insurance 75 (2008) 365–386.
[9] Y. Nakano, Efficient hedging with coherent risk measure, Journal of Mathematical Analysis and Applications 293 (2004) 345–354.
[10] J. Cai, K.T. Tan, Optimal retention for a stop loss reinsurance under the *VaR* and *CTE* risk measures, ASTIN Bulletin 37 (1) (2007) 93–112.
[11] R.T. Rockafellar, S. Uryasev, M. Zabarankin, Optimality conditions in portfolio analysis with general deviations measures, Mathematical Programming, Series B 108 (2006) 515–540.
[12] A. Ruszczynski, A. Shapiro, Optimization of convex risk functions, Mathematics of Operations Research 31 (3) (2006) 433–452.
[13] D.G. Luenberger, Optimization by Vector Spaces Methods, John Wiley & Sons, 1969.
[14] A. Balbás, R. Balbás, S. Mayoral, Portfolio choice problems and optimal hedging with general risk functions: a simplex-like algorithm, European Journal of Operational Research 192 (2) (2009) 603–620.
[15] A. Balbás, B. Balbás, A. Heras, Optimal reinsurance with general risk measures, Insurance: Mathematics and Economics 44 (2009) 374–384.
[16] S.S. Wang, A class of distortion operators for pricing financial and insurance risks, Journal of Risk and Insurance 67 (2000) 15–36.
[17] J. Horvàth, Topological Vector Spaces and Distributions, Vol I, Addison Wesley, Reading, MA, 1966.
[18] J. Staum, Fundamental theorems of asset pricing for good deal bounds, Mathematical Finance 14 (2004) 141–161.
[19] A. De Wagenaere, P.P. Wakker, Nonmonotonic Choquet integrals, Journal of Mathematical Economics 36 (1) (2001) 45–60.
[20] J. Barbarin, P. Devolder, Risk measure and fair valuation of an investment guarantee in life insurance, Insurance: Mathematics and Economics 37 (2) (2005) 297–323.
[21] M. Kaluszka, Optimal reinsurance under convex principles of premium calculation, Insurance: Mathematics and Economics 36 (2005) 375–398.
[22] O. Deprez, U. Gerber, On convex principles of premium calculation, Insurance: Mathematics and Economics 4 (1985) 179–189.