VaR as the CVaR sensitivity: Applications in risk optimization

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

[a] *University Carlos III of Madrid. C/ Madrid, 126. 28903 Getafe (Madrid, Spain). alejandro.balbas@uc3m.es.*

[b] *University of Castilla La Mancha. Avda. Real Fábrica de Seda, s/n. 45600 Talavera (Toledo, Spain). beatriz.balbas@uclm.es.*

[c] *University Complutense of Madrid. Somosaguas. 28223 Pozuelo de Alarcón (Madrid, Spain). raquel.balbas@ccee.ucm.es.*

Abstract VaR minimization is a complex problem playing a critical role in many actuarial and financial applications of mathematical programming. The usual methods of convex programming do not apply due to the lack of sub-additivity. The usual methods of differentiable programming do not apply either, due to the lack of continuity. Taking into account that the CVaR may be given as an integral of VaR, one has that VaR becomes a first order mathematical derivative of CVaR. This property will enable us to give accurate approximations in VaR optimization, since the optimization VaR and CVaR will become quite closely related topics. Applications in both finance and insurance will be given.

Key words VaR Optimization, CVaR Sensitivity, Approximation Methods, Optimality Conditions, Actuarial and Financial Applications.

A.M.S. Classification 90C59, 90C30, 91B30, 91G10, 90B50.

J.E.L. Classification C02, C61, G11, G22.

1 Introduction

VaR has many applications in finance and insurance. Risk management, capital requirements, financial reporting, asset allocation, *bonus-malus* systems, optimal reinsurance, etc. just compose a brief list of topics closely related to VaR. Beyond VaR, risk measurement is an open problem provoking a growing interest and discussion in recent years. Since Artzner *et al.* (1999) introduced their coherent measures of risk much more approaches have been proposed. Very important examples are the expectation bounded measures of risk (Rockafellar *et al.,* 2006), consistent risk measures (Goovaerts *et al., 2004*), actuarial risk measures (Goovaerts and Laeven, 2008), indices of riskiness (Aumann and Serrano, 2008, Foster and Hart, 2009, Bali *et al.*, 2011), etc.

The existence of alternative risk measures implies that many risk-linked problems may be studied without dealing with VaR. Moreover, VaR is not sub-additive (Artzner *et al.,* 1999), it is difficult to optimize (Gaivoronski and Pflug, 2005) and it presents some more drawbacks which may recommend to deal with other risk measures such as $CVaR$ (Rockafellar and Uryasev, 2000). Nevertheless, for several reasons VaR still plays a critical role for many practitioners, institutions and researchers. Firstly, regulation (Basel for banks, Solvency for insurers, etc.) still assigns a vital role to VaR. Secondly, VaR never becomes infinity, while the rest of usual risk measures may attain this value. For instance, $CVaR$ becomes infinity for random risks whose expected losses equal infinity too (for instance, positive random variables with unbounded expectation). Infinite values may provoke analytical and mathematical problems quite difficult to overcome, specially if several heavy tails are simultaneously involved (Chavez-Demoulin *et al.,* 2006). Heavy tails are usual in some actuarial topics (Zajdenwebe, 1996), some operational risk topics (Mitra *et al.,* 2015) and other issues. Thirdly, sub-additivity may be undesirable for some actuarial and financial problems, as pointed out by Dhaene *et al.* (2008), who suggested the use of VaR for some merger-linked problems, for instance. Fourthly, for very important financial problems VaR often provides valuable solutions from both theoretical (Basak and Shapiro, 2001, Assa, 2015) and empirical (Annaert *et al.,* 2009) viewpoints, and VaR also facilitates the use probabilities in both the objective function and/or the constraints of several financial optimization problems (Dupacová and Kopa, 2014, Zhao and Xiao, 2016, etc.).

The optimization of VaR is much more complicated than the optimization of other risk measures (Rockafellar and Uryasev, 2000, Larsen *et al.,* 2002, Gaivoronski and Pflug, 2005, Shaw, 2011, Wozabal, 2012, etc.). Since VaR is neither convex nor differentiable, one may face the existence of many local minima, and they may become undetectable by means of the standard optimization methods. There are many and quite different approaches addressing the optimization of VaR (Larsen *et al.,* 2002, Gaivoronski and Pflug, 2005, Shaw, 2011, Wozabal, 2012, etc.). All of them yield interesting algorithms or optimality conditions allowing us to find adequate solutions under different assumptions, but non of them solves the problem in an exhaustive manner. There are many cases which cannot be treated with the existent methodologies.

A very interesting approach may be found in Wozabal *et al.* (2010) and Wozabal (2012). The authors deal with discrete probability spaces composed of finitely many atoms, and they prove that VaR equals the difference of two convex functions. This property allows them to provide efficient optimizing algorithms. Nevertheless, it is easy to show that the property above does not hold for general probability spaces. Since there are many problems involving VaR and continuous random variables (Shaw, 2011, Zhao and Xiao, 2016, etc.), further extensions containing general probability spaces should be welcome.

This paper deals with a very simple idea. If the $CVaR$ (also called $AVaR$, or average value at risk) may be given as an integral of VaR, then VaR must become a first order mathematical derivative of $CVaR$. Consequently, an approximation of VaR must be given by the change in $CVaR$ over the change in level of confidence (or, in other words, by a quotient of increments). Hence, an approximation of VaR must be given by the difference of two convex functionals, and the result of Wozabal (2012) will become true in general probability spaces if one takes a limit.

Ideas above will be formalized in Section 2, where it will be proved that VaR is the limit of the difference of convex functionals. We will also explain why one does not need to take any limit in the discrete case. In Section 3 we will consider a sequence of optimization problems whose objective function has a limit, and we will analyze the relationship between the sequence of solutions and the solution optimizing the limit. As a consequence, we will establish conditions under which the optimization of VaR may be solved by optimizing the difference of two convex functionals. In Section 4 we will focus on a methodology proposed in Balbás *et al.* $(2010a)$ and we will address the minimization of the difference of two convex functionals in arbitrary probability spaces. Several optimality conditions will be found. Applications in finance (optimal investment) and insurance (optimal reinsurance) will be given in Section 5. Though the purpose of Section 5 is merely illustrative, these examples will be general enough, since they will apply in both static and dynamic frameworks and for discrete or continuous price/claim processes. Section 6 will summarize the paper.

2 Preliminaries and notations

We will deal with the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ composed of the set Ω, the $\sigma-$algebra $\mathcal{F}$ and the probability measure $\mathbb{P}$. We can consider $1 \leq p < \infty$ and the space L^p (also denoted by $L^p(\mathbb{P})$ or $L^p(\Omega, \mathcal{F}, \mathbb{P})$) of real-valued random variables y such that $\mathbb{E}(|y|^p) < \infty$, $\mathbb{E}()$ representing the mathematical expectation. Recall that L^q is the dual space of L^p, where $1 < q \leq \infty$, $1/p + 1/q = 1$, and L^∞ is composed of the essentially bounded random variables (Riesz Representation Theorem, Rudin,1987). Recall also that the usual norm of L^p is

$$\|y\|_p := (\mathbb{E}(|y|^p))^{1/p} \tag{1}$$

if $1 \leq p < \infty$ and $\|y\|_\infty := Ess\_Sup(|y|)$, Ess_Sup denoting "essential supremum".

For $1 \leq p \leq p' \leq \infty$ we have that $L^p \supset L^{p'}$. In particular, $L^1 \supset L^p \supset L^\infty$ for every $1 \leq p \leq \infty$. Recall also that for $1 \leq p \leq \infty$ we have that L^p may be endowed with the topology $\sigma(L^p, L^q)$, which is weaker than the norm topology. Furthermore, if $1 < p < \infty$ then every convex, closed and bounded subset of L^p is $\sigma(L^p, L^q)-$compact (Hanhn-Banach's Theorem and Alaoglu's Theorem). If Ω is a finite set then L^p becomes a finite-dimensional space for every $1 \leq p \leq \infty$, $L^p = L^{p'}$ for every $1 \leq p \leq p' \leq \infty$, and all of the introduced topologies of L^p coincide. Further details about Banach spaces of random variables may be found in Rudin (1973), (1987) and Kopp (1984).

The space L^0 containing every real-valued random variable may be endowed with the usual convergence in probability, in which case L^0 becomes a metric (but not Banach) space. The usual distance in L^0 is given by $d(y,z) = \mathbb{E}(Min\ (1, |y-z|))$, and it is known that $L^1 \subset L^0$ (Rudin, 1987).

Finally, we will deal with many topological properties. All of them may be found in Kelly (1955).

Let us fix a confidence level $1 - \mu \in (0,1)$. As usual, for a random variable $y \in L^0$,[1] the

[1] (3) is the usual definition of $VaR_{1-\mu}(y)$ if y represents a future random wealth (or income). In many actuarial and financial applications y represents random capital losses, in which case (3) is replaced by

$$VaR_{1-\mu}(y) := Sup\ \{x \in \mathbb{R};\ \mathbb{P}(y \leq x) < 1 - \mu\}. \tag{2}$$

Throughout this paper we will deal with (3), but a parallel analysis could be implemented for (2).

Value at Risk $VaR_{1-\mu}(y)$ of y is given by

$$VaR_{1-\mu}(y) := -Inf\left\{x \in \mathbb{R};\ \mathbb{P}(y \leq x) > \mu\right\}, \tag{3}$$

and for $y \in L^1 \subset L^0$ the Conditional Value at Risk $CVaR_{1-\mu}(y)$ is

$$CVaR_{1-\mu}(y) := \frac{1}{\mu}\int_0^{\mu} VaR_{1-t}(y)\,dt. \tag{4}$$

According to Rockafellar *et al.* (2006), $CVaR_{1-\mu}(y)$ may be also given by

$$CVaR_{1-\mu}(y) = Max\ \left\{-\mathbb{E}(yz);\ 0 \leq z \leq 1/\mu,\ \mathbb{E}(z) = 1\right\}, \tag{5}$$

and the set

$$\Delta_\mu := \left\{z \in L^\infty;\ 0 \leq z \leq 1/\mu,\ \mathbb{E}(z) = 1\right\}, \tag{6}$$

which does not depend on y, is called the $CVaR_{1-\mu}-$sub-gradient, it is included in L^q for every $1 \leq q \leq \infty$, and it is convex and $\sigma(L^q, L^p)-$compact for every $1 < q \leq \infty$. An obvious implication of (5) is the equality

$$-CVaR_{1-\mu}(y) = Min\ \left\{\mathbb{E}(yz);\ 0 \leq z \leq 1/\mu,\ \mathbb{E}(z) = 1\right\} \tag{7}$$

for every $y \in L^1$. A second implication of (5) is the L^1-norm continuity of the function

$$L^1 \ni y \to CVaR_{1-\mu}(y) \in \mathbb{R}, \tag{8}$$

along with its $\sigma(L^1, L^\infty)-$lower semi-continuity.[2]

Fix $y \in L^1$. It is known that the function

$$(0,1) \ni t \to VaR_{1-t}(y) \in \mathbb{R} \tag{9}$$

[2]It is easy to see that $L^p \ni y \to VaR_{1-\mu}(y) \in \mathbb{R}$ is not continuous if $p = 0$ or $1 \leq p < \infty$. Indeed, take $\mu = 0.1$, $\Omega = (0,1)$, $\mathcal{F}$ the Borel $\sigma-$algebra, and $\mathbb{P}$ the Lebesgue measure. Take the sequence of random variables

$$(0,1) \ni \omega \to y_n(\omega) = \begin{cases} -1, & if\ 0 < \omega < 0.1 + 1/(2n) \\ 0, & otherwise \end{cases}$$

$n = 1, 2, ...,$ and take

$$(0,1) \ni \omega \to y_0(\omega) = \begin{cases} -1, & if\ 0 < \omega < 0.1 \\ 0, & otherwise \end{cases}$$

Then, $Lim_{n\to\infty}(y_n) = y_0$ in the norm topology of L^p for $1 \leq p < \infty$ and in the metric topology of L^0. Besides, $VaR_{1-\mu}(y_0) = 0$ and $VaR_{1-\mu}(y_n) = 1$, for $n > 0$.

is non-increasing, right-continuous and (Lebesgue) integrable in $(0,1)$. Thus, if one considers the function

$$(0,1) \ni \mu \rightarrow \varphi_y(\mu) := \mu CVaR_{1-\mu}(y) \in \mathbb{R}, \tag{10}$$

which may be also given by (see (4))

$$\varphi_y(\mu) = \int_0^{\mu} VaR_{1-t}(y)\, dt,$$

then the First Fundamental Theorem of Calculus guarantees that

$$\varphi_y'^{+}(\mu) = VaR_{1-\mu}(y) \tag{11}$$

for every $\mu \in (0,1)$, $\varphi_y'^{+}$ denoting the right-hand side derivative of φ_y.

For $n \in \mathbb{N}$ "large enough", Expressions (10) and (11) suggest the approximation

$$VaR_{1-\mu}(y) \approx \frac{(\mu + 1/n)\, CVaR_{1-\mu-1/n}(y) - \mu CVaR_{1-\mu}(y)}{1/n},$$

i.e.,

$$VaR_{1-\mu}(y) \approx (n\mu + 1)\, CVaR_{1-\mu-1/n}(y) - n\mu CVaR_{1-\mu}(y). \tag{12}$$

More accurately,

$$VaR_{1-\mu}(y) = Lim_{n\rightarrow\infty}\left((n\mu + 1)\, CVaR_{1-\mu-1/n}(y) - n\mu CVaR_{1-\mu}(y)\right) \tag{13}$$

holds for every $y \in L^1$. Consequently, if $Y \subset L^1$, the optimization problems

$$Min\ \left\{VaR_{1-\mu}(y);\ y \in Y\right\} \tag{14}$$

and

$$Min\ \left\{\frac{n\mu + 1}{n\mu} CVaR_{1-\mu-1/n}(y) - CVaR_{1-\mu}(y);\ y \in Y\right\} \tag{15}$$

could have "similar solutions". Section 3 will be devoted to analyzing several relationships between Problems (14) and (15), and some methods solving Problem (15) will be presented in Section 4.

3 Connecting the optimization of VaR and CVaR

Many relationships between the solution of (14) and the solution of (15) may be proved in a more general setting. Thus, let us give two general lemmas that will apply in our particular framework.

Lemma 1 *Consider a set* A *and a sequence* $(f_n)_{n=0}^{\infty}$ *of real valued functions on* A *(i.e.,* $f_n : A \to \mathbb{R}$, $n = 0, 1, 2, ...$*) such that* $(f_n)_{n=1}^{\infty} \to f_0$*, pointwise convergence on* A*. Consider* $x_n \in A$ *solving* $Min\ \{f_n(x);\ x \in A\}$ *for every* $n \geq 1$*.*

a) $f_0(x) \geq Lim\_Sup_{n\to\infty} f_n(x_n)$ *for every* $x \in A$*,* $Lim\_Sup_{n\to\infty} f_n(x_n)$ *denoting the limit superior of the sequence* $(f_n(x_n))_{n=1}^{\infty}$*.*

b) Consider a vector space E *and a convex cone* C *such that* $A \subset C \subset E$*. Suppose that* $f_n : A \to \mathbb{R}$ *can be extended to* C*,* $n = 0, 1, 2...$ *and becomes positively homogeneous (i.e.,* $f_n(\lambda x) = \lambda f_n(x)$ *for* $x \in C$*,* $\lambda \geq 0$ *and* $n = 0, 1, 2, ...$*). Suppose finally that* Inf $\left\{f_0(x);\ x \in \bigcup_{\lambda \geq 1}(\lambda A)\right\} > -\infty$*. Then,* $f_0(x) \geq Lim\_Sup_{n\to\infty} f_n(x_n)$ *for every* $x \in \bigcup_{\lambda \geq 1}(\lambda A)$*.*

Proof. $a)$ Fix $x \in A$. If $\varepsilon > 0$, it is sufficient to prove the expression $f_0(x) \geq -\varepsilon + Lim\_Sup_{n\to\infty} f_n(x_n)$. Consider $n_0 \in \mathbb{N}$ such that $|f_n(x) - f_0(x)| < \varepsilon$ holds for every $n \geq n_0$. Then, $f_0(x) \geq f_n(x) - \varepsilon \geq f_n(x_n) - \varepsilon$ holds for every $n \geq n_0$.

$b)$ Consider $x \in A$ and let us prove that $f_0(x) \geq 0$. Indeed, otherwise we would have

$$Inf \left\{ f_0(z);\ z \in \bigcup_{\lambda \geq 1} (\lambda A) \right\} \geq Inf\ \{\lambda f_0(x);\ \lambda > 0\} = -\infty,$$

contradicting the assumptions. Since f_0 is positively homogeneous, we have that $f_0(x) \geq 0$ holds for every $x \in \bigcup_{\lambda \geq 1}(\lambda A)$. If $Lim\_Sup_{n\to\infty} f_n(x_n) < 0$ the assertion becomes obvious, so let us assume that $Lim\_Sup_{n\to\infty} f_n(x_n) \geq 0$. For $\lambda \geq 1$ we have

$$\lambda Lim\_Sup_{n\to\infty} f_n(x_n) \geq Lim\_Sup_{n\to\infty} f_n(x_n).$$

Consider $z = \lambda x$ with $x \in A$. Assertion $a)$ implies that $f_0(x) \geq Lim\_Sup_{n\to\infty} f_n(x_n)$. Hence, $f_0(z) = \lambda f_0(x) \geq \lambda Lim\_Sup_{n\to\infty} f_n(x_n) \geq Lim\_Sup_{n\to\infty} f_n(x_n)$. □

Lemma 2 *Consider a set* A *and a sequence* $(f_n)_{n=0}^{\infty}$ *of real valued functions on* A *such that* $(f_n)_{n=1}^{\infty} \to f_0$ *uniformly on* A*. Consider* $x_n \in A$ *solving* $Min\ \{f_n(x);\ x \in A\}$ *for every* $n \geq 1$*.*

a) Suppose that $x_0 \in A$ *and there exists a topology on* A *such that* x_0 *is an agglomeration point of* $(x_n)_{n=1}^{\infty}$ *(in particular, if* $x_0 = Lim_{n\to\infty}(x_n)$*) and* f_0 *is lower semi-continuous at* x_0*. Then,* $f_0(x_0) = Lim\_Sup_{n\to\infty} f_n(x_n)$ *and* x_0 *solves* $Min\ \{f_0(x);\ x \in A\}$*.*

b) (A pseudo-converse of a) also holds). Suppose that $Min\ \{f_0(x)\,;\ x \in A\}$ *is solvable. Then, there exists a topology on* A *such that* f_0 *is lower semi-continuous on* A *and* $(x_n)_{n=1}^{\infty}$ *has an agglomeration point* $x_0 \in A$ *solving* $Min\ \{f_0(x)\,;\ x \in A\}$ *and such that* $f_0(x_0) = Lim\_Sup_{n\to\infty} f_n(x_n)$. *In particular, if the optimization problem* $Min\ \{f_0(x)\,;\ x \in A\}$ *is solvable then* $Lim\_Sup_{n\to\infty} f_n(x_n)$ *is its optimal value.*

c) Consider a vector space E *and a convex cone* C *such that* $A \subset C \subset E$. *Suppose that* $f_n : A \to \mathbb{R}$ *can be extended to* C, $n = 0, 1, 2...$ *and becomes positively homogeneous. Suppose that* $Inf\ \left\{f_0(x)\,;\ x \in \bigcup_{\lambda\geq 1}(\lambda A)\right\} > -\infty$. *Suppose finally that* $x_0 \in A$, *and there exists a topology on* A *such that* x_0 *is an agglomeration point of* $(x_n)_{n=1}^{\infty}$ *and* f_0 *is lower semi-continuous at* x_0. *Then,* $f_0(x_0) = Lim\_Sup_{n\to\infty} f_n(x_n)$ *and* x_0 *solves* $Min\ \left\{f_0(x)\,;\ x \in \bigcup_{\lambda>0}(\lambda A)\right\}$.

Proof. *a)* Consider $\varepsilon > 0$. There exist a neighborhood V of x_0 and $n_0 \in \mathbb{N}$ such that $f_0(x) > f_0(x_0) - \varepsilon$ holds for every $x \in V$ and $|f_n(x) - f_0(x)| < \varepsilon$ holds for every $x \in A$ and every $n \geq n_0$. Consequently, if $x \in V$ and $n \geq n_0$,

$$f_0(x_0) < f_0(x) + \varepsilon < f_n(x) + 2\varepsilon. \tag{16}$$

Besides, there exists a natural number, still denoted by n_0, such that for every $m \geq n_0$ there exists $k \geq m$ such that $x_k \in V$ and, therefore, $f_0(x_0) < f_k(x_k) + 2\varepsilon$. The obvious implication is that $f_0(x_0) \leq Lim\_Sup_{n\to\infty} f_n(x_n) + 2\varepsilon$ and, therefore, $f_0(x_0) \leq Lim\_Sup_{n\to\infty} f_n(x_n)$. Hence, the conclusion trivially follows from Lemma 1*a*.

b) It is easy to see that the family $\{\varnothing\} \cup \left\{f_0^{-1}(\mathbb{R})\right\} \cup \left\{f_0^{-1}(x,\infty)\,;\ x \in \mathbb{R}\right\}$ of subsets of A is a topology on A making f_0 lower semi-continuous. Suppose that A is compact. Then, $(x_n)_{n=1}^{\infty}$ will have an agglomeration point $x_0 \in A$ (Kelly, 1955), which will satisfy $f_0(x_0) = Lim\_Sup_{n\to\infty} f_n(x_n)$ and will solve $Min\ \{f_0(x)\,;\ x \in A\}$ due to Lemma 2*a*. In order to see that A is compact, consider a family of open sets satisfying

$$A \subset \bigcup_x f_0^{-1}(x,\infty) = f_0^{-1}\left(\bigcup_x (x,\infty)\right) = f_0^{-1}(Inf_x x, \infty),$$

where $Inf_x x$ is the obvious infimum. If $a \in A$ and $f_0(a) = Min\ \{f_0(x)\,;\ x \in A\}$, then $a \in f_0^{-1}(Inf_x x, \infty)$ implies that $f_0(a) > Inf_x x$, so there exists $(\tilde{x}, \infty)$ in the given family of open sets such that $f_0(a) > \tilde{x}$. Therefore, $a \in f_0^{-1}(\tilde{x}, \infty)$ and, consequently, $A \subset f_0^{-1}(\tilde{x}, \infty)$ because $f_0(x) \geq f_0(a) > \tilde{x}$ will hold for every $x \in A$.

c) Bearing in mind Lemma 1b, it is sufficient to prove that $f_0(x_0) \leq Lim\_Sup_{n\to\infty} f_n(x_n) + 2\varepsilon$ for every $\varepsilon > 0$. As in the proof of a), there exist a neighborhood $V \subset A$ of x_0 and $n_0 \in \mathbb{N}$ such that $f_0(x) > f_0(x_0) - \varepsilon$ holds for every $x \in V$ and $|f_n(x) - f_0(x)| < \varepsilon$ holds for every $x \in A$ and every $n \geq n_0$. Consequently, if $x \in V$ and $n \geq n_0$ then (16) holds. As in a), for every $m \geq n_0$ there exists $k \geq m$ such that $x_k \in V$ and, therefore, $f_0(x_0) < f_k(x_k) + 2\varepsilon$. □

Next, let us see that solutions of (15) always yield a lower bound for (14).

Proposition 3 *Consider $Y \subset L^1$ and $y_n \in Y$ solving (15) for every $n \in \mathbb{N}$.*

$$VaR_{1-\mu}(y) \geq Lim\_Sup_{n\to\infty} \left((n\mu + 1) CVaR_{1-\mu-1/n}(y_n) - n\mu CVaR_{1-\mu}(y_n) \right) \tag{17}$$

holds for every $y \in Y$.

Proof. This is an obvious consequence of (13) and Lemma 1a. □

Besides, under additional conditions, solutions of (15) lead to solutions of (14).

Theorem 4 *Consider $Y \subset L^1$ and suppose that (13) holds uniformly on Y. Consider $n_0 \in \mathbb{N}$ and $y_n \in Y$ solving (15) for every $n \geq n_0$.*

a) Consider $y_0 \in Y$ and suppose that there exists a topology on Y such that y_0 is an agglomeration point of $(y_n)_{n=n_0}^{\infty}$ and $Y \ni y \to VaR_{1-\mu}(y) \in \mathbb{R}$ is lower semi-continuous at y_0. Then, y_0 solves (14) and $VaR_{1-\mu}(y_0)$ equals the right hand side of (17).

b) If $1 < p < \infty$, $Y \subset L^p$ is convex, closed and bounded, and $Y \ni y \to VaR_{1-\mu}(y) \in \mathbb{R}$ is lower $\sigma(L^p, L^q)$ −semi-continuous on Y, then $(y_n)_{n=n_0}^{\infty}$ has an agglomeration point $y_0 \in Y$ solving (14) and $VaR_{1-\mu}(y_0)$ equals the right hand side of (17).

c) If $(y_n)_{n=n_0}^{\infty}$ has an agglomeration point $y_0 \in Y$ in the L^1−norm topology of Y, then y_0 solves (14) and $VaR_{1-\mu}(y_0)$ equals the right hand side of (17).

Proof. a) is a consequence of Lemma 2a. b) follows from a) if one bears in mind that Y is $\sigma(L^p, L^q)$ −closed (Hanhn-Banach's Theorem, Rudin, 1973) and $\sigma(L^p, L^q)$ −compact

(Alaoglu's Theorem), and therefore $(y_n)_{n=n_0}^{\infty} \subset Y$ has a $\sigma(L^p, L^q)-$agglomeration point $y_0 \in Y$ (Kelly, 1955). Finally, $c)$ follows from $a)$ if one bears in mind that

$$Y \ni y \to (n\mu + 1)\, CVaR_{1-\mu-1/n}(y) - n\mu CVaR_{1-\mu}(y) \in \mathbb{R}$$

is L^1-norm continuous (see (8)), and therefore so is $Y \ni y \to VaR_{1-\mu}(y) \in \mathbb{R}$ due to the uniform convergence of (13) on Y. □

Lemmas 1b and 2c also have interesting implications in VaR optimization.

Theorem 5 *Consider $Y \subset L^1$ and suppose that $Inf\ \left\{VaR_{1-\mu}(y);\ y \in \bigcup_{\lambda \geq 1}(\lambda Y)\right\} > -\infty$. Consider $n_0 \in \mathbb{N}$ and $y_n \in Y$ solving (15) for every $n \geq n_0$.*

a) Problem

$$Min \left\{ VaR_{1-\mu}(y);\ y \in \bigcup_{\lambda \geq 1}(\lambda Y) \right\} \tag{18}$$

is bounded and the right hand side of (17) is a lower bound of its optimal value.

b) Suppose that (13) holds uniformly on Y. Suppose that $y_0 \in Y$, and there exists a topology on Y such that y_0 is an agglomeration point of $(y_n)_{n=1}^{\infty}$ and $Y \ni y \to VaR_{1-\mu}(y) \in \mathbb{R}$ is lower semi-continuous at y_0. Then, y_0 solves (18) and $VaR_{1-\mu}(y_0)$ equals the right hand side of (17).

c) Suppose that (13) holds uniformly on Y. If $1 < p < \infty$, $Y \subset L^p$ is convex, closed and bounded, and $Y \ni y \to VaR_{1-\mu}(y) \in \mathbb{R}$ is lower $\sigma(L^p, L^q)-$semi-continuous on Y, then $(y_n)_{n=n_0}^{\infty}$ has an agglomeration point $y_0 \in Y$ which solves (18) and $VaR_{1-\mu}(y_0)$ equals the right hand side of (17).

d) Suppose that (13) holds uniformly on Y. If $(y_n)_{n=n_0}^{\infty}$ has an agglomeration point $y_0 \in Y$ in the L^1-norm topology of Y, then y_0 solves (18) and $VaR_{1-\mu}(y_0)$ equals the right hand side of (17).

Proof. Both $VaR_{1-\mu}$ and $(n\mu + 1)\, CVaR_{1-\mu-1/n} - n\mu CVaR_{1-\mu}$ are defined on L^1 and are positively homogeneous (Rockafellar and Uryasev, 2000). Hence, $a)$ is a particular case of Lemma 1b, and $b)$ is a particular case of Lemma 2a. Besides, $c)$ and $d)$ follow from $b)$ if one bears in mind the same arguments as in the proofs of Theorem 4b and 4c. □

According to Theorems 4 and 5, it is important to give conditions guaranteeing the uniform convergence of (13).

Proposition 6 *a) Consider a non-increasing (and therefore Lebesgue integrable) function* $f : [a,b] \to \mathbb{R}$*. For* $0 < h \le b - a$*,*

$$0 \le f(a) - \frac{1}{h}\int_a^{a+h} f(t)\,dt \le f(a) - f(a+h).$$

b) Consider $y \in L^1$*. Then, for* $n = 1, 2, 3, \ldots$*,*

$$\begin{aligned} &0 \le VaR_{1-\mu}(y) - \left((n\mu+1)\,CVaR_{1-\mu-1/n}(y) - n\mu CVaR_{1-\mu}(y)\right) \\ &\le VaR_{1-\mu}(y) - VaR_{1-\mu-1/n}(y). \end{aligned} \tag{19}$$

c) Consider $Y \subset L^1$*. If*

$$Lim_{n\to\infty}\left(VaR_{1-\mu-1/n}(y)\right) = VaR_{1-\mu}(y) \tag{20}$$

uniformly on $y \in Y$ *then (13) holds uniformly on* $y \in Y$*.*

Proof. $a)$ $hf(a) \ge \int_a^{a+h} f(t)\,dt \ge hf(a+h)$ because f is non-increasing. Hence, $f(a) \ge \frac{1}{h}\int_a^{a+h} f(t)\,dt \ge f(a+h)$. The first inequality implies that $0 \le f(a) - \frac{1}{h}\int_a^{a+h} f(t)\,dt$, while the second one implies that $-\frac{1}{h}\int_a^{a+h} f(t)\,dt \le -f(a+h)$ and therefore $f(a) - \frac{1}{h}\int_a^{a+h} f(t)\,dt \le f(a) - f(a+h)$.

$b)$ The result trivially follows from $a)$ because for $h = \frac{1}{n}$ we have that (4) leads to

$$\frac{1}{h}\int_\mu^{\mu+h} VaR_{1-t}(y)\,dt = (n\mu+1)\,CVaR_{1-\mu-1/n}(y) - n\mu CVaR_{1-\mu}(y).$$

$c)$ The result trivially follows from (19). □

Remark 1 *Suppose that* Ω *is a finite set. Suppose that there are no elements of* Ω *with null probability. We can consider the set* $\mathcal{C}$ *of couples* (F,ω) *such that* $F \subset \Omega$*,* $\omega \in \Omega$*,* $\omega \notin F$*,* $\sum_{j\in F}\mathbb{P}(\omega_j) \le \mu$ *and* $\mathbb{P}(\omega) + \sum_{j\in F}\mathbb{P}(\omega_j) > \mu$*. Obviously,* $\mathcal{C}$ *is non void and finite (notice that* $F = \varnothing$ *is accepted). Consider a random variable* y*. According to (3),* $VaR_{1-\mu}(y)$ *is characterized by an element* $(F,\omega) \in \mathcal{C}$*. Indeed, consider an order*

$\Omega = \{\omega_1(y), \omega_2(y), ..., \omega_k(y)\}$ on Ω such that $y(\omega_1(y)) \leq y(\omega_2(y)) \leq .... \leq y(\omega_k(y))$, take $F = \{\omega_1(y), \omega_2(y), ..., \omega_j(y)\}$ with $\sum_{i=1}^{j} \mathbb{P}(\omega_i(y)) \leq \mu$ and $\sum_{i=1}^{j+1} \mathbb{P}(\omega_i(y)) > \mu$, and take $\omega = \omega_{j+1}(y)$. Then, $-VaR_{1-\mu}(y) = y(\omega)$. Moreover, for every $\mu + 1/n < \sum_{i=1}^{j+1} \mathbb{P}(\omega_i(y))$ one still has $-VaR_{1-\mu-1/n}(y) = y(\omega)$, i.e., $VaR_{1-\mu}(y) = VaR_{1-\mu-1/n}(y)$. Since $\mathcal{C}$ is finite, there exists $n_0 \in \mathbb{N}$ such that for every $n \geq n_0$ and every $(F, \omega) \in \mathcal{C}$ one has $\mathbb{P}(\omega) + \sum_{j \in F} \mathbb{P}(\omega_j) > \mu + 1/n$. Therefore $VaR_{1-\mu}(y) = VaR_{1-\mu-1/n}(y)$ for every $n \geq n_0$ and every random variable $y \in L^1$ (notice that $L^1 = L^0$ in this particular case). Consequently, (19) implies that (12) holds as an equality for $n \geq n_0$ and every random variable $y \in L^1$. In order words, the sequence of (13) remains constant for every y and every $n \geq n_0$. It is obvious that the uniform convergence of (13) holds in the whole space L^1, and therefore Theorems 4 and 5 apply. Moreover, we do not have to take any limit because Problem (14) and Problem (15) are exactly the same problem for every $Y \subset L^1$ and every $n \geq n_0$. A similar result is proved in Wozabal et al. (2010) and Wozabal (2012) for portfolio choice problems and other optimization problems involving $VaR_{1-\mu}$ and a finite set Ω. Further extensions applying for infinitely many states of nature require the use of limits and the analysis above. □

4 Optimizing the CVaR-linked approximations

This section will be devoted to solving Problem (15). More accurately, for $k > 0$, $\nu > 0$, $1 - \mu - \nu \in (0, 1)$ and $Y \subset L^1$, we will study Problem

$$Min \ \{kCVaR_{1-\mu-\nu}(y) - CVaR_{1-\mu}(y); \ y \in Y\}. \tag{21}$$

Since Rockafellar and Uryasev (2000) presented their famous method to optimize the $CVaR$, many authors have extended the discussion for other risk measures and frameworks (Ruszczynski and Shapiro, 2006, Balbás *et al.,* 2010*a*, etc.). With respect to Problem (21), Wozabal *et al.* (2010) and Wozabal (2012) proposed new procedures applying under discrete probability spaces with finitely many atoms. In order to find optimality conditions for Problem (21) and general probability spaces, we will follow the approach of Balbás *et al.* (2010*a*), since it has proved to be very efficient in both actuarial (Balbás *et al.,* 2015) and financial (Balbás *et al.,* 2010*b*, Balbás *et al.,* 2016*a* or Balbás *et al.,* 2016*b*) applications.

Lemma 7 *Consider $y^* \in Y$. y^* solves (21) if and only if there exist $\theta^* \in \mathbb{R}$ and $z^* \in L^\infty$*

such that (y^*, θ^*, z^*) *solves (see (6))*

$$Min\ \theta + \mathbb{E}(yz) \begin{cases} \theta + k\mathbb{E}(yw) \geq 0, & \forall w \in \Delta_{\mu+\nu} \\ \mathbb{E}(z) = 1 \\ z \geq 0 \\ z \leq 1/\mu \\ y \in Y \\ (y, \theta, z) \in L^1 \times \mathbb{R} \times L^\infty \end{cases} \tag{22}$$

(y, θ, z) *being the decision variable. If so, then we have that* $\theta^* = kCVaR_{1-\mu-\nu}(y^*)$, $\mathbb{E}(y^* z^*) = -CVaR_{1-\mu}(y^*)$, *and the optimal values of both (21) and (22) coincide.*

Proof. Suppose that y^* solves (21), and consider (y, θ, z) (22)-feasible. Take $\theta^* = kCVaR_{1-\mu-\nu}(y^*)$ and $z^* \in \Delta_\mu$ such that (see (6) and (7)) $\mathbb{E}(y^* z^*) = -CVaR_{1-\mu}(y^*)$. (5) and (7) show that (y^*, θ^*, z^*) is (22)-feasible. Since (y, θ, z) is (22)-feasible, (5) and (7) show that $\theta \geq kCVaR_{1-\mu-\nu}(y)$ and $-CVaR_{1-\mu}(y) \leq \mathbb{E}(yz)$. Hence,

$$\theta + \mathbb{E}(yz) \geq kCVaR_{1-\mu-\nu}(y) - CVaR_{1-\mu}(y) \geq$$
$$kCVaR_{1-\mu-\nu}(y^*) - CVaR_{1-\mu}(y^*) = \theta^* + \mathbb{E}(y^* z^*).$$

Conversely, suppose that (y^*, θ^*, z^*) solves (22). If $y \in Y$ then (5) and (7) show existence of $z \in \Delta_\mu$ with $\mathbb{E}(yz) = -CVaR_{1-\mu}(y)$ and the (22)-feasibility of

$$(y, \theta = kCVaR_{1-\mu-\nu}(y), z).$$

Hence,

$$kCVaR_{1-\mu-\nu}(y) - CVaR_{1-\mu}(y) = \theta + \mathbb{E}(yz) \geq \theta^* + \mathbb{E}(y^* z^*). \tag{23}$$

Let us prove that

$$\theta^* = kCVaR_{1-\mu-\nu}(y^*). \tag{24}$$

Indeed, otherwise θ^* could be replaced by $kCVaR_{1-\mu-\nu}(y^*) < \theta^*$ and we would still have a (22)-feasible solution due to (5). Thus, $\theta^* + \mathbb{E}(y^* z^*) > kCVaR_{1-\mu-\nu}(y^*) + \mathbb{E}(y^* z^*)$ would imply a contradiction.

Next let us prove that

$$\mathbb{E}(y^* z^*) = -CVaR_{1-\mu}(y^*). \tag{25}$$

Indeed, otherwise (7) would imply $\mathbb{E}(y^* z^*) > -CVaR_{1-\mu}(y^*)$, and we could find $z^{**} \in \Delta_\mu$ with (y^*, θ^*, z^{**}) (22)-feasible and $\mathbb{E}(y^* z^{**}) = -CVaR_{1-\mu}(y^*) < \mathbb{E}(y^* z^*)$. Thus, $\theta^* + \mathbb{E}(y^* z^*) > \theta^* + \mathbb{E}(y^* z^{**})$ would be a contraction again.

Finally, (23), (24) and (25) imply that y^* solves (21). □

Notice that Problem (22) is linear in the $\theta-$variable and bilinear in the $(y,z)-$variable. In particular, if one fixes y or z, then (22) becomes linear, and therefore it is easy to find its optimality conditions.

Theorem 8 *(Optimality conditions). Suppose that is Y convex and $y^* \in Y$ solves (21). There exists $(w^*, z^*, \alpha, \alpha_0, \alpha_\mu) \in \Delta_{\mu+\nu} \times \Delta_\mu \times \mathbb{R} \times L^1 \times L^1$ such that*

$$\begin{cases} \alpha_0 z^* = 0 \\ \alpha_\mu (1/\mu - z^*) = 0 \\ y^* = \alpha + \alpha_0 - \alpha_\mu \\ \alpha_0 \geq 0, \ \ \alpha_\mu \geq 0 \\ \mathbb{E}(y^*(z^* - kw^*)) \leq \mathbb{E}(y(z^* - kw^*)), \quad \forall y \in Y \\ \mathbb{E}(y^* w^*) \leq \mathbb{E}(y^* w), \quad \forall w \in \Delta_{\mu+\nu} \end{cases} \tag{26}$$

Furthermore, if $1 \leq p \leq \infty$ and $y^* \in L^p$, then $(\alpha_0, \alpha_\mu) \in L^p \times L^p$.

Proof. There exists (y^*, θ^*, z^*) solving (22). Thus, $z^* \in \Delta_\mu$ and (y^*, θ^*) solves the linear problem

$$Min \ \theta + \mathbb{E}(yz^*) \ \begin{cases} \theta + k\mathbb{E}(yw) \geq 0, \quad \forall w \in \Delta_{\mu+\nu} \\ y \in Y \\ (y, \theta) \in L^1 \times \mathbb{R} \end{cases} \tag{27}$$

As in Balbás *et al.* $(2010a)$, the dual problem of (27) is

$$\begin{cases} Max \quad \Gamma(w) = Inf \ \ \{\mathbb{E}(y(z^* - kw)); \ y \in Y\} \\ \qquad\quad w \in \Delta_{\mu+\nu} \end{cases}, \tag{28}$$

there is no duality gap between (27) and (28), and the complementary slackness conditions between (27) and (28) lead to the fifth and sixth condition of (26).

Besides, (θ^*, z^*) solves

$$Min\ \theta + \mathbb{E}\left(y^* z\right) \begin{cases} \theta + k\mathbb{E}\left(y^* w\right) \geq 0, \quad \forall w \in \Delta_{\mu+\nu} \\ \mathbb{E}\left(z\right) = 1 \\ z \geq 0 \\ z \leq 1/\mu \\ (\theta, z) \in \mathbb{R} \times L^\infty \end{cases} \tag{29}$$

As in Balbás *et al.* (2010*a*), the Lagrangian function of (29) is

$$\mathcal{L}\left(z, w, \alpha_0, \alpha_\mu\right) = \mathbb{E}\left(y^* z\right) - k\mathbb{E}\left(y^* w\right) - \int_\Omega z\alpha_0\left(d\omega\right) + \int_\Omega z\alpha_\mu\left(d\omega\right),$$

where $\alpha_0 \geq 0$ and $\alpha_\mu \geq 0$ belong to the dual space of L^∞. $\left(z, w, \alpha_0, \alpha_\mu\right)$ will be (29)-dual feasible if and only if $\mathcal{L}\left(z, w, \alpha_0, \alpha_\mu\right)$ has a finite lower bound in the affine space $\{z \in L^\infty; \mathbb{E}\left(z\right) = 1\}$, which is equivalent to the existence of $\alpha \in \mathbb{R}$ such that

$$\mathbb{E}\left(y^* z\right) - \int_\Omega z\alpha_0\left(d\omega\right) + \int_\Omega z\alpha_\mu\left(d\omega\right) = \alpha\mathbb{E}\left(z\right).$$

for every $z \in L^\infty$. Thus, $y^* - \alpha_0 + \alpha_\mu = \alpha$ proves the third condition of (26) and, moreover, the complementary slackness conditions between (29) and its dual lead to the first a second one. Finally, the first, second a third condition in (26) imply that

$$\alpha_0 = \begin{cases} y^* + \alpha, & z^* < 1/\mu \\ 0, & otherwise \end{cases} \quad and\ \alpha_\mu = \begin{cases} \alpha - y^*, & z^* > 0 \\ 0, & otherwise \end{cases},$$

and therefore $\left(\alpha_0, \alpha_\mu\right) \in L^p \times L^p$ if $y^* \in L^p$. □

5 Examples

Risk optimization plays a critical role in finance and insurance. There are many classical problems very frequently visited and revisited in the literature. We have selected two examples. This is not at all an exhaustive list and we are aware of that, but we just have an illustrative purpose. We would like to show how the theory of sections above may be useful in both actuarial and financial applications. We will not completely solve the selected examples because it is beyond the scope of this paper, whose focus is on the VaR minimization. Nevertheless, we will see how the developed theory may enable us to find adequate solutions.

The first example is actuarial. We have chosen the optimal reinsurance problem (ORP) because it was among the most studied actuarial optimization problems during many years (Kaluszka, 2005, Cai and Tan, 2007, Chi and Tan, 2013, Balbás *et al.,* 2015, Zhuang *et al.*, 2016, etc.). Similarly, our second choice, portfolio selection and asset allocation ($PCAA$), was the focus of many papers during many years too (Shaw, 2011, Dupacová and Kopa, 2014., Zhao and Xiao, 2016, Balbás *et al.,* 2016a, etc.).

5.1 Optimal reinsurance

Consider an insurance company having to pay the random indemnification $u \geq 0$ at a future date T. The company can buy a reinsurance whose retained risk u_r and ceded risk u_c will satisfy $u = u_r + u_c$. The choice of $u_r \geq 0$ and $u_c \geq 0$ is the focus of the ORP. Since the solution is often achieved with a *stop_loss* contract $u_c = (u - U)^+ = Max\ \{u - U, 0\}$ for some $U \geq 0$, which could provoke reinsurer moral hazard, we will prevent the feasibility of this contract by following the approach of Balbás *et al.* (2015) (see also Zhuang *et al.*, 2016).[3] Hence, consider the Banach space X composed of the out of a countable set continuous functions $x : [0, \infty) \rightarrow \mathbb{R}$ with finite lower and upper bound, endowed with its usual norm $\|x\|_\infty := Sup\ \{|x(t)|\,;\ t \geq 0\}$. If the expectation $\mathbb{E}(u)$ and variance σ_u^2 of u satisfy $\mathbb{E}(u) < \infty$ and $\sigma_u^2 < \infty$, $\mathbb{P}$ is the probability measure generated by u on $[0, \infty)$ (*i.e.*, $\mathbb{P}(B)$ is the probability of the event $u \in B$ for every Borel set $B \subset [0, \infty)$) and $J : X \rightarrow L^2(\mathbb{P})$ is given by

$$J(x)(t) = \int_0^t x(s)\, ds,$$

then, it is easy to see that J is well defined, linear and continuous. Indeed, we will have that

$$|J(x)(t)| \leq \int_0^t \|x\|_\infty\, ds = \|x\|_\infty\, t$$

for every $x \in X$ and every $t \geq 0$,

$$\int_0^\infty J(x)^2\, \mathbb{P}(dx) \leq \|x\|_\infty^2 \int_0^\infty t^2 \mathbb{P}(dx) = \|x\|_\infty^2 \left(\sigma_u^2 + \mathbb{E}(u)^2\right) \tag{30}$$

for every $x \in X$, and the result will trivially follow from properties very standard in functional analysis (Rudin, 1973). In practice, and out of Lebesgue null sets, x may be

[3] If the approach of Balbás *et al.* (2015) is not implemented, and one deals with more classical frameworks (Cai and Tan, 2007, Chi and Tan, 2013, etc.), then, under some straightforward modifications, the rest of the example essentially remains the same.

understood as the sensitivity (or first order derivative) of the retained risk with respect to claims (Balbás *et al.,* 2015). In fact, one can identify u with $u(t) = J(1)(t) = t$ for every $t \geq 0$. Thus, if $x \in X$ is the chosen reinsurance contract, then $u_r = J(x)$ and $u_c = J(1-x)$ will be the chosen retained and ceded risk, respectively. The reinsurer may prevent her/his moral hazard by imposing the constraint $x \geq h$ for a selected "threshold of the retained sensitivity" $h \in X$, $h \geq 0$. Obviously, if the reinsurer accepts $stop - loss$ contracts, this constraint becomes irrelevant by selecting $h = 0$. On the contrary, if $stop - loss$ contracts are not accepted, they will become infeasible by choosing $h \geq \varepsilon$ for some $\varepsilon > 0$.

Let $C > 0$ be a loading rate and consider the reinsurance price

$$(1+C)\,\mathbb{E}(u_c) = (1+C)\,\mathbb{E}(J(1-x))$$

computed with the expected value premium principle. Alternative premium principles may be considered too (Kaluszka, 2005, Balbás *et al.,* 2015, etc.), but, as said above, we only attempt to illustrate the interest of Sections 2, 3 and 4. The final wealth of the insurer will be

$$W(x) = \Pi - J(x) - (1+C)\,\mathbb{E}(J(1-x)) \tag{31}$$

Π being the amount of money paid by the insurer clients. If $VaR_{1-\mu}(W(x))$ reflects the insurer risk, then the ORP may become the vector optimization problem

$$\begin{cases} Max & \mathbb{E}(W(x)) \\ Min & VaR_{1-\mu}(W(x)) \\ & x \in X,\ h \leq x \leq 1 \end{cases}$$

Since $C > 0$, $\Pi \in \mathbb{R}$, $(1+C)\,\mathbb{E}(J(1)) \in \mathbb{R}$ and $VaR_{1-\mu}$ is translation invariant (Artzner *et al.,* 1999), (31) implies the equivalence between this problem and

$$\begin{cases} Min & -\mathbb{E}(J(x)) \\ Min & VaR_{1-\mu}(-J(x)) - (1+C)\,\mathbb{E}(J(x)) \\ & x \in X,\ h \leq x \leq 1 \end{cases}$$

As usual in vector optimization, this problem may be solved by means of positive weights. If $W_0 > 0$ is the weight of $\mathbb{E}(J(x))$ and 1 is the weight of $VaR_{1-\mu}(-J(x)) - (1+C)\,\mathbb{E}(J(x))$, then the objective function of ORP will be $VaR_{1-\mu}(-J(x)) - (1+C+W_0)\,\mathbb{E}(J(x))$, and the ORP final version will become (take $W = 1+C+W_0 > 1$ and recall again that $VaR_{1-\mu}$

is translation invariant)

$$\begin{cases} Min & VaR_{1-\mu}\left(W\mathbb{E}\left(J\left(x\right)\right)-J\left(x\right)\right) \\ & x\in X,\ h\leq x\leq 1 \end{cases} \tag{32}$$

Next, let us see that the proposed methodology applies to solve (32). Indeed, first of all notice that (32) is a particular case of (14) if

$$Y=\left\{W\mathbb{E}\left(J\left(x\right)\right)-J\left(x\right);\ x\in X,\ h\leq x\leq 1\right\}.$$

Secondly, $Y\subset L^2\left(\mathbb{P}\right)$ and it is convex and bounded due to (30). Suppose that (13) holds uniformly on Y. Then, the closure of Y will satisfy the conditions of Theorem 4b, and every agglomeration point of the sequence of solutions of (15) will satisfy the conditions of Theorem 4c. In order to see that (13) holds uniformly on Y, let us draw on Proposition 6c and (20). We only have to prove that

$$Lim_{n\to\infty}\left(VaR_{1-\mu-1/n}\left(W\mathbb{E}\left(J\left(x\right)\right)-J\left(x\right)\right)\right)=VaR_{1-\mu}\left(W\mathbb{E}\left(J\left(x\right)\right)-J\left(x\right)\right)$$

uniformly on $x\in X$, $h\leq x\leq 1$. Since $VaR_{1-\nu}$ is translation invariant for every $1-\nu\in(0,1)$, it is sufficient to show that

$$Lim_{n\to\infty}\left(VaR_{1-\mu-1/n}\left(-J\left(x\right)\right)\right)=VaR_{1-\mu}\left(-J\left(x\right)\right) \tag{33}$$

uniformly on $x\in X$, $h\leq x\leq 1$. Notice that $h\leq x\leq 1$ implies that $J\left(x\right)$ and $J\left(1-x\right)$ are co-monotone (Assa and Karai, 2013). Since $VaR_{1-\nu}$ is co-monotone additive for every $1-\nu\in(0,1)$ (Assa and Karai, 2013), we have that

$$\begin{aligned} &VaR_{1-\mu-1/n}\left(-J\left(1\right)\right)=VaR_{1-\mu-1/n}\left(-J\left(x\right)\right)+VaR_{1-\mu-1/n}\left(-J\left(1-x\right)\right) \\ &VaR_{1-\mu}\left(-J\left(1\right)\right)=VaR_{1-\mu}\left(-J\left(x\right)\right)+VaR_{1-\mu}\left(-J\left(1-x\right)\right) \end{aligned}$$

and therefore,

$$\begin{aligned} &VaR_{1-\mu}\left(-J\left(x\right)\right)-VaR_{1-\mu-1/n}\left(-J\left(x\right)\right)= \\ &VaR_{1-\mu}\left(-J\left(1\right)\right)-VaR_{1-\mu}\left(-J\left(1-x\right)\right) \\ &-\left(VaR_{1-\mu-1/n}\left(-J\left(1\right)\right)-VaR_{1-\mu-1/n}\left(-J\left(1-x\right)\right)\right)= \\ &VaR_{1-\mu}\left(-J\left(1\right)\right)-VaR_{1-\mu-1/n}\left(-J\left(1\right)\right) \\ &-\left(VaR_{1-\mu}\left(-J\left(1-x\right)\right)-VaR_{1-\mu-1/n}\left(-J\left(1-x\right)\right)\right)\leq \\ &VaR_{1-\mu}\left(-J\left(1\right)\right)-VaR_{1-\mu-1/n}\left(-J\left(1\right)\right) \end{aligned}$$

because (9) is non-increasing function. Thus, the uniform convergence of (33) trivially follows from the right-continuity of (9) for $y = J(1)$.

Once we know that Proposition 3 and Theorems 4*b* and 4*c* apply, it only remains to verify Theorem 8 and (26). They lead to $(w^*, z^*, \alpha, \alpha_0, \alpha_\mu) \in \Delta_{\mu+\nu} \times \Delta_\mu \times \mathbb{R} \times L^2 \times L^2$, $\tilde{z}^* = z^* - kw^*$, and

$$\begin{cases} \alpha_0 z^* = 0 \\ \alpha_\mu \left(1/\mu - z^*\right) = 0 \\ -J\left(x^*\right) = W\mathbb{E}\left(J\left(x^*\right)\right) + \alpha + \alpha_0 - \alpha_\mu \\ \alpha_0 \geq 0,\ \ \alpha_\mu \geq 0 \\ \mathbb{E}\left(\left(W\mathbb{E}\left(J\left(x^*\right)\right) - J\left(x^*\right)\right)\tilde{z}^*\right) \leq \mathbb{E}\left(\left(W\mathbb{E}\left(J\left(x\right)\right) - J\left(x\right)\right)\tilde{z}^*\right), \qquad \forall h \leq x \leq 1 \\ \mathbb{E}\left(\left(W\mathbb{E}\left(J\left(x^*\right)\right) - J\left(x^*\right)\right)w^*\right) \leq \mathbb{E}\left(\left(W\mathbb{E}\left(J\left(x^*\right)\right) - J\left(x^*\right)\right)w\right), \ \ \forall w \in \Delta_{\mu+\nu} \end{cases} \tag{34}$$

We will not solve System (34) because it would significantly enlarge the paper. Nevertheless, similar systems have been solved in Balbás *et al.* (2015) and Balbás *et al.* (2016*b*), where the authors optimize the $CVaR$ by means of closely related equations.

5.2 Optimal investment

Let us introduce the $PCAA$ by means of the Balbás *et al.* (2010*b*) approach. It is very general because it applies for both static and dynamic frameworks, and it simplifies some aspects by means of the stochastic discount factor (SDF).

Consider a time interval $[0, T]$ and suppose that marketed claims at T are given by random variables on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Suppose that every marketed claim y is in L^2 and that the market is complete, *i.e.*, every $y \in L^2$ is a marketed claim (or reachable pay-off). Completeness is not necessary (Balbás *et al.,* 2010*b*), but it simplifies the exposition and, as said above, we only try to illustrate several possibilities of previous sections. Current prices are given by the linear and continuous function (pricing rule) $L^2 \ni y \to \Pi(y) = \mathbb{E}(z_\Pi y) \in \mathbb{R}$. $z_\Pi \in L^2$ is the SDF and must satisfy $\mathbb{P}(z_\Pi > 0) = 1$ in order to prevent the arbitrage. We will also impose $\mathbb{E}(z_\Pi) = 1$ or, equivalently, the riskless rate vanishes. Once again this assumption may be removed, but it simplifies some notations.

The $PCAA$ will focus on both risk and expected pay-off per invested dollar. Thus, if $R > 1$ is the desired expected return (notice that $R = 1$ can be reached with the riskless security), the $PCAA$ will become

$$Min\ VaR_{1-\mu}(y)\ \begin{cases} \mathbb{E}(y) \geq R,\ \mathbb{E}(z_{\Pi}y) \leq 1 \\ y \in L^2 \end{cases} \tag{35}$$

Since $y = (y - R) + R$, $VaR_{1-\mu}(y) = -R + VaR_{1-\mu}(y - R)$, $\mathbb{E}(y) \geq R \Leftrightarrow \mathbb{E}(y - R) \geq 0$ and $\mathbb{E}(z_{\Pi}y) \leq 1 \Leftrightarrow \mathbb{E}(z_{\Pi}(y - R)) \leq 1 - R$, replacing y with $y - R$ and denoting again by y the decision variable, (35) becomes

$$Min\ VaR_{1-\mu}(y)\ \begin{cases} \mathbb{E}(y) \geq 0,\ \mathbb{E}(z_{\Pi}y) \leq -\alpha \\ y \in L^2 \end{cases} \tag{36}$$

with $\alpha = R - 1 > 0$. Problem (36) is feasible under very weak conditions, but is is often unbounded (Balbás *et al.,* 2010*b*). For instance, it is unbounded for the Black and Scholes pricing model and for many stochastic volatility pricing models. If it is bounded and solvable, then, for M large enough, the solution y^* will satisfy $\|y^*\|_2 \leq M$ (see (1)). Therefore, it will also solve

$$Min\ VaR_{1-\mu}(y)\ \begin{cases} \mathbb{E}(y) \geq 0,\ \mathbb{E}(z_{\Pi}y) \leq -\alpha \\ y \in L^2,\ \ \|y\|_2 \leq M \end{cases} \tag{37}$$

Obviously, the feasible set of (37) is closed, bounded and $\sigma(L^2, L^2)-$compact. Therefore, since λy is trivially (36)-feasible if $\lambda \geq 1$ and y is (37)-feasible, Theorem 5a will apply. Furthermore, according to Proposition 6c, if (13) holds uniformly on the (37)-feasible set, then Theorems 5c and 5d will apply too. The uniform fulfillment of (13) in the (37)-feasible set will not hold in general. Nevertheless, the solvability of (36) will often fail as well. If appropriate constraints are added in (36) so as to recover solvability (for instance, if some bounds for the usual Delta or other Greeks are imposed), then the uniform convergence of (13) will be proved with similar arguments to those used in Example (32). With respect to Theorem 8, as already done in (34) for Problem (32), System (26) is easily adapted to Problem (37). As already said, this section only has illustrative purposes, and we will not present a profound analysis of (36) because it would significantly enlarge the paper content.

6 Conclusion

The optimization of VaR is still very important in finance and insurance, among many other fields. Though there are alternative risk measures with valuable properties, several authors have justified the usefulness of VaR in many applications.

The optimization of VaR is much more complicated than the optimization of other risk measures. Since VaR is neither convex nor differentiable, the standard methods of mathematical programming are frequently difficult to apply. There are many and quite different approaches addressing the optimization of VaR. All of them yield interesting algorithms or optimality conditions, but non of them solves the problem in an exhaustive manner. There are many cases which cannot be treated with the existent methodologies.

This paper has proved that a VaR approximation may be given with a linear combination of two $CVaRs$ with different confidence level. More accurately, VaR is a $CVaR$ derivative, and therefore it is the limit of a sequence of linear combination of $CVaRs$ with different confidence level. This property has been used in order to provide new methods to optimize both VaR and linear combinations of $CVaRs$ in general probability spaces. Applications in finance (optimal investment) and insurance (optimal reinsurance) have been given. They show the practical effectiveness of the provided new methodologies.

Acknowledgments. This research was partially supported by "*Ministerio de Economía*" (Spain, Grant $ECO2012-39031-C02-01$). The usual caveat applies.

References

[1] Annaert, J., S. Van Osselaer and B. Verstraete, 2009. Performance evaluation of portfolio insurance strategies using stochastic dominance criteria. *Journal of Banking & Finance*, 33, 272-280.

[2] Artzner, P., F. Delbaen, J.M. Eber and D. Heath, 1999. Coherent measures of risk. *Mathematical Finance*, 9, 203-228.

[3] Assa, H., 2015. Trade-off between robust risk measurement and market principles. *Journal of Optimization Theory and Applications*, 166, 306-320.

[4] Assa, H. and K.M. Karai, 2013. Hedging, Pareto optimality and good deals. *Journal of Optimization Theory and Applications*, 157, 900-917.

[5] Aumann, R.J. and R. Serrano, 2008. An economic index of riskiness. *Journal of Political Economy*, 116, 810-836.

[6] Balbás, A., B. Balbás and R. Balbás, 2010a. Minimizing measures of risk by saddle point conditions. *Journal of Computational and Applied Mathematics*, 234, 2924-2931.

[7] Balbás, A., B. Balbás and R. Balbás, 2010b. $CAPM$ and $APT-$like models with risk measures. *Journal of Banking & Finance*, 34, 1166–1174.

[8] Balbás, A., B. Balbás and R. Balbás, 2016a. Good deals and benchmarks in robust portfolio selection. *European Journal of Operational Research*, 250, 666 - 678.

[9] Balbás, A., Balbás, B. and R. Balbás, 2016b. Outperforming benchmarks with their derivatives: Theory and empirical evidence. *The Journal of Risk*, forthcoming.

[10] Balbás, A., B. Balbás, R. Balbás and A. Heras, 2015. Optimal reinsurance under risk and uncertainty. *Insurance: Mathematics and Economics*, 60, 61 - 74.

[11] Basak, S. and A. Shapiro, 2001. Value at risk based risk management. *Review of Financial Studies*, 14, 371-405.

[12] Bali, T.G., N. Cakici and F. Chabi-Yo, 2011. A generalized measure of riskiness. *Management Science*, 57, 8, 1406-1423.

[13] Cai, J. and K.S. Tan, 2007. Optimal retention for a stop loss reinsurance under the VaR and CTE risk measures. *ASTIN Bulletin*, 37, 1, 93-112.

[14] Chavez-Demoulin, V., P. Embrechts and J. Neslehová, 2006. Quantitative models for operational risk: Extremes, dependence and aggregation. *Journal of Banking & Finance*, 30, 2635–2658.

[15] Chi, Y. and K.S. Tan, 2013. Optimal reinsurance with general premium principles. *Insurance: Mathematics and Economics*, 52, 180-189.

[16] Dhaene, J., R.J. Laeven, S. Vanduffel, G. Darkiewicz and M.J. Goovaerts, 2008. Can a coherent risk measure be too subadditive? *Journal of Risk and Insurance*, 75, 365-386.

[17] Dupacová, J. and M. Kopa, 2014. Robustness of optimal portfolios under risk and stochastic dominance constraints. *European Journal of Operational Research*, 234, 434 - 441.

[18] Foster, D.P. and S. Hart, 2009. An operational measure of riskiness. *Journal of Political Economy*, 117, 785-814.

[19] Gaivoronski, A. and G. Pflug, 2005. Value at risk in portfolio optimization: Properties and computational approach. *The Journal of Risk*, 7,2, 1 - 31.

[20] Goovaerts, M.J. and R. Laeven, 2008. Actuarial risk measures for financial derivative pricing. *Insurance: Mathematics and Economics*, 42, 540-547.

[21] Goovaerts, M.J., R. Kaas, J. Dhaene and Q. Tang, 2004. A new classes of consistent risk measures. *Insurance: Mathematics and Economics*, 34, 505-516.

[22] Kaluszka, M., 2005. Optimal reinsurance under convex principles of premium calculation. *Insurance: Mathematics and Economics*, 36, 375 - 398.

[23] Kelly, J.L., 1955. *General topology*. Springer.

[24] Kopp, P.E., 1984, *Martingales and stochastic integrals*. Cambridge University Press.

[25] Larsen, N., H. Mausser and S. Uryasev, 2002. *Algorithms for optimization of value-at-risk*. In: Pardalos P, Tsitsiringos V (eds) Financial Engineering, e-Commerce and Supply Chain, Kluwer Academic Publishers, Dordrecht, Netherlands, 129–157.

[26] Mitra, S., A. Karathanasopoulos, G. Sermpinis, C. Dunis and J. Hood, 2015. Operational risk: Emerging markets, sectors and measurement. *European Journal of Operational Research*, 241, 122-132.

[27] Rockafellar, R.T. and S. Uryasev, 2000. Optimization of conditional-value-at-risk. *The Journal of Risk*, 2, 21 – 42.

[28] Rockafellar, R.T., S. Uryasev and M. Zabarankin, 2006. Generalized deviations in risk analysis. *Finance & Stochastics*, 10, 51-74.

[29] Rudin, W., 1973. *Functional analysis*. McGraw-Hill.

[30] Rudin, W., 1987. *Real and complex analysis*. Third Edition. McGraw-Hill, Inc.

[31] Ruszczynski, A. and A. Shapiro, 2006. Optimization of convex risk functions. *Mathematics of Operations Research*, 31, 3, 433-452.

[32] Shaw, W.T., 2011. *Risk, VaR, CVaR and their associated portfolio optimizations when Asset returns have a multivariate Student T distribution*. Available at SSRN: http://ssrn.com/abstract=1772731 or http://dx.doi.org/10.2139/ssrn.1772731.

[33] Wozabal, D., 2012. Value-at-Risk optimization using the difference of convex algorithm. *OR Spectrum*, 34, 861-883.

[34] Wozabal, D., R. Hochreiter and G. Pflug, 2010. A D.C. formulation of value-at-risk constrained optimization. *Optimization*, 59, 377–400.

[35] Zajdenwebe, D., 1996. Extreme values in business interruption insurance. *Journal of Risk and Insurance*, 63, 95-110.

[36] Zhao, P. and Q. Xiao, 2016. Portfolio selection problem with Value-at-Risk constraints under non-extensive statistical mechanics. *Journal of Computational and Applied Mathematics*, 298, 74-91.

[37] Zhuang, S.C., C. Weng, K.S. Tan and H. Assa, 2016. Marginal indemnification function formulation for optimal reinsurance. *Insurance: Mathematics and Economics*, 67, 65-76.