Working Paper n° 11/06

Optimizing measures of risk: A simplex-like algorithm

Alejandro Balbás
Departamento de Economía de la Empresa
Universidad Carlos III de Madrid

Raquel Balbás
PhD Student
Facultad de Ciencias Económicas y Empresariales
Universidad Autónoma de Madrid

Silvia Mayoral
Facultad de Ciencias Económicas y Empresariales
Universidad de Navarra

Optimizing measures of risk: A simplex-like algorithm
Alejandro Balbás, Raquel Balbás and Silvia Mayoral
Working Paper No.11/06
September 2006
JEL No. G11, C02.

ABSTRACT

The minimization of general risk or dispersion measures is becoming more and more important in Portfolio Choice Theory. There are two major reasons. Firstly, the lack of symmetry in the returns of many assets provokes that the classical optimization of the standard deviation may lead to dominated strategies, from the point of view of the second order stochastic dominance. Secondly, but not less important, many institutional investors must respect legal capital requirements, which may be more easily studied if one deals with a risk measure related to capital losses.
This paper proposes a new method to simultaneously minimize several risk or dispersion measures. The representation theorems of risk measures are applied to transform the general risk minimization problem in a minimax problem, and later in a linear programming problem between infinite-dimensional Banach spaces. Then, new necessary and sufficient optimality conditions are stated and a simplex-like algorithm is developed. The algorithm solves the dual (and therefore the primal) problem and provides both optimal portfolios and their sensitivities.
The approach is general enough and does not depend on any particular risk measure, but some of the most important cases are specially analyzed.

Alejandro Balbás
Depto. de Economía de la Empresa
Universidad Carlos III de Madrid
C/ Madrid, 126 – 28903. Getafe. Madrid.
alejandro.balbas@uc3m.es

Raquel Balbás
PhD Student.
Depto. de Financiación e Investigación Comercial
Universidad Autónoma de Madrid
Ciudad Universitaria de Cantoblanco.
28049. Madrid

Silvia Mayoral
Universidad de Navarra
Depto. Métodos Cuantitativos
Campus Universitario
31080 Pamplona
smayoral@unav.es

Optimizing measures of risk: A simplex-like algorithm

Alejandro Balbás,[*]Raquel Balbás[†]and Silvia Mayoral[‡]

Abstract

The minimization of general risk or dispersion measures is becoming more and more important in Portfolio Choice Theory. There are two major reasons. Firstly, the lack of symmetry in the returns of many assets provokes that the classical optimization of the standard deviation may lead to dominated strategies, from the point of view of the second order stochastic dominance. Secondly, but not less important, many institutional investors must respect legal capital requirements, which may be more easily studied if one deals with a risk measure related to capital losses.

This paper proposes a new method to simultaneously minimize several risk or dispersion measures. The representation theorems of risk measures are applied to transform the general risk minimization problem in a minimax problem, and later in a linear programming problem between infinite-dimensional Banach spaces. Then, new necessary and sufficient optimality conditions are stated and a simplex-like algorithm is developed. The algorithm solves the dual (and therefore the primal) problem and provides both optimal portfolios and their sensitivities.

The approach is general enough and does not depend on any particular risk measure, but some of the most important cases are specially analyzed.

Key Words. Risk Measure. Deviation Measure. Portfolio Selection. Infinite-Dimensional Linear Programming. Simplex-Like Method.

AMS Classification. 90C48, 90C47, 90C34.

JEL Classification. G11, C02.

1. Introduction

Modern risk analysis must face two major drawbacks affecting most of the available securities and many investment strategies: Asymmetric returns and fat tails. They recently caused important capital losses, that were difficult to predict when managers and supervisors were using classical risk functions like the standard deviation or

---

[*]Corresponding Author. Universidad Carlos III. CL. Madrid 126. 28903 Getafe (Madrid, Spain). alejandro.balbas@uc3m.es

[†]Universidad Autónoma de Madrid. PhD-Student.

[‡]Universidad de Navarra. Edificio Bibliotecas. 31080 Pamplona (Navarra, Spain) smayoral@unav.es

1

sensitivities of the portfolio value with regard to economic or financial variables. This motivated the group $G - 10$ to create the Basle Committee on Banking Supervision, which provided a minimal set of capital requirements in 1988. Besides, during the late $1980's$ JP Morgan developed a firm-wide VaR system able to predict possible losses under very negative scenarios (although they assumed normal distributions).

Researchers have also focused on this problem and Artzner et al. (1999) published a seminal paper introducing the notion of "coherent measure of risk". They attempted to establish a minimal set of axioms that a risk measure should satisfy in order to adequately reflect capital requirements, and they derived important properties and representation theorems from the axioms. Their analysis has been extended or modified by many authors. Important contributions are, amongst many others, Rockafellar et al. (2006a) and (2006b), where the concepts or "bounded expectation risk measure" and "deviation risk measure" are introduced, studied, related and applied to some portfolio choice linked problems, or Ruszczynski and Shapiro (2006), where interesting portfolio optimization topics are addressed too.

Another line of research focuses on those risk functions compatible with the second order stochastic dominance. For instance, Ogryczak and Ruszczynski (2002) point out that the standard deviation does not satisfy this condition if asymmetric returns are involved, whereas the absolute deviation and semi-deviation really do. This justifies that many authors consider the absolute deviation in portfolio selection problems. For instance, Konno et al. (2005), who can reduce the optimization problem to a linear one because they only consider discrete return distributions (generated from recent samples).

Finally, recent literature has also optimized modern risk functions in order to price and hedge in incomplete or imperfect markets. Interesting papers are Föllmer and Schied (2002) and Nakano (2004), among others.

The optimization of dispersions or risk functions reflecting capital requirements is often complex. Indeed, as pointed out by many papers, one frequently must deal with a non-differentiable problem. Though it is usually convex, the subgradient-linked optimality conditions are not so easily treated in practice. Therefore, authors usually look for an equivalent alternative optimization problem. For instance, this is done by Föllmer and Schied (2002) and Nakano (2004). As said above, Konno et al. (2005) can deal with a linear problem, but they minimize a particular dispersion and involve discrete returns. Benati (2003) also gets a linear problem (that he combines with an integer fractional problem) to optimize the "worst conditional expectation", a coherent measure introduced in Artzner et al. (1999), but discrete sample-linked returns are used once more.

The present article proposes a linear programming approach that applies for many risk functions. Amongst them, many coherent, expectation bounded or deviation risk measures are included. The two crucial keys are the representation theorems of risk

measures and the use of Banach spaces to address those problems with infinitely-many constraints. The representation theorems allow us to transform a general portfolio choice problem into a minimax problem. Minimax problems are usual in finance and often lead to a linear problem (see, for instance, Young, 1998), but there are cases for which there are no equivalent linear formulations (Barber and Copper, 1998). Here we follow an idea of Balbás and Romera (2006), where the authors transform a minimax problem, providing hedging strategies against the interest rate risk, into an infinite-dimensional linear programming problem between Banach spaces.

Our linear programming approach yields new optimality conditions and algorithms. So, the paper outline is as follows. Section 2 presents the basic notations and assumptions, as well as a general portfolio choice vector optimization problem that becomes a minimax one. Section 3 transforms the minimax problem into a dual couple of linear ones between Banach spaces. The involved dual Banach spaces are $\mathcal{C}(\Delta)$ and $\mathcal{M}(\Delta)$, continuous functions and inner regular Borel measures on the compact space Δ. Theorem 2, the most important result in this section, provides necessary and sufficient optimality conditions that show important differences with respect to those of previous literature. In fact, rather than subgradient-linked properties, we draw on the existence of probability measures satisfying appropriate requirements, in the line of the complementary slackness conditions of linear programming. Section 4 focuses on special portfolio choice problems that consider a finite number of available assets in the market. The linear problems become semi-infinite, which makes it easier to formulate the optimality conditions of Theorem 2. Section 5 characterizes the extreme points of the dual problem (Theorem 7) and develops a simplex-like algorithm to solve it. The complementary slackness conditions allow us to obtain the primal solution (optimal portfolio). Thus we have both the optimal portfolio and its sensitivity with respect to the involved parameters. Section 6 particularizes the theory for several risk functions, with special focus on some classical ones and those that can be given by distorting functions (Wang, 2000). These risk functions allow us to construct "alternative indexes", whose combinations with the riskless asset lead to the efficient portfolios in a new framework such that risk levels are measured taking into account asymmetries and fat tails. Section 7 concludes the paper.

2. Preliminaries and notations

Let $(\Omega, \mathcal{F}, \mu)$ be a probability space, $p \in [1, \infty)$ and $q \in (1, \infty]$ such that $(1/p) + (1/q) = 1$, where, as usual, we take the convention $(1/\infty) = 0$. It is well known that $L^q = L^q(\Omega, \mathcal{F}, \mu)$ is the dual space of L^P. Consider the functions $\rho_i : L^p \longmapsto \mathbb{R}$, $i = 1, 2, ..., \gamma$, given by

$$\rho_i(y) = Sup\{-E(yz) : z \in \Delta_i\} \tag{1}$$

each Δ_i being a non-void convex and $\sigma\left(L^q, L^p\right)-$compact subset of L^q and

$$E\left(h\right) = \int_\Omega h\left(\omega\right) d\mu\left(\omega\right) = \int_\Omega hd\mu$$

denoting the mathematical expectation of every random variable $h \in L^1$. Notice that $L^q \ni z \longmapsto E\left(yz\right) \in \mathbb{R}$ is $\sigma\left(L^q, L^p\right)-$continuous for every $y \in L^p$ and, therefore, Sup may be replaced by Max in (1).

Many coherent risk measures, in the sense of Artzner et al. (1999), or expectation bounded risk measures, in the sense of Rockafellar et al. (2006a), satisfy Expression (1), as pointed out in Rockafellar et al. (2006a) and (2006b), where Δ_i is assumed to be $\sigma\left(L^q, L^p\right)-$closed. However, as we will show in Section 6, owing to the Alaoglu's Theorem (see Holmes, 1975) the $\sigma\left(L^q, L^p\right)-$compactness of Δ_i frequently holds in practice.

Besides, Rockafellar et al. (2006a) introduce the notion of deviation measure, and characterize those cases for which

$$D(y) = R(y) + E(y), \ \forall y \in L^p, \tag{2}$$

R and D denoting an expectation bounded risk measure and a deviation measure respectively. Obviously, if R satisfies (1) then so does D once Δ_i is substituted by

$$\Delta_i - \{1\} = \{z - 1; z \in \Delta_i\}. \tag{3}$$

Whence, our analysis may apply for both risk and deviation measures.

Fix $m \in \mathbb{N}$, a set of real numbers $\{b_j\}_{j=1}^m$ and a set of random variables $\{q_j\}_{j=1}^m \subset L^q$. We will deal with the vector optimization problem

$$\left.\begin{array}{l} Min \ \rho(y) \\ E\left(yq_j\right) \le b_j, \quad j = 1, 2, ..., m \\ y \in Y \end{array}\right\} \tag{4}$$

$Y \subset L^p$ being an arbitrary convex cone, and $\rho : L^P \longmapsto \mathbb{R}^\gamma$ denoting the vector function with components ρ_i, $i = 1, 2, ..., \gamma$. (4) may be understood as a portfolio selection problem such that there is no a clear agreement about the most convenient way to measure the risk level.[1] Constraints $E\left(yq_j\right) \le b_j$ (or, equivalently, $-E\left(yq_j\right) \ge$

---

[1] Jouini et al. (2004) have introduced the concept of "vector-valued coherent risk measure" as a set-valued function. We do not draw on this notion here. We just use the vector function ρ because we assume that there is no a clear way to measure the risk level. It is quite coherent with previous literature since there is no a total consensus with regard to the concrete risk measure that must apply for each portfolio selection problem.

$-b_j$) may reflect several practical situations. For example, a maximum amount of money to invest, a minimum required expected pay-off, etc.

Expression (1) points out that ρ is a convex function and, consequently, for every optimal solution y_0 of (4) there exists $\alpha = (\alpha_1, \alpha_2, ..., \alpha_\gamma)$ such that $\alpha \neq 0$, $\alpha \geq 0$ and y_0 solves the scalar optimization problem

$$\left. \begin{array}{l} Min \ \sum_{i=1}^{\gamma} \alpha_i \rho_i(y) \\ E\left(yq_j\right) \leq b_j, \qquad j = 1, 2, ..., m \\ y \in Y \end{array} \right\}. \tag{5}$$

Conversely, if $\alpha_i > 0$, $i = 1, 2, ..., \gamma$, then every solution of (5) also solves (4). Thus, by solving Problem (5), with α arbitrary, we will get the whole set of minimal solutions of (4) and, possibly, some more points. All of them compose the so called "set of weak solutions of (4)". Henceforth we will fix a non-null and non-negative arbitrary vector α, so as to obtain those weak solutions.[2]

3. Minimizing risk measures by linear optimization in Banach spaces: Optimality conditions

In this section we will present several problems whose solution leads to the solution of Problem (5). Furthermore, most of the yielded alternative problems will be linear, which will allow us to characterize their solutions by means of duality relationships and complementary slackness conditions. The dual variables generate economic interpretations applying in practical portfolio choice problems. Finally, a future section will provide a simplex-like algorithm permitting us to solve the dual (and therefore the primal) problem.

Denote by $\Delta = \Delta_1 \times \Delta_2 \times ... \times \Delta_\gamma$ the usual product of $(\Delta_i)_{i=1}^{\gamma}$ endowed with the product of the weak*$-$topologies. Then Δ is obviously convex and compact.

Consider Problem

$$\left. \begin{array}{ll} Min \ \theta \\ \theta + \sum_{i=1}^{\gamma} \alpha_i E\left(yz_i\right) \geq 0, & \forall z = (z_1, z_2, ..., z_\gamma) \in \Delta \\ E\left(yq_j\right) \leq b_j, & j = 1, 2, ..., m \\ \theta \in \mathbb{R}, \ y \in Y \end{array} \right\} \tag{6}$$

$(\theta, y) \in \mathbb{R} \times L^p$ being the decision variable. We have the following result whose proof is very simple and therefore omitted.

---

[2]If a vector optimization problem is solved by using this procedure we are applying the scalarization method. An alternative way is provided by the notion of "balance point" (see Galperin, 1997, or Balbás et al., 2002, for further details). If the balance point approach were used then the remainder of this paper should be appropriately adapted, although the main results would still hold.

Proposition 1. *Suppose that $\theta_0 \in \mathbb{R}$ and $y_0 \in Y$. Then, y_0 solves (5) and*

$$\theta_0 = \sum_{i=1}^{\gamma} \alpha_i \rho_i (y_0)$$

if and only if (θ_0, y_0) solves (6).[3]   ¤

Remark 1. *Problem (6) is equivalent to Problem*

$$\left.\begin{array}{l} Min\ \theta \\ \theta + E\left(yz\right) \geq 0, \quad \forall z \in \Delta^{\alpha} \\ E\left(yq_j\right) \leq b_j, \qquad j = 1, 2, ..., m \\ \theta \in \mathbb{R},\ y \in Y \end{array}\right\}$$

Δ^{α} being the convex and $\sigma\left(L^q, L^p\right) - $compact subset of L^q given by

$$\Delta^{\alpha} = \sum_{i=1}^{\gamma} \alpha_i \Delta_i = \left\{ z \in L^q;\ z = \sum_{i=1}^{\gamma} \alpha_i z_i,\ z_i \in \Delta_i,\ i = 1, 2, ..., \gamma \right\}.$$

This new formulation would simplify the notation of future analyses, although we will focus on Problem (6) because it might be complicated to determine Δ^{α} in practical examples.   ¤

Notice that Problem (6) is linear and its first constraint is established in the Banach space $\mathcal{C}\left(\Delta\right)$ of real valued and continuous functions on the compact space Δ endowed with the supremum norm. According to the Riesz Representation Theorem, its dual space, denoted by $\mathcal{M}\left(\Delta\right)$, is composed of the real valued inner regular and $\sigma-$additive measures on the Borel $\sigma-$algebra of Δ endowed with the norm of the total variation.

The Lagrangian function

$$\mathcal{L}: \mathbb{R} \times L^p \times \mathbb{R}^m \times \mathcal{M}\left(\Delta\right) \longmapsto \mathbb{R}$$

of (6) becomes

$$\begin{array}{l} \mathcal{L}\left(\theta, y, \lambda, \nu\right) = \theta - \int_{\Delta} \left(\theta + \sum_{i=1}^{\gamma} \alpha_i E\left(yz_i\right)\right) d\nu(z) + \sum_{j=1}^{m} \lambda_j \left(E\left(yq_j\right) - b_j\right) \\ = \theta \left(1 - \nu\left(\Delta\right)\right) - \sum_{i=1}^{\gamma} \alpha_i \left(\int_{\Delta} E\left(yz_i\right) d\nu(z)\right) + \sum_{j=1}^{m} \lambda_j \left(E\left(yq_j\right) - b_j\right). \end{array} \tag{7}$$

---

[3]Among others, Shimizu and Aiyoshy (1980) already introduced a new variable in order to simplify a minimax problem. The proposition above just applies this idea.

In order to simplify some expressions, if convenient we will represent by $\nu_i = \pi_i(\nu)$ the standard projection of ν on Δ_i, $i = 1, 2, ..., \gamma$, and it is obvious that

$$\int_\Delta E\left(yz_i\right) d\nu(z) = \int_{\Delta_i} E\left(yz_i\right) d\nu_i(z_i)$$

$i = 1, 2, ..., \gamma$. Moreover, the integration variables z or z_i may be omitted.

According to Anderson and Nash (1987), a couple $(\lambda, \nu) \in \mathbb{R}^m \times \mathcal{M}(\Delta)$ is dual feasible for Problem (6) if and only if $\lambda \geq 0$, $\nu \geq 0$ and

$$Inf\left\{\theta\left(1 - \nu\left(\Delta\right)\right) - \sum_{i=1}^\gamma \alpha_i \left(\int_\Delta E\left(yz_i\right) d\nu\right) + \sum_{j=1}^m \lambda_j \left(E\left(yq_j\right) - b_j\right); (\theta, y) \in \mathbb{R} \times Y\right\} > -\infty,$$

in which case its dual objective value is given by the infimum above. Thus, the dual problem becomes

$$\left.\begin{aligned} & Max \ -\sum_{j=1}^m b_j\lambda_j \\ & \sum_{j=1}^m \lambda_j E\left(yq_j\right) - \sum_{i=1}^\gamma \alpha_i\left(\int_\Delta E\left(yz_i\right) d\nu\right) \geq 0, \quad \forall y \in Y \\ & \nu(\Delta) = 1 \\ & (\lambda, \nu) \in \mathbb{R}^m \times \mathcal{M}(\Delta) \\ & \lambda \geq 0, \ \nu \geq 0 \end{aligned}\right\} \tag{8}$$

$(\lambda, \nu) \in \mathbb{R}^m \times \mathcal{M}(\Delta)$ being the decision variable.

It is trivial to see that the inequality

$$\theta \geq -\sum_{j=1}^m b_j\lambda_j \tag{9}$$

holds whenever (θ, y) and (λ, ν) are (6) and (8)-feasible respectively.

Since (6) and (8) involve infinite-dimensional spaces the absence of duality gap is not guaranteed. To solve this minor drawback we will impose the usual Slater Qualification (see Anderson and Nash, 1987). Furthermore, in practical examples it will not be realistic to assume that the risk level may tend to $-\infty$. Thus, we will consider that (6) is bounded.

Assumption I. There exists $y_0 \in Y$ such that $E\left(y_0 q_j\right) < b_j, j = 1, 2, ..., m$. Moreover, there exists θ^* such that $\theta \geq \theta^*$ whenever (θ, y) is (6)$-$feasible.   ¤

The latter assumption implies that (6) has a finite infimum value and (8) attains its optimal value, *i.e.*, (8) is solvable. We will denote by θ_α the optimal value of both problems.

The second and the last constraints in (8) reveal that (8)-feasible inner regular measures ν have to be probabilities. Hereafter we will denote by

$$\mathcal{P}(\Delta) = \{\nu \in \mathcal{M}(\Delta);\ \nu \geq 0, \nu(\Delta) = 1\}.$$

The Alaoglu's Theorem easily leads to the compactness of $\mathcal{P}(\Delta)$ when endowed with the $\sigma(\mathcal{M}(\Delta), \mathcal{C}(\Delta))-$topology (Holmes, 1975, or Anderson and Nash, 1987). The same notations and comments apply if Δ is replaced by Δ_i, $i = 1, 2, ..., \gamma$.

Notice that the influence of the ν variable in (8) only depends on $\{\nu_i\}_{i=1}^{\gamma}$, in the sense that two probability measures with similar projections have similar role in this optimization problem. Thus, if convenient, the dual variable ν may be substituted by its projections $\{\nu_i\}_{i=1}^{\gamma}$.

The first constraint of (8) involves the primal variable $y \in L^p$. This makes some notations "rather complex", so it is worthwhile to eliminate y. Notice that

$$L^p \ni y \longmapsto \sum_{i=1}^{\gamma} \alpha_i \left(\int_{\Delta_i} E(y z_i)\, d\nu_i \right) \in \mathbb{R}$$

is linear and continuous for every $\nu \in \mathcal{M}(\Delta)$. Thus it may be represented by a vector of L^q that we will denote by

$$\sum_{i=1}^{\gamma} \alpha_i \left(\int_{\Delta_i} E(-)\, d\nu_i \right).$$

So, in order to remove the y variable from (8) we will draw on the notation

$$\sum_{j=1}^{m} q_j \lambda_j - \sum_{i=1}^{\gamma} \alpha_i \left(\int_{\Delta_i} E(-)\, d\nu_i \right) \geq_Y 0. \tag{10}$$

Then, the dual problem becomes

$$\left. \begin{array}{l} Max \ -\sum_{j=1}^{m} b_j \lambda_j \\ \sum_{j=1}^{m} q_j \lambda_j - \sum_{i=1}^{\gamma} \alpha_i \left(\int_{\Delta_i} E(-)\, d\nu_i \right) \geq_Y 0 \\ \nu(\Delta) = 1 \\ (\lambda, \nu) \in \mathbb{R}^m \times \mathcal{M}(\Delta) \\ \lambda \geq 0,\ \nu \geq 0 \end{array} \right\} \tag{11}$$

The absence of duality gap permits us to characterize primal and dual solutions by means of a system of equations and inequalities.

Theorem 2. *Consider $(\theta, y) \in \mathbb{R} \times L^p$ and $(\lambda, \nu) \in \mathbb{R}^m \times \mathcal{P}(\Delta)$. They solve (6) and (11) if and only if they solve the following system of complementary slackness conditions*

$$
\left.
\begin{array}{ll}
\sum_{j=1}^m \lambda_j E\left(yq_j\right) - \sum_{i=1}^\gamma \alpha_i \left(\int_{\Delta_i} E\left(yz_i\right) d\nu_i\right) = 0 & \\
\lambda_j \left(E\left(yq_j\right) - b_j\right) = 0, & j = 1, 2, ..., m \\
\theta + \sum_{i=1}^\gamma \alpha_i \left(\int_{\Delta_i} E\left(yz_i\right) d\nu_i\right) = 0 & \\
\theta + \sum_{i=1}^\gamma \alpha_i E\left(yz_i\right) \geq 0, & \forall z \in \Delta \\
E\left(yq_j\right) \leq b_j, & j = 1, 2, ..., m \\
\sum_{j=1}^m q_j \lambda_j - \sum_{i=1}^\gamma \alpha_i \left(\int_{\Delta_i} E\left(-\right) d\nu_i\right) \geq_Y 0 & \\
y \in Y, \lambda \geq 0 &
\end{array}
\right\}
\quad (12)
$$

Proof. The four inequalities of the system above point out that both (θ, z) and (λ, ν) are feasible. Besides, the first equality gives

$$
-\sum_{j=1}^m \lambda_j E\left(yq_j\right) = -\sum_{i=1}^\gamma \alpha_i \left(\int_{\Delta_i} E\left(yz_i\right) d\nu_i\right).
$$

From the second equality,

$$
-\sum_{j=1}^m \lambda_j b_j = -\sum_{i=1}^\gamma \alpha_i \left(\int_{\Delta_i} E\left(yz_i\right) d\nu_i\right).
$$

From the third equality,

$$
-\sum_{j=1}^m \lambda_j b_j = \theta
$$

and (9) shows that we are facing the solutions of both problems.

Conversely, suppose that (θ, y) and (λ, ν) solve both problems. Then they must be feasible and hence they must satisfy the four inequalities. On the other hand, the complementary slackness conditions of linear programming (Anderson and Nash, 1987) lead to the second and third expressions of (12). Finally, (θ, y) must minimize the Lagrangian Function (7) on $\mathbb{R} \times Y$ (Anderson and Nash, 1987). Since the dual constrains trivially show that the minimal value of (7) is attained at $(0, 0)$ and reaches the value $-\sum_{j=1}^m \lambda_j b_j$, we easily obtain the first equality of (12).   ¤

Remark 2. *System (12) provides necessary and sufficient optimality conditions for Problems (5) and (6) that do not use the subgradients of the risk vector measure ρ.*

On the contrary, these conditions draw on the existence of inner regular probability measures $\{\nu_i\}_{i=1}^{\gamma}$ *on* $\{\Delta_i\}_{i=1}^{\gamma}$ *whose supports are contained on the zeros of*

$$\theta + \sum_{i=1}^{\gamma} \alpha_i E\left(yz_i\right),$$

and multipliers $\{\lambda_j\}_{j=1}^{m}$ *such that (10) holds. This fact may be an interesting alternative to solve many portfolio choice problems in practice, although it will not be easy to achieve the explicit solution of System (12).[4] However, to deal with the system becomes far easier if the dual solution may be computed by means of alternative procedures. This is the reason why in Section 5 we will develop a simplex-like method for the dual problem that applies for those cases such that the convex cone* Y *is generated by a finite set of available securities.* ¤

4. Semi-infinite problems

In this section we will consider $n = r + s \in \mathbb{N}$ available assets whose pay-offs are $\{y_h\}_{h=1}^{n} \subset L^p$. Strategies will be represented by $x \in \mathbb{R}^n$, and the closed convex cone

$$X = \left\{ x = (x_h)_{h=1}^{r+s} \in \mathbb{R}^{r+s}; \; x_h \geq 0, \; h = r+1, r+2, ..., r+s \right\}$$

will reflect that short-sales of some securities are not allowed.[5]

---

[4]Recall that dealing with subgradients may also generate significant problems in practice. See Rockafellar et al. (2006a) or Ruszczynski and Shapiro (2006) for complementary analyses.

[5]The cone X above may be replaced by a more general convex closed cone $\tilde{X} \subset \mathbb{R}^n$. Then, the Representation Theorem of closed convex subsets of $\mathbb{R}^n$ (Holmes, 1975) shows that $\tilde{X} = L_{\tilde{X}} + C_{\tilde{X}}$,, $L_{\tilde{X}}$ being the vector space

$$L_{\tilde{X}} = \left\{ x \in \mathbb{R}^n; \; x + \tilde{X} = \tilde{X} \right\},$$

and $C_{\tilde{X}}$ being the closed pointed convex cone (or recession cone)

$$C_{\tilde{X}} = \left\{ x \in \mathbb{R}^n; \; x + \left(\tilde{X} \cap L_{\tilde{X}}^{\vert} \right) \subset \tilde{X} \cap L_{\tilde{X}}^{\vert} \right\},$$

$L_{\tilde{X}}^{\vert}$ representing the orthogonal subspace of $L_{\tilde{X}}$. In the most important cases one has that

$$L_{\tilde{X}} = \mathbb{R}^n \text{ and } C_{\tilde{X}} = \{0\}$$

for $\tilde{X} = \mathbb{R}^n$ and

$$L_{\tilde{X}} = \{0\} \text{ and } C_{\tilde{X}} = \mathbb{R}_+^n$$

for $\tilde{X} = \mathbb{R}_+^n$. More generally,

$$L_{\tilde{X}} = \{0\} \text{ and } C_{\tilde{X}} = \tilde{X}$$

if $\tilde{X}$ is pointed. By using the above representation of $\tilde{X}$ and applying the procedure proposed in Balbás and Romera (2006), we can adapt the variables of our problem so that a $X-$like cone can play the role of $\tilde{X}$.

The set Y (of reachable pay-offs) will be given by

$$Y = \left\{ y \in L^p;\ y = \sum_{h=1}^{n} x_h y_h,\ x = (x_h)_{h=1}^{n} \in X \right\}.$$

It is easy to verify that Y is a convex cone.

Problem (6) becomes

$$\left.\begin{array}{ll}
Min\ \theta \\
\theta + \sum_{h=1}^{n} \left(\sum_{i=1}^{\gamma} \alpha_i E\left(y_h z_i\right) \right) x_h \geq 0, & \forall z \in \Delta \\
\sum_{h=1}^{n} E\left(y_h q_j\right) x_h \leq b_j, & j = 1, 2, ..., m \\
\theta \in \mathbb{R},\ x_h \geq 0 & h = r+1, ..., r+s
\end{array}\right\} \qquad (13)$$

$(\theta, x) \in \mathbb{R}^{1+n}$ being the decision variable. Notice that linear constraints with the form $\sum_{h=1}^{n} l_h x_h \leq L$ are easily incorporated in (13). Indeed, take $\tilde{q} \in L^q$ such that $E\left(y_h \tilde{q}\right) = l_h,\ h = 1, 2, ..., n$, and apply the second constraint in (13). Furthermore, the existence of $\tilde{q}$ holds under weak conditions like the linear independence of $\{y_h\}_{h=1}^{n}$. In order to prevent the existence of duality gaps we will still assume the fulfillment of the Slater Qualification along with the existence of primal lower bounds (Assumption I).

In order to adapt the dual problem (11) we need to analyze the constraint (10). Manipulating we have

$$\sum_{h=1}^{n} E\left[\left(\sum_{j=1}^{m} q_j \lambda_j \right) y_h \right] x_h \geq \sum_{h=1}^{n} \left[\sum_{i=1}^{\gamma} \alpha_i \left(\int_{\Delta_i} E\left(y_h z_i\right) d\nu_i \right) \right] x_h$$

for every $x \in X$. Then, (11) obviously becomes

$$\left.\begin{array}{ll}
Max\ -\sum_{j=1}^{m} b_j \lambda_j \\
\sum_{j=1}^{m} E\left(q_j y_h\right) \lambda_j = \sum_{i=1}^{\gamma} \alpha_i \left(\int_{\Delta_i} E\left(y_h z_i\right) d\nu_i \right), & h = 1, 2, ..., r \\
\sum_{j=1}^{m} E\left(q_j y_h\right) \lambda_j \geq \sum_{i=1}^{\gamma} \alpha_i \left(\int_{\Delta_i} E\left(y_h z_i\right) d\nu_i \right), & h = r+1, ..., r+s \\
(\lambda, \nu) \in \mathbb{R}^m \times \mathcal{P}\left(\Delta\right) \\
\lambda \geq 0
\end{array}\right\} \qquad (14)$$

$(\lambda, \nu) \in \mathbb{R}^m \times \mathcal{P}\left(\Delta\right)$ being the decision variable.

According to Theorem 2, the complementary slackness conditions (12) are necessary and sufficient optimality conditions. In our case they are

$$
\left.
\begin{array}{l}
\sum_{j=1}^{m} \lambda_j E\left(y_h q_j\right) - \sum_{i=1}^{\gamma} \alpha_i \left(\int_{\Delta_i} E\left(y_h z_i\right) d\nu_i\right) = 0, \qquad h = 1, 2, ..., r \\
\left[\sum_{j=1}^{m} \lambda_j E\left(y_h q_j\right) - \sum_{i=1}^{\gamma} \alpha_i \left(\int_{\Delta_i} E\left(y_h z_i\right) d\nu_i\right)\right] x_h = 0, \quad h = r+1, ..., r+s \\
\lambda_j \left[\sum_{h=1}^{n} E\left(y_h q_j\right) x_h - b_j\right] = 0, \qquad\qquad\qquad\qquad j = 1, 2, ..., m \\
\theta + \sum_{h=1}^{n} \left[\sum_{i=1}^{\gamma} \alpha_i \left(\int_{\Delta_i} E\left(y_h z_i\right) d\nu_i\right)\right] x_h = 0
\end{array}
\right\}
$$
(15)

along with the constraints of both problems. As already said, it may be difficult in practice to solve the system above in the unknowns θ, x, λ and ν. However, if λ and ν are known, then (15) becomes

$$
\left.
\begin{array}{l}
\left[\sum_{j=1}^{m} \lambda_j E\left(y_h q_j\right) - \sum_{i=1}^{\gamma} \alpha_i \left(\int_{\Delta_i} E\left(y_h z_i\right) d\nu_i\right)\right] x_h = 0, \quad h = r+1, ..., r+s \\
\sum_{h=1}^{n} E\left(y_h q_j\right) x_h = b_j, \qquad\qquad\qquad j = 1, 2, ..., m, \; \lambda_j \neq 0 \\
\theta + \sum_{h=1}^{n} \left[\sum_{i=1}^{\gamma} \alpha_i \left(\int_{\Delta_i} E\left(y_h z_i\right) d\nu_i\right)\right] x_h = 0
\end{array}
\right\},
$$
(16)

which is a simple linear system in θ and x.

5. The simplex-like algorithm

This section is devoted to present a simplex-like algorithm so as to solve the semi-infinite linear programming problem (14). Then, (16) will generate the primal solution too.

First of all we will introduce the slackness variables $(\xi_h)_{h=r+1}^{r+s}$ in order to get equality constraints. Thus, consider the equivalent Problem

$$
\left.
\begin{array}{l}
Max \; -\sum_{j=1}^{m} b_j \lambda_j \\
\sum_{j=1}^{m} E\left(q_j y_h\right) \lambda_j - \sum_{i=1}^{\gamma} \alpha_i \left(\int_{\Delta_i} E\left(y_h z_i\right) d\nu_i\right) = 0, \qquad h = 1, 2, ..., r \\
\sum_{j=1}^{m} E\left(q_j y_h\right) \lambda_j - \xi_h - \sum_{i=1}^{\gamma} \alpha_i \left(\int_{\Delta_i} E\left(y_h z_i\right) d\nu_i\right) = 0, \quad h = r+1, ..., r+s \\
(\lambda, \xi, \nu) \in \mathbb{R}^m \times \mathbb{R}^s \times \mathcal{P}(\Delta) \\
\lambda \geq 0, \xi \geq 0
\end{array}
\right\}.
$$
(17)

As said above, $\mathcal{P}(\Delta)$ is convex and $\sigma\left(\mathcal{M}(\Delta), \mathcal{C}(\Delta)\right)$ −compact. Besides, given $z \in \Delta$ we will denote by $\delta_z \in \mathcal{P}(\Delta)$ the usual Dirac delta that concentrates the mass on $\{z\}$, i.e., $\delta_z(\{z\}) = 1$ and $\delta_z(\Delta \setminus \{z\}) = 0$. It is known that the set of extreme points of $\mathcal{P}(\Delta)$ is given by

$$
ext\left(\mathcal{P}(\Delta)\right) = \{\delta_z; z \in \Delta\},
$$

though we will not have to draw on this result. Similar properties hold if Δ is substituted by Δ_i, $i = 1, 2, ..., \gamma$.

Lemma 3. *Denote by $\mathcal{A}$ the feasible set of (17) and by*

$$\mathcal{A}_0 = \{(\lambda, \xi) \in \mathbb{R}^m \times \mathbb{R}^s; \text{ there exists } \nu \in \mathcal{P}(\Delta) \text{ with } (\lambda, \xi, \nu) \in \mathcal{A}\}.$$

Then, $\mathcal{A}_0$ is convex and closed.

Proof. It is easy to see that $\mathcal{A}_0$ is convex, so let us prove that it is closed. Indeed, take the sequence $\{(\lambda_k, \xi_k)\}_{k=1}^{\infty} \subset \mathcal{A}_0$ and suppose that

$$Lim_{k \longmapsto \infty} \ (\lambda_k, \xi_k) = (\lambda, \xi).$$

We must prove that $(\lambda, \xi) \in \mathcal{A}_0$. Take $\{(\nu_k)\}_{k=1}^{\infty} \subset \mathcal{P}(\Delta)$ such that $(\lambda_k, \xi_k, \nu_k) \in \mathcal{A}$. Since $\mathcal{P}(\Delta)$ is compact there exists ν, agglomeration point of $\{(\nu_k)\}_{k=1}^{\infty}$. Therefore, (λ, ξ, ν) is an agglomeration point of $\{(\lambda_k, \xi_k, \nu_k)\}_{k=1}^{\infty}$. Since the points of this sequence are in $\mathcal{A}$ and $\mathcal{A}$ is closed, (λ, ξ, ν) is in $\mathcal{A}$ too. Hence, $(\lambda, \xi) \in \mathcal{A}_0$.   ¤

Lemma 4. *Consider the sets $\mathcal{A}_0$ and $\mathcal{A}$ above and their extreme sets, $ext(\mathcal{A}_0)$ and $ext(\mathcal{A})$. Then, there exists $(\lambda, \xi, \nu) \in ext(\mathcal{A})$ such that $- \sum_{j=1}^{m} b_j \lambda_j = \theta_\alpha$, i.e., there exist $(\lambda, \xi, \nu) \in ext(\mathcal{A})$ solving (17). Moreover, $(\lambda, \xi) \in ext(\mathcal{A}_0)$.*[6]

Proof. The previous lemma shows that $\mathcal{A}_0$ is convex and closed. Consequently, the Representation Theorem of closed convex subsets apply (Holmes, 1975). Since $\mathcal{A}_0$ is included in the non-negative cone of $\mathbb{R}^m \times \mathbb{R}^s$ it does not contain any affine manifold, and therefore

$$\mathcal{A}_0 = Co\left[ext(\mathcal{A}_0)\right] + Rc(\mathcal{A}_0), \tag{18}$$

Co denoting convex hulls and Rc denoting recession cones.

Assumption I guarantees that (17) is solvable, and, accordingly, (18) ensures the existence of $(\lambda, \xi) \in ext(\mathcal{A}_0)$ such that $- \sum_{j=1}^{m} b_j \lambda_j = \theta_\alpha$. Fix (λ, ξ) and set

$$\mathcal{A}_1 = \{\nu \in \mathcal{P}(\Delta); \ (\lambda, \xi, \nu) \in \mathcal{A}\}.$$

It is easy to check that $\mathcal{A}_1$ is closed, and therefore compact because it is included in $\mathcal{P}(\Delta)$. The Krein Milman Theorem (Holmes, 1975, or Anderson and Nash, 1987) ensures that $\mathcal{A}_1$ is the $\sigma(\mathcal{M}(\Delta), \mathcal{C}(\Delta))$ −closed convex hull of its extreme set. In particular, there exists $\nu \in ext(\mathcal{A}_1)$, and (λ, ξ, ν) is the required element of $ext(\mathcal{A})$. ¤

---

[6] Given a linear programming problem between infinite-dimensional Banach spaces, if the feasible set is weakly compact and the problem is bounded, then there exists a extreme point of the feasible set where the objective function is as close as desired to the optimal value (Anderson and Nash, 1987). However, since our objective does not depend on the infinite-dimensional variable ν, we can prove much more.

Lemma 5. *Let be $N \in \mathbb{N}$ and $C \subset \mathbb{R}^N$ a convex compact set. Let c_0 be a extreme point of C. Then, there exists a linear map $L : \mathbb{R}^N \longmapsto \mathbb{R}$ with $L(c_0) > L(c)$ for every $c \in C \setminus \{c_0\}$.*

Proof. Let us apply the induction method on the dimension of C, $Dim(C)$ (*i.e.*, the dimension of the minimal affine manifold containing C). The result is clear if $Dim(C)$ equals zero or one. Suppose that it also holds for dimensions $1, 2, ..., Dim(C) - 1$. Without loss of generality we can assume that $Dim(C) = N$. The Separation Theorems for convex sets (Holmes, 1975) point out the existence of a non-null linear map $L' : \mathbb{R}^N \longmapsto \mathbb{R}$ with $L'(c_0) \geq L'(c)$ for every $c \in C$. Obviously,

$$Dim \ \{c \in C; \ L'(c) = L'(c_0)\} \leq N - 1$$

(this set is included in a $(N-1)-$dimensional affine manifold). Due to the induction hypothesis it is easy to establish the existence of $L'' : \mathbb{R}^N \longmapsto \mathbb{R}$ with $L''(c_0) \geq L''(c)$ for every $c \in C$ and $L''(c_0) > L''(c)$ whenever $c \in C$ and $L'(c_0) = L'(c)$. Obviously, $L = L' + L''$ is the required linear map. ¤

Lemma 6. *Let be $N \in \mathbb{N}$ and $f : \Delta \longmapsto \mathbb{R}^N$ a $\sigma(L^q, L^p)-$continuous function. Then, for every $\nu \in \mathcal{P}(\Delta)$ there exist a finite subset*

$$\left\{z^1, z^2, ...z^K\right\} \subset \Delta$$

and a linear convex combination

$$\sum_{k=1}^{K} t_k \delta_{z^k}, \ t_1, t_2, ...t_K \geq 0, \ \sum_{k=1}^{K} t_k = 1$$

such that

$$\int_\Delta f d\nu = \int_\Delta f d \left(\sum_{k=1}^{K} t_k \delta_{z^k}\right).$$

Proof. Since f is continuous then $f(\Delta)$ is connected and compact. Consider the function

$$\mathcal{M}(\Delta) \ni \nu \longrightarrow \phi_f(\nu) = \int_\Delta f d\nu \in \mathbb{R}^N.$$

ϕ_f is clearly linear and $\sigma(\mathcal{M}(\Delta), \mathcal{C}(\Delta))-$continuous and therefore $\phi_f(\mathcal{P}(\Delta))$ is convex and compact. Furthermore, $f(z) = \int_\Delta f d\delta_z$ implies that

$$f(\Delta) \subset \phi_f(\mathcal{P}(\Delta)). \tag{19}$$

We will present the complete proof in two steps.

Step1. The result holds for $N = 1$ with $K = 1$. Indeed, in such a case take

$$m = Min \ \{f(z); \ z \in \Delta\} \leq Max \ \{f(z); \ z \in \Delta\} = M.$$

Since $f(\Delta)$ is connected we have that

$$f(\Delta) = [m, M]. \tag{20}$$

Moreover

$$m = \int_{\Delta} m d\nu \leq \int_{\Delta} f d\nu \leq \int_{\Delta} M d\nu = M$$

holds for every $\nu \in \mathcal{P}(\Delta)$, from where (20) leads to the existence of $z_{\nu} \in \Delta$ with $f(z_{\nu}) = \int_{\Delta} f d\nu$ for every $\nu \in \mathcal{P}(\Delta)$.

Step2. The result holds for every $N \in \mathbb{N}$. Indeed, since $\phi_f(\mathcal{P}(\Delta))$ is convex and compact we have that

$$\phi_f(\mathcal{P}(\Delta)) = Co\left(ext\left(\phi_f(\mathcal{P}(\Delta))\right)\right).$$

Thus, for each $\nu \in \mathcal{P}(\Delta)$ there exists a finite set $\{u_1, u_2, ..., u_K\} \subset ext\left(\phi_f(\mathcal{P}(\Delta))\right)$ generating $\int_{\Delta} f d\nu$ as a linear convex combination

$$\int_{\Delta} f d\nu = \sum_{h=1}^{K} t_h u_h.$$

It is sufficient to see that every element in the set $\{u_1, u_2, ..., u_K\}$ above takes the form $u_k = f(z^k)$, $z^k \in \Delta$, so fix $u \in ext\left(\phi_f(\mathcal{P}(\Delta))\right)$ and let us prove the latter expression. From the previous lemma there exists a linear function $L : \mathbb{R}^N \longmapsto \mathbb{R}$ with

$$L(u) > L(w) \tag{21}$$

for every $w \in \phi_f(\mathcal{P}(\Delta))$, $w \neq u$. Obviously,

$$\int_{\Delta} (L \circ f) \, d\nu = L\left(\int_{\Delta} f d\nu\right)$$

for every $\nu \in \mathcal{P}(\Delta)$, so

$$\int_{\Delta} (L \circ f) \, d\nu_u = L(u)$$

for those $\nu_u \in \mathcal{P}(\Delta)$ with $\int_{\Delta} f d\nu_u = u$ (whose existence follows from $u \in \phi_f(\mathcal{P}(\Delta))$). According to the results stated in Step 1 there exists $z \in \Delta$ with

$$L(f(z)) = (L \circ f)(z) = \int_{\Delta} (L \circ f) \, d\nu_u = L(u)$$

from where, bearing in mind that $f(z) \in f(\Delta) \subset \phi_f(\mathcal{P}(\Delta))$ (see (20) and (21)), we have that $f(z) = u$. ¤

Theorem 7. *With the notations above let (λ, ξ) be a extreme point of $\mathcal{A}_0$ and let (λ, ξ, ν) be a solution of (17) and a extreme point of $\mathcal{A}$. Then, there exist $\left\{z^1, z^2, ...z^K\right\} \subset \Delta$ and a linear convex combination*

$$\sum_{k=1}^{K} t_k \delta_{z^k}, \ t_1, t_2, ...t_K \geq 0, \ \sum_{k=1}^{K} t_k = 1,$$

such that $\left(\lambda, \xi, \sum_{k=1}^{K} t_k \delta_{z^k}\right)$ also solves (17) and is a extreme point of $\mathcal{A}$. Furthermore, if K^* is the number of strictly positive components of (λ, ξ) then $K^* + K \leq r + s + 1$.

Proof. The existence of (λ, ξ, ν) follows from Lemma 4, and Lemma 6 guarantees the existence of $\sum_{k=1}^{K} t_k \delta_{z^k}$ with

$$\sum_{i=1}^{\gamma} \alpha_i \left(\int_{\Delta_i} E\left(y_h z_i\right) d\nu_i\right) = \sum_{i=1}^{\gamma} \alpha_i \left(\int_{\Delta_i} E\left(y_h z_i\right) d\left(\sum_{k=1}^{K} t_k \delta_{z^k}\right)\right),$$

$h = 1, 2, ..., n$. Thus, $\left(\lambda, \xi, \sum_{k=1}^{K} t_k \delta_{z^k}\right) \in \mathcal{A}$ and we can assume that $t_k \neq 0$, $k = 1, 2, ..., K$. If $\left(\lambda, \xi, \sum_{k=1}^{K} t_k \delta_{z^k}\right)$ is not a extreme point of $\mathcal{A}$ then $\sum_{k=1}^{K} t_k \delta_{z^k}$ can be replaced by a new linear convex combination $\sum_{k=1}^{K} \tau_k \delta_{z^k}$ such that $(\tau_k)_{k=1}^{K}$ is a extreme point of the set composed of those $(\tilde{\tau}_k)_{k=1}^{K}$ with non-negative components, $\sum_{k=1}^{K} \tau_k = 1$ and

$$\sum_{i=1}^{\gamma} \alpha_i \left(\int_{\Delta_i} E\left(y_h z_i\right) d\nu_i\right) = \sum_{i=1}^{\gamma} \alpha_i \left(\int_{\Delta_i} E\left(y_h z_i\right) d\left(\sum_{k=1}^{K} \tilde{\tau}_k \delta_{z^k}\right)\right).$$

Now it is easy to see that $\left(\lambda, \xi, \sum_{k=1}^{K} \tau_k \delta_{z^k}\right)$ is a extreme point of $\mathcal{A}$.

Finally, it only remains to see that $K^* + K \leq n + 1$ (we still denote by K the number of nun-null components of $(\tau_k)_{k=1}^{K}$). Since $\left(\lambda, \xi, \sum_{k=1}^{K} \tau_k \delta_{z^k}\right) \in \mathcal{A}$ one has

$$\left. \begin{array}{ll} \sum_{j=1}^{m} E\left(q_j y_h\right) \lambda_j - \sum_{k=1}^{K} \tau_k \left[\int_{\Delta} \left(\sum_{i=1}^{\gamma} \alpha_i E\left(y_h z\right)\right)\right] d\delta_{z^k} = 0, & h = 1, 2, ..., r \\ \sum_{j=1}^{m} E\left(q_j y_h\right) \lambda_j - \xi_h - \sum_{k=1}^{K} \tau_k \left[\int_{\Delta} \left(\sum_{i=1}^{\gamma} \alpha_i E\left(y_h z\right)\right)\right] d\delta_{z^k} = 0, & h = r + 1, ..., r + s \\ \sum_{k=1}^{K} \tau_k = 1 & \\ \lambda \geq 0, \xi \geq 0, \tau \geq 0 & \end{array} \right\}$$

$$(22)$$

If one fixes the whole set of parameters in the system above except (λ, ξ) and $\tau = (\tau_k)_{k=1}^{K}$ we already know that there is a solution in the unknown (λ, ξ, τ). Moreover,

since our concrete (λ, ξ, τ) is a extreme point of $\mathcal{A}$ it is also a extreme point of the set of solutions of the system above. Hence, (λ, ξ, τ) is a Basic Feasible Solution (Anderson and Nash, 1987) and whence it cannot have more than $n + 1$ (number of equations) strictly positive components.   ¤

Remark 3. *Notice that the necessary and sufficient optimality conditions (15) may be simplified by using the latter theorem.*   ¤

Remark 4. (Simplex-like Algorithm). *Anderson and Nash (1987) presented a simplex-like algorithm that solves some semi-infinite linear optimization problems. Since their approach does not exactly fit our Problem (17), we will adapt the method. However, we will just present the algorithm without proofs, because they are quite similar to those provided by the authors above.*

Step1. *Fix a initial Basic Feasible Solution (extreme point)* $\left(\lambda, \xi, \sum_{k=1}^{K} t_k \delta_{z^k}\right)$ *satisfying the conditions stated in Theorem 7. In order to find this first Basic Feasible Solution (BFS) one can follow those procedures indicated in Anderson and Nash (1987).*
Consider the matrix of System (22) above

$$A = \begin{pmatrix} \left(E\left(y_h q_j\right)\right)_{h=1,j=1}^{h=r,j=m} & (0)_{r \times s} & \left(\left[\int_\Delta \left(\sum_{i=1}^{\gamma} \alpha_i E\left(y_h z\right)\right)\right] d\delta_{z^k}\right)_{h=1,k=1}^{h=r,k=K} \\ \left(E\left(y_h q_j\right)\right)_{h=r+1,j=1}^{h=r+s,j=m} & -I_{s \times s} & \left(\left[\int_\Delta \left(\sum_{i=1}^{\gamma} \alpha_i E\left(y_h z\right)\right)\right] d\delta_{z^k}\right)_{h=r+1,k=1}^{h=r+s,k=K} \\ (0)_{1 \times m} & (0)_{1 \times s} & \left((1,1,...,1)_{1 \times K}\right. \end{pmatrix},$$

whose dimensions equal $(n + 1) \times (m + s + K)$. Consider a sub-matrix B taking the columns of A associated with non-null elements of $\left(\lambda, \xi, \sum_{k=1}^{K} t_k \delta_{z^k}\right)$. Theorem 7 guarantees that B has less than $(n + 1)$ columns or exactly $(n + 1)$ columns. In the first case this BFS is said to be degenerated and non-degenerated in the second one. If we were facing degeneration then we would add some columns of A so as to reach a square and regular B. Consider finally the row matrix

$$c = (-b_1, -b_2, ..., -b_m, 0, .0, ..., 0)$$

with $(m + s + K)$ columns and the row matrix c_B with $(n + 1)$ columns that is obtained taking from c those elements associated with the columns of A composing B (henceforth, basic columns).

Step2. *Compute the matrices*

$$B^{-1}A \text{ and } c_B B^{-1}A - c.$$

It is trivial to show that those columns of $B^{-1}A$ associated with the basic columns generate the identity matrix and, consequently, those elements of $c_B B^{-1}A - c$ associated with the basic columns must vanish. In particular, all of the elements associated with $\sum_{k=1}^{K} t_k \delta_{z^k}$ must vanish.

Consider the function

$$\Delta \ni z \longmapsto \Phi(z) = c_B B^{-1} \begin{pmatrix} \left(\left[\int_\Delta \left(\sum_{i=1}^\gamma \alpha_i E(y_h z) \right) \right] d\delta_z \right)_{h=1}^{h=r} \\ \left(\left[\int_\Delta \left(\sum_{i=1}^\gamma \alpha_i E(y_h z) \right) \right] d\delta_z \right)_{h=r+1}^{h=n} \\ 1 \end{pmatrix} \in \mathbb{R} \qquad (23)$$

We can extend the matrix $B^{-1}A$ by adding a new row and a two new columns. We obtain the simplex tableau

$$\begin{pmatrix} & c_B B^{-1}A - c. & -\sum_{j=1}^m b_j \lambda_j \\ \begin{pmatrix} (j_l)_{j_l} \\ (h_l)_{h_l} \\ (z_{k_l})_{k_l} \end{pmatrix} & B^{-1}A, & B^{-1} \begin{pmatrix} 0 \\ . \\ . \\ . \\ 1 \end{pmatrix} \end{pmatrix} \qquad (24)$$

indicating those variables that are basic (first column), their values (last column) and the objective level.

Optimality Criterion. *If there are no negative elements in $c_B B^{-1}A - c$ and*

$$\Phi(z) \geq 0 \qquad (25)$$

for every $z \in \Delta$ then $\left(\lambda, \xi, \sum_{k=1}^K t_k \delta_{z^k} \right)$ solves (17). The algorithm ends here. Otherwise we must go to Step 3.

Step3. Scenario 1. Assume that $c_B B^{-1}A - c$ contains a negative element. It is associated with a non-basic column of A,

$$\begin{pmatrix} (E(y_h q_j))_{h=1,j=1}^{h=r,j=m} \\ (E(y_h q_j))_{h=r+1,j=1}^{h=r+s,j=m} \\ . \\ . \\ . \\ 0 \end{pmatrix} \quad \text{or} \quad (0,...,0,0,...,-1,...,0,0)^t$$

related to a non-basic variable λ_{j_0} or ξ_{h_0} that will become basic in a new iteration of the algorithm. Denote by A_0 the column above and compute the column matrix (which is a column of $B^{-1}A$)

$$\left(\eta_1, \eta_2, ..., \eta_{n+1} \right)^t = B^{-1}A_0. \qquad (26)$$

It may be easily proved that there exists a positive element in the column matrix above since, otherwise, Problem (17) would be unbounded, contradicting Assumption I. Take

$$\eta = Min \left\{ \left\{ \frac{\lambda_{j_l}}{\eta_l}; \ \eta_l > 0 \right\} \cup \left\{ \frac{\xi_{h_l}}{\eta_l}; \ \eta_l > 0 \right\} \cup \left\{ \frac{t_{k_l}}{\eta_l}; \ \eta_l > 0 \right\} \right\}, \qquad (27)$$

with the obvious notations $\left(\left(\lambda_{j_l} \right), \left(\xi_{h_l} \right), (t_{k_l}) \right) = B^{-1} \left(0, 0, ..., 1 \right)^t$ is in the last column of (24) and represents here the "old" or non-optimal basic feasible solution). The minimum value is attained at the element representing the variable that will give up being basic in the next iteration. Therefore we have modified the set a basic variables, and we can solve System (22) so as to obtain a new BFS. The objective value has been improved in the new BFS and we have to go back to Step 1.

However, some computations can be accelerated if one proceeds as follows: Call "pivot" P to that element of $B^{-1}A$ in the position where the minimum (27) above is reached. Update the row of $B^{-1}A$ and Tableau (24) containing the pivot by dividing the whole row by P (one will obtain the new value 1 instead of P). Update $c_B B^{-1}A - c$ and the remainder rows of $B^{-1}A$ and (24) by subtracting from the old one the new row containing the pivot multiplied by P_l, element in the corresponding row and in the same column as the pivot. Modify the subscripts included in the first column of (24) so as to reflect the new basic variables. We can now go back to Step 2 *and we have already updated $\left(\lambda, \xi, \sum_{k=1}^{K} t_k \delta_{z^k} \right)$, $B^{-1}A$, $c_B B^{-1}A - c$ and (24).*

Step3. Scenario 2. Assume that $c_B B^{-1}A - c$ does not contain any negative element but there exists $\tilde{z} \in \Delta$ with $\Phi \left(\tilde{z} \right) < 0$. Obviously, $\tilde{z}$ does not belong to the set $\left\{ z^1, z^2, ...z^K \right\}$. If possible, choose $\tilde{z}$ so as to solve

$$Min \ \left\{ \Phi \left(z \right); \ z \in \Delta \right\}. \qquad (28)$$

Otherwise choose an arbitrary $\tilde{z}$ making negative the value of Φ. $\tilde{z}$ will belong to the next BFS and we have to determine to variable to leave the basis. In order to do that we follow the same way as in the previous scenario. Thus, let

$$A_0 = \begin{pmatrix} \left(\left[\int_{\Delta} \left(\sum_{i=1}^{\gamma} \alpha_i E \left(y_h \tilde{z} \right) \right) \right] d\delta_{\tilde{z}} \right)_{h=1}^{h=r} \\ \left(\left[\int_{\Delta} \left(\sum_{i=1}^{\gamma} \alpha_i E \left(y_h \tilde{z} \right) \right) \right] d\delta_{\tilde{z}} \right)_{h=r+1}^{h=n} \\ 1 \end{pmatrix}$$

compute (26), and the value η of (27) will indicate the variable to stop being basic. Now we can solve System (22) so as to obtain a new BFS and we have to go back to Step 1. *We can simplify some computations if we update $\left(\lambda, \xi, \sum_{k=1}^{K} t_k \delta_{z^k} \right)$, $B^{-1}A$,*

$c_B B^{-1} A - c$ and (24) by proceeding as in the previous case. If so, it is worth to extend "the old $B^{-1}A$, $c_B B^{-1}A - c$ and (24)" before the computations by adding the new column

$$\begin{pmatrix} \Phi\left(\tilde{z}\right) \\ B^{-1} \begin{pmatrix} \left(\left[\int_{\Delta}\left(\sum_{i=1}^{\gamma}\alpha_i E\left(y_h \tilde{z}\right)\right)\right] d\delta_{\tilde{z}}\right)_{h=1}^{h=r} \\ \left(\left[\int_{\Delta}\left(\sum_{i=1}^{\gamma}\alpha_i E\left(y_h \tilde{z}\right)\right)\right] d\delta_{\tilde{z}}\right)_{h=r+1}^{h=n} \\ 1 \end{pmatrix} \end{pmatrix}.$$

Final Note. If the optimality criterion of Step 2 does not hold and we go to the second scenario of Step 3 then we should choose $\tilde{z} \in \Delta$ such that $\Phi\left(\tilde{z}\right)$ were as small (negative) as possible (we should even solve Problem (28)), in order to achieve a fast convergence of the algorithm to the solution. Thus, this pivotal step might be important, although it is worth to recall that the objective function does not directly depend on the ν variable, whose unique effect is on the constraints of the problem. Anyway, since Φ is linear in the z variable (see (23)), one can check the value of Φ in the extreme points of Δ in order to get an element $\tilde{z}$ as close as possible to the solution of (28)). ¤

6. Some significant risk functions

This section will be devoted to present several examples of risk functions satisfying (1), *i.e.*, such that the developed methodology applies. We will not be exhaustive and will only summarize the most important properties of some significant particular cases.

All the examples are positively homogeneous and convex functions. Thus, if we maximize a generalized Sharpe ratio (GSR), quotient between the expected risk premium and the risk excess with respect to the risk-free security, then we will obtain a special strategy that composes the efficient portfolios in a "new risk/return framework". Indeed, once we fix the risk function, we can outperform every feasible portfolio by adequately combining the riskless asset and the strategy maximizing the GSR. This property is similar to that well known in the $CAPM$ model, where the risk function is the standard deviation and efficiency is achieved if and only if one diversifies between the riskless security and the Market Portfolio. Obviously, the maximization of the GSR is equivalent to the minimization of the risk level under an appropriate linear constraint on the expected risk premium, *i.e.*, the here proposed methodology applies. In some sense one can interpret that we are constructing a "new index" that solves those problems generates by asymmetries and fat tails.

Example 1. Standard Deviation. The standard deviation σ_2 is the dispersion measure used in classical portfolio selection and the equilibrium models $CAPM$ and

APT. It is defined over the Banach space L^2. According to the results of Rockafellar et al. (2006a), Expression (1) holds if (see (2) and (3))

$$\tilde{\Delta}^{\sigma_2} = \Delta^{\sigma_2} - \{1\} = \{z \in L^2;\ E(z) = 1,\ \|z - 1\|_2 \le 1\} - \{1\}$$
$$= \{z \in L^2;\ E(z) = 0,\ \|z\|_2 \le 1\}$$
,

which is a $\sigma(L^2, L^2)-$compact set because it is bounded (recall the Alaoglu's Theorem). The related risk measure also satisfies (1) if we take the compact space Δ^σ. Since, for a general probability space $(\Omega, \mathcal{F}, \mu)$, Δ^σ contains non-positive random variables, the risk measure is not decreasing, and therefore it is not coherent in the sense of Artzner et al. (1999) (see Rockafellar et al., 2006a). This is a drawback because risk measures try to represent capital requirements, and higher pay-offs might lead to higher initial capital reserves. As already said, the deviation measure σ_2 is not compatible with the second order stochastic dominance ($SOSD$) if asymmetric returns are involved (Ogryczak and Ruszczynski, 2002).

Example 2. $p-$**Deviation.** The $p-$deviation is given by $L^p \ni y \longmapsto \sigma_p(y) = \|(y - E(y))\|_p = [E\,|y - E(y)|^p]^{1/p} \in \mathbb{R}$, where $p \in [1, \infty)$. Since $\|.\|_q$ is the dual norm of $\|.\|_p$ we have that $\|y\|_p = Sup\left\{E(yz);\ z \in L^q,\ \|z\|_q \le 1\right\}$ holds for every $y \in L^p$. Hence,

$$\sigma_p(y) = \|(y - E(y))\|_p = Sup\left\{E\left[(y - E(y))\,z\right];\ z \in L^q,\ \|z\|_q \le 1\right\}$$
$$= Sup\left\{E(yz) - E(y)E(z);\ z \in L^q,\ \|z\|_q \le 1\right\}$$
$$= Sup\left\{E\left[y\,(z - E(z))\right];\ z \in L^q,\ \|z\|_q \le 1\right\}$$
$$= Sup\left\{E(y(-z));\ z = E(z') - z',\ z' \in L^q,\ \|z'\|_q \le 1\right\} = Sup\left\{E(y(-z));\ z \in \tilde{\Delta}^{\sigma_p}\right\}.$$

Since $L^q \ni z \longmapsto E(z) - z \in L^q$ is $\sigma(L^q, L^p)-$continuous and the unit ball of L^q is $\sigma(L^q, L^p)-$compact (Alaoglu's Theorem), $\tilde{\Delta}^{\sigma_p}$ is $\sigma(L^q, L^p)-$compact, *i.e.*, Expression (1) and the theory developed in this paper apply.

As in the previous example, $\Delta^{\sigma_p} = \tilde{\Delta}^{\sigma_p} + \{1\}$ contains non-positive random variables and, therefore, the associated risk measure is non-decreasing and non-coherent in general.

The absolute deviation $\sigma_1(y) = E\,|y - E(y)|$, $y \in L^1$, presents a significant property. Indeed, it is always compatible with the $SOSD$ (Ogryczak and Ruszczynski, 2002).

Example 3. **Worst Conditional Expectation.** The worst conditional expectation (WCE) is a coherent risk measure introduced in Artzner et al. (1999) and

defined on the space L^1. It is expectation bounded and can be represented if we take

$$\Delta^{WCE} = \bar{C}o \left\{ \frac{\chi_{\Omega_0}}{\mu(\Omega_0)}; \ \Omega_0 \in \mathcal{F}, \ \mu(\Omega_0) > \mu_0 \right\} \subset L^\infty$$

$\mu_0 \in (0,1)$ being the level of confidence (see Rockafellar et al., 2006a). Δ^{WCE} is $\sigma(L^\infty, L^1)$−compact with the same arguments as in the previous case.

Example 4. Conditional Value at Risk. The conditional value at risk ($CVaR$), defined on L^1, is becoming a very important risk measure, quite analyzed in theoretical studies and quite used by traders in practice. It is coherent and Rockafellar et al. (2006a) stated that it is expectation bounded and can be represented with

$$\Delta^{CVaR} = \left\{ z \in L^\infty; \ 0 \leq z \leq \frac{1}{\mu_0}, \ E(z) = 1 \right\} \subset L^\infty$$

$\mu_0 \in (0,1)$ being the level of confidence. Δ^{CVaR} is obviously convex and it is also $\sigma(L^\infty, L^1)$−compact for the same reasons as Examples 2 and 3. The extreme set of Δ^{CVaR} is important to make it easier to deal with Condition (25) and Problem (28). If μ is atomless then it is composed of those random variables taking the form $(1/\mu_0)\chi_{\Omega_0}$, $\Omega_0 \in \mathcal{F}$, $\mu(\Omega_0) = \mu_0$. Rockafellar et al. (2006a) show that $CVaR$ and WCE are closely related. The deviation associated with $CVaR$ can also my treated with our methodology since it is represented by the same convex compact set with -1, $(1/\mu_0) - 1$ and $E(z) = 0$ rather than 0, $(1/\mu_0)$ and $E(z) = 1$. Despite its growing importance, $CVaR$ is not compatible with the $SOSD$, as we will see in the next example.

Example 5. Distortion Functions. The examples above do not satisfy the whole set of "suitable properties" (coherent and bounded expectation risk measures such that they or their deviations are compatible with the $SOSD$) but we will "solve it" in our last example.

Wang (2000) considers a non-decreasing function $g : [0,1] \longmapsto [0,1]$ with $g(0) = 0$ and $g(1) = 1$, and a general risk measure with the "heuristic form" $R_g(y) = -\int_0^1 F_y^{-1}(t) dg(t)$, where F_y stands for the distribution function of the random variable y, and F_y^{-1} must be adequately defined. If g is continuous in $[0,1]$ and piecewise differentiable in $(0,1)$ then the expression above gives

$$R_g(y) = -\int_0^1 F_y^{-1}(t) g'(t) dt. \tag{29}$$

A special important case is

$$g(t) = \begin{cases} (1/\mu_0)\, t, & t \leq \mu_0 \\ 1, & t \geq \mu_0 \end{cases}$$

that leads to $CVaR$. Wang justifies his proposal by several reasons. Among them, he shows simple numerical examples illustrating that $CVaR$ might present some drawbacks in particular practical problems. The idea of Wang is formalized in Whirch and Hardy (2006), where it is established that R_g is compatible with the $SOSD$ if and only if g is strictly concave, case in which R_g is also coherent.

There are many risk measures given by (29) satisfying a set of desirable conditions for which our methodology applies. For example, suppose that g is two times continuously differentiable in $[0,1]$ and with strictly negative second derivative in $(0,1)$. Then R_g is defined in L^1, compatible with the $SOSD$, and it may be stated that (1) holds if Δ^g is the weakly$^*-$closed convex hull of

$$\{0\} \cup \left\{ z \in L^\infty;\ z(\omega) = g'\left[\mu\left(y \leq y\left(\omega\right)\right)\right],\ y \in L^1 \right\}.$$

Since g' is non-negative and bounded from above, Δ^g is $\sigma\left(L^\infty, L^1\right)-$compact and (according to Rockafellar et al., 2006a) R_g is coherent and expectation bounded, that is, R_g satisfies the whole set of "ideal properties" and the methodology of this paper applies. A important example in Actuarial Sciences is the "dual power transform", generated by $g_\varkappa(t) = 1 - (1-t)^\varkappa$, $\varkappa \geq 2$.

Wang also proposed the distorting function $g(t) = \Psi\left(a + \Psi^{-1}(t)\right)$, with $a > 0$ and Ψ denoting the cumulative function of the standard normal distribution. Then R_g is defined in L^2, compatible with the $SOSD$, and it may be stated that (1) holds if Δ^g is the weakly$-$closed convex hull of

$$\{0\} \cup \left\{ z \in L^2;\ z(\omega) = g'\left[\mu\left(y \leq y\left(\omega\right)\right)\right],\ y \in L^2 \right\}.$$

Since g' is non-negative and belongs to $L^2[0,1]$ (when $[0,1]$ is endowed with the Measure of Lebesgue), Δ^g is $\sigma\left(L^2, L^2\right)-$compact and (according to Rockafellar et al., 2006a) R_g is coherent and expectation bounded.

7. Conclusions

Capital requirements of Financial Institutions, fat-tailed return distributions and asymmetric returns have provoked a growing interest in modern risk analysis. Researchers, regulators and practitioners are sharing in the development of new methods measuring risk levels of investment strategies.

Portfolio optimization is complex in practice when general risk functions (measures) are involved. These functions are usually non-differentiable and it is also complex to apply the properties of Convex (or convex-like) Programming. Thus, the development of appropriate optimality conditions and algorithms is becoming an important topic.

Due to the Representation Theorems of Risk Measures we have transformed a (vector) Risk Minimization Problem of Portfolio Choice Theory into a Minimax Problem,

and consequently, into a dual pair of Linear Programming Problems between infinite-dimensional Banach spaces of continuous functions and inner regular $\sigma-$additive measures. Accordingly, necessary and sufficient optimality conditions have been established. They do not draw on the concept of subgradient, and some probability measures are involved as multipliers.

Under very general conditions the dual problem becomes semi-infinite, allowing for a simplex-like algorithm leading to both primal and dual solutions. Thus, one has a practical method to obtain optimal strategies and their sensitivities. Furthermore, in the semi-infinite case the dual solution is achieved at a convex combination of Dirac deltas, allowing for new versions and interpretations of the optimality conditions.

The theory applies for many coherent, expectation bounded and/or deviation risk measures and, therefore, it can be particularized to address those special cases we may be interested in.

¤

References

[1] Anderson, E.J. and P. Nash, 1987. "*Linear programming in infinite-dimensional spaces*". John Wiley & Sons, New York.

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

[3] Balbás, A., P.J. Guerra and M.J. Muñoz-Bouzo, 2002. "The balance space approach in optimization with Riesz spaces valued objectives: An application to financial markets". *Computers & Mathematics with Applications*, 44, 7, 887-897.

[4] Balbás, A. and R. Romera, 2006. "Hedging bond portfolios by optimization in Banach spaces". *Journal of Optimization Theory and Applications*, 131, 1, forthcoming.

[5] Barber, J.R. and M.L. Copper, 1998. "A minimax risk strategy for portfolio immunization." *Insurance: Mathematics and Economics*, 23, 173-177.

[6] Benati, S., 2003. "The optimal portfolio problem with coherent risk measure constraints". *European Journal of Operational Research*, 150, 572-584.

[7] Föllmer, H. and A. Schied, 2002. "Convex measures of risk and trading constraints". *Finance & Stochastics*, 6, 429-447.

[8] Galperin E.A., 1997. "Pareto analysis vis-à-vis balance space approach in multiobjective global optimization". *Journal of Optimization Theory and Applications*, 93, 3, 533-545.

[9] Holmes, R.B.,1975. "*Geometrical functional analysis and its applications.*" Graduate Texts in Mathematics, No 24, Springer-Verlag, New York-Heidelberg.

[10] Jouini, E., M. Meddeb and N. Touci, 2004. "Vector-valued coherent risk measures". *Finance & Stochastics*, 8, 531-552

[11] Konno, H., K. Akishino and R. Yamamoto, 2005. "Optimization of a long-short portfolio under non-convex transaction costs". *Computational Optimization and Applications*, 32, 115-132.

[12] Nakano, Y., 2004. "Efficient hedging with coherent risk measure". *Journal of Mathematical Analysis and Applications*, 293, 345-354.

[13] Ogryczak, W. and A. Ruszczynski, 2002. "Dual stochastic dominance and related mean risk models". *SIAM Journal of Optimization*, 13, 60-78.

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

[15] Rockafellar, R.T., S. Uryasev and M. Zabarankin, 2006b. "Optimality conditions in portfolio analysis with general deviations measures". *Mathematical Programming*, Ser. B, 108, 515-540.

[16] Ruszczynski, A. and A. Shapiro, 2006. "Optimization of risk measures". *Mathematics of Operations Research*, forthcoming.

[17] Shimizu, K. and E. Aiyoshy, 1980. "Necessary conditions for Min-Max problems and algorithms by a relaxation procedure". *Transactions on Automatic Control*, 25, 62-67.

[18] Wang, S.S., 2000. "A class of distortion operators for pricing financial and insurance risks". *Journal of Risk and Insurance*, 67, 15-36.

[19] Whirch, J.L., and M.R. Hardy, 2006. "Distortion risk measures. Coherence and stochastic dominance". Unpublished Working Paper (http://www.gloriamundi.org).

[20] Young, M.R., 1998. "A minimax portfolio selection rule with linear programming solution.". *Management Science*, 43, 673-683.