UNIVERSIDAD CARLOS III DE MADRID
working papers

Working Paper 10-19
Statistics and Econometrics Series 08
June 2010

Departamento de Estadística
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34) 91 624-98-49

MULTIVARIATE EXTREMALITY MEASURE

Henry Laniado* **Rosa E. Lillo**** **Juan Romo****

Abstract

We propose a new multivariate order based on a concept that we will call extremality". Given a unit vector, the extremality allows to measure the "farness" of a point with respect to a data cloud or to a distribution in the vector direction. We establish the most relevant properties of this measure and provide the theoretical basis for its nonparametric estimation. We include two applications in Finance: a multivariate Value at Risk (VaR) with level sets constructed through extremality and a portfolio selection strategy based on the order induced by extremality.

Keywords: extremality; oriented cone; value at risk; portfolio selection

* Department of Statistics, Universidad Carlos III de Madrid, 28911, Leganés, Madrid, Spain. e-mail: hlaniado@est-econ.uc3m.es .

** Department of Statistics, Universidad Carlos III de Madrid, 28903, Getafe, Madrid, Spain. e-mail: lillo@est-econ.uc3m.es (Rosa E. Lillo) romo@est-ecom.uc3m.es (Juan Romo).

Multivariate Extremality Measure

Henry Laniado* Rosa E. Lillo Juan Romo†

June 16, 2010

Abstract

We propose a new multivariate order based on a concept that we will call "extremality". Given a unit vector, the extremality allows to measure the "farness" of a point in $\Re^n$ with respect to a data cloud or to a distribution in the vector direction. We establish the most relevant properties of this measure and provide the theoretical basis for its nonparametric estimation. We include two applications in Finance: a multivariate Value at Risk (VaR) with level sets constructed through extremality and a portfolio selection strategy based on the order induced by extremality.

1 Introduction

A multivariate order is a valuable tool to analyze the data properties and to obtain direct analogues for multivariate data of univariate order concepts such as median, range, extremes, quantiles or order statistics. Generalization of these concepts to the multivariate case is not easy due to the difficulty of defining total orders in $\mathbb{R}^n$. Chaudhuri[7] and references therein has studied different ways to generalize quantiles, but the lack of a unique criterion for ordering multivariate observations is the key problem in extending these concepts to several dimensions. Barnett [3] was among the first to give an extension of univariate order concepts such as median, extremes and ranges to the higher dimensional case. A flexible way to summarize properties of multivariate data are processes based on generalized quantile functions which are studied in Einmahl and Mason [11].

Multivariate orders allow comparisons and decision making in multiple output scenarios; for example, in psychology and sociology to compare individuals by their characteristics; in the financial industry is important to compare portfolios and performance of investment funds (Zani et al [29]). The detection of outliers in multivariate data is also a relevant application of the multivariate orders (Cerioli and Riani [6]).

Several extensions of usual orders from $\mathbb{R}$ to $\mathbb{R}^n$, such as the Pareto-dominance types and componentwise order, have the drawback of not being total orders. To

*Department of Statistics, Universidad Carlos III de Madrid, 28911, Leganés, Madrid, Spain (e-mail: hlaniado@est-econ.uc3m.es)

†Department of Statistics, Universidad Carlos III de Madrid, 28903, Getafe, Madrid, Spain (e-mail: lillo@est-econ.uc3m.es; juan.romo@uc3m.es)

facilitate the total comparison in the multivariate case the antisymmetry is waived; as a consequence, preorders are obtained instead of orders. An example is defining some function of interest $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ and ordering the data according to its f-value i.e., $x \leq y \Longleftrightarrow f(x) \leq f(y)$. Orders defined through either norms or projections onto some vector $\vec{u}$ such as order by average or weighted average are of this type (see Barnett [3]). A depth function assigns each point in $\mathbb{R}^n$ a measure of centrality with respect to the data cloud or probability distribution. This measure decreases from the center outward (see, e. g., Zuo and Serfling [30] and Liu et al. [19]) and thus a depth function provides a multivariate order that allows to define multivariate versions of median, order statistics, multivariate spacing and tolerance regions (Li and Liu [20]). Another example is majorization (Marshall and Olkin [22]). The majorization order is based on the idea of homogeneity between the components of a vector in $\mathbb{R}^n$ and is used in economy to compare the distribution of wealth in populations. Other orders can be characterized by a Euclidean convex cone C; for instance, for $x,\ y \in \mathbb{R}^n$ $x \leq y \Longleftrightarrow y - x \in C$. This is the case of the componentwise order, where $C = R^+ \cup \{0\}$ or $C = R^+$. These are two of the most important convex cones: the *non-negative orthant* and the *positive orthant* and are useful in the theory of inequalities. It is customary to write $x \leq y$, if $y - x$ belongs to the non-negative orthant. (see Rockafellar [24]).

Next, we introduce the concept of extremality for multivariate data. The extremality of $x \in \mathbb{R}^n$ in the direction $\vec{u}$ is one minus the probability of a oriented convex cone with vertex in x. This cone will be called oriented sub-orthant. Different $\vec{u}$ unit vectors define different ways to rank a multivariate sample. An important step in multivariate data analysis taking into account directions has been made in Kong and Mizera [18] where is adopted a very simple and quite natural projection- based definition of quantiles. More recently Hallin et al [15] proposed a new multivariate quantile based on a directional version of traditional regression quantiles, which also are associated with a vector $\vec{u}$. In the same paper they showed that the contours generated by the directional quantiles coincide with the classical halfspace depth contours. Our proposal of extremality is also based on directions. To calculate the extremality of a point $x \in \mathbb{R}^n$ we move the non-negative orthant in the direction given by $\vec{u}$ and translating the origin to x. The above determines a isomorph cone to the non-negative orthant and will be denoted by $\mathcal{C}_x^{\vec{u}}$. Thus, the extremality of x in direction $\vec{u}$ will be $1 - \mathbb{P}\left(\mathcal{C}_x^{\vec{u}}\right)$. Unlike Hallin et al [15] where $\vec{u}$ indicates the direction of the "vertical" axis in the regression, in this paper $\vec{u}$ is "bisectrix" of the oriented cone. For example, if $\vec{u} = \frac{1}{\sqrt{n}}1_n$ then, $\mathcal{C}_x^{\vec{u}} = x + \mathbb{R}_+^n$. As a consequence of the extremality concept we propose a order for multivariate data that allows to establish the "farness" of $x \in \mathbb{R}^n$ respect to points cloud or to a distribution function. Thus, this extremality measure provides a statistical methodology for segmenting a multivariate data sample, since the set of $x^* \in \mathbb{R}^n$ such that $\mathbb{P}\left(\mathcal{C}_{x^*}^{\vec{u}}\right) = q$ can be interpreted as a multidimensional quantile in the sense of Tibiletti [27] when $\vec{u} = \frac{1}{\sqrt{n}}[\pm 1, \ldots, \pm 1]'$. In fact, extremality is a starting point to study segmentation by considering other type of directions such as the first principal component. The purpose of this paper is to analyze structural properties of this multivariate order and to initiate a theory of nonparametric statistical estimation of the extremality in the sample case. We introduce an estimator and prove its weak and strong consistency.

From an applied perspective to insurance and finance, first we propose a version of multivariate Value at Risk (VaR) based on extremality. The VaR as risk measure has taken place as benchmark in the risk management techniques. Its approach is based on a general notion of risk as the probability of not exceeding a certain threshold quantity considered as dangerous. It has been strongly criticized from Artzner [2] for not to encourage the diversification and defended by Heyde et al. [16] for the robustness. For univariate risks, the VaR is simply the $\alpha-$ quantile of the loss distribution function so the VaR is a risk measure easily interpretable and still remains the most popular measure used by risk managers. However, for the multivariate case to define VaR is more complicated due to existence of manyfold definitions of multidimensional quantile (see Einmahl and Mason [11], Tibiletti [27], Chaudhuri [7], Serfling [25] and Hallin et al. [15] for definitions of multidimensional quantile). Bivariate versions of VaR have been studied in Arbia [1], Tibiletti [28], Nappo and Spizzichino [23] and in general for multivariate case in Embrechts and Pucceti [12] and more recently in Cascos and Molchanov [5]. We propose a multivariate VaR based on extremality notion as its set levels. It enables to identify those relevant events for risk management in the direction $\vec{u}$. Specifically $\vec{u} = \frac{1}{\sqrt{2}}[\pm 1, \pm 1]'$ our VaR coincides with the VaR in sense Tibiletti [28]. Now, $\vec{u} = \frac{-1}{\sqrt{n}}[1, \ldots, 1]'$ and $\vec{u} = \frac{1}{\sqrt{n}}[1, \ldots, 1]'$ the VaR in this paper coincides with multivariate lower orthant VaR and multivariate upper orthant VaR respectively, discussed in Embrechts and Pucceti [12]. However, taking into account other directions we can obtain conservative types of VaR.

As a second application of the multivariate order based on extremality we propose a portfolio selection strategy. Portfolio selection problem was considered in Markowitz [21] whose philosophy is that a investor should hold a portfolio on the set of couples risk-return which one cannot improve both at the same time. This set was denoted as the efficient frontier. From Markowitz [21], several criteria have been studied (see for instance, DeMiguel et al. [10] and references there in) for portfolio selection. We propose to sort feasible portfolios according to the order induced by the direction $\vec{u} = \frac{1}{2}[1, -1]$ that favors the risk and does not favors the returns so we must select the smallest portfolio. We also show that the portfolio selected under this strategy belongs to the efficient frontier.

The paper is organized as follows. Section 2 introduces the definition and properties of the oriented sub-orthant and how it is constructed. In Section 3, we present the extremality measure and the induced multivariate order. The main properties and consistency results are discussed in Section 4. A multivariate VaR is proposed in Section 5 and a portfolio selection strategy is constructed in Section 6, where we compare it with strategies previously used in the literature. Finally, in Section 7 we summarize the main conclusions.

2 Preliminaries

We introduce in this section definitions and preliminary results needed throughout the paper. Recall that a binary relation $\preceq$ on an arbitrary set C is called a *partial order* if it satisfies: *reflexivity* ($x \preceq x$, for all $x \in C$), *transitivity* ($x \preceq y$ and $y \preceq z \Longrightarrow x \preceq z$) and *antisymmetry* ($x \preceq y$ and $y \preceq x \Longrightarrow x = y$). Orderings that satisfy *reflexivity* and *transitivity* are called *preorders.* A subset C of $\mathbb{R}^n$ is said to be *convex* if $(1-\lambda)x + \lambda y \in C$ whenever $x \in C, y \in C$ and $0 \leq \lambda \leq 1$.

Definition 1 (Extreme Point) *Let $C \subseteq \mathbb{R}^n$ be a convex set. Then $x \in C$ is an extreme point of C if there not exist $x_1, x_2 \in C$, $x_1 \neq x_2$, such that*

$$x = \lambda x_1 + (1-\lambda)x_2, \quad \textit{for some} \quad 0 < \lambda < 1.$$

An *extreme point* of a convex set C does not belong to the segment between any two points in C. Points 1 and 3 in the left panel of Figure 1 are extreme, while the right panel has a unique extreme point.

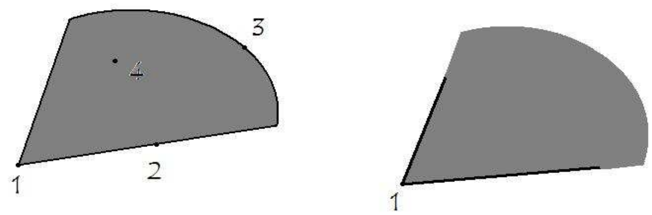


Figure 1: Points 1 and 3 are extreme, and 2 and 4 are not extreme points

Definition 2 *A subset C of $\mathbb{R}^n$ is a cone with vertex in v if $v + \lambda(x - v) \in C$, for all $x \in C$ and $\lambda > 0$.*

C is a *convex cone* if it is a convex set and satisfies Definition 2. Clearly the nonnegative upper orthant is a convex cone with vertex in 0. In this paper, we are interested in rotations of this cone. To formalize the idea, we will use the QR factorization.

Definition 3 *Let A be a $m \times n$ matrix with $m \geq n$. Then A can be factorized as*

$$A = QR,$$

where Q is an orthogonal matrix and

$$R = \begin{pmatrix} R_1 \\ 0 \end{pmatrix},$$

with R_1 an upper triangular matrix.

Matrix Q can be obtained by using, for instance, Householder Reflections, Givens Rotations or Gram-Schmidt Transformations (see Gentle [14], pages 95-103). Since Q is an orthogonal matrix, $Q' = Q^{-1}$. If the diagonal entries of R are required to be nonnegative, Q and R are unique (we will assume nonnegative elements in R along the paper). The next result establishes that R is the first element of the canonical basis in $\mathbb{R}^n$ in the QR factorization of any unit vector.

Proposition 1 *Let $\vec{u} = [u_1, \ldots, u_n]'$ be a vector with Euclidean norm $\|\vec{u}\|_2 = 1$. If $\vec{u} = QR$, then $R = [1, 0, \ldots, 0]'$.*

Proof. We have that

$$1 = \vec{u}'\vec{u} = R'Q'QR = R'IR = R'R.$$

Therefore, R has to be $[1, 0, \ldots, 0]'$ according to Definition 3. ■

Consider the unit vectors $\vec{e} = \frac{1}{\sqrt{n}}[1, \ldots, 1]'$ and $\vec{u} \in \mathbb{R}^n$. Writing

$$\vec{e} = Q_1 R_1 \text{ and } \vec{u} = Q_2 R_2,$$

$R_2 = R_1 = [1, 0, \ldots, 0]'$ from Proposition 1
. Hence, $Q_2'\vec{u} = Q_1'\vec{e}$ and $Q_1 Q_2' \vec{u} = \vec{e}$. Thus,

$$\mathcal{R}_{\vec{u}} = Q_1 Q_2' \tag{1}$$

is an orthogonal matrix transforming $\vec{u}$ into a unit vector with identical components. This transformation will send each vector x to a new orthogonal coordinates system, where $\vec{u}$ has all its angles equal with respect to the new nonnegative axis coordinates, that is, $\mathcal{R}_{\vec{u}}\vec{u} = \vec{e}$. This transformation (1) allows to define the following special oriented cones.

Definition 4 (Oriented sub-orthant $\mathcal{C}_x^{\vec{u}}$) *Given a unit director vector $\vec{u} \in \mathbb{R}^n$ and a vertex $x \in \mathbb{R}^n$, an oriented sub-orthant $\mathcal{C}_x^{\vec{u}}$ is the convex cone given by*

$$\mathcal{C}_x^{\vec{u}} = \{z \in \mathbb{R}^n \mid \mathcal{R}_{\vec{u}}(z - x) \geq 0\}, \tag{2}$$

where the inequality is componentwise.

$\mathcal{C}_x^{\vec{u}}$ is a convex cone with vertex in x obtained moving the nonnegative orthant and translating the origin to x. Besides, according to Definition 1, $\mathcal{C}_x^{\vec{u}}$ is a convex set with a single *extreme point* in x, the semi-line

$$l = \{z \in \mathbb{R}^n \mid z = x + \lambda\vec{u}, \qquad \lambda \geq 0\} \tag{3}$$

is totally contained in $\mathcal{C}_x^{\vec{u}}$ and its angles with respect to the new nonnegative semi-axis coordinates are equal to $\cos^{-1}\left(\frac{1}{\sqrt{n}}\right)$. Note that when $\vec{u} = \frac{1}{\sqrt{n}}[\pm 1 \cdots \pm 1]'$ and $v = 0$, $\mathcal{C}_v^{\vec{u}}$ coincides with the 2^n orthants in $\mathbb{R}^n$.

Example 1 *Consider $\mathbb{R}^2$. If with $\vec{u} = [u_1, u_2]'$ and $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ then,*

$$\mathcal{C}_x^{\vec{u}} = \left\{ \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \in \mathbb{R}^2 \ : \ \frac{\sqrt{2}}{2} \begin{pmatrix} u_1 + u_2 & u_2 - u_1 \\ u_1 - u_2 & u_1 + u_2 \end{pmatrix} \begin{pmatrix} z_1 - x_1 \\ z_2 - x_2 \end{pmatrix} \geq \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\}. \tag{4}$$

Example 2 *In $\mathbb{R}^2$, the director vector $\vec{u}$ can be determined by an angle $0 \leq \theta \leq 2\pi$ indicating the direction of the cone. Then $\vec{u} = [\cos\theta, \sin\theta]'$ and*

$$\mathcal{C}_x^{\vec{u}} = \left\{ \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \in \mathbb{R}^2 : \begin{pmatrix} \cos(\theta - \frac{\pi}{4}) & \sin(\theta - \frac{\pi}{4}) \\ -\sin(\theta - \frac{\pi}{4}) & \cos(\theta - \frac{\pi}{4}) \end{pmatrix} \begin{pmatrix} z_1 - x_1 \\ z_2 - x_2 \end{pmatrix} \geq \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\}. \tag{5}$$

Thus, $\mathcal{C}_x^{\vec{u}}$ is a convex cone obtained rotating the non-negative quadrant by an angle $(\theta - \frac{\pi}{4})$ and translating the origin to $(x_1, x_2)'$. Besides, the semi-line (3) will be bisectrix of $\mathcal{C}_x^{\vec{u}}$ with angles $\cos^{-1}\left(\frac{1}{\sqrt{2}}\right)$ with respect to the rotated nonnegative semi-axis.

Figure 2 presents the *oriented sub-orthants* $\mathcal{C}_A^{\vec{u_1}}, \mathcal{C}_B^{\vec{u_1}}, \mathcal{C}_C^{\vec{u_2}}$ and $\mathcal{C}_D^{\vec{u_3}}$ with vertices in A, B, C, D and $\vec{u} = [\cos\theta, \sin\theta]$, for $\theta = \frac{\pi}{3}, \frac{\pi}{4}, \frac{5\pi}{4}, \frac{\pi}{2}$, respectively.

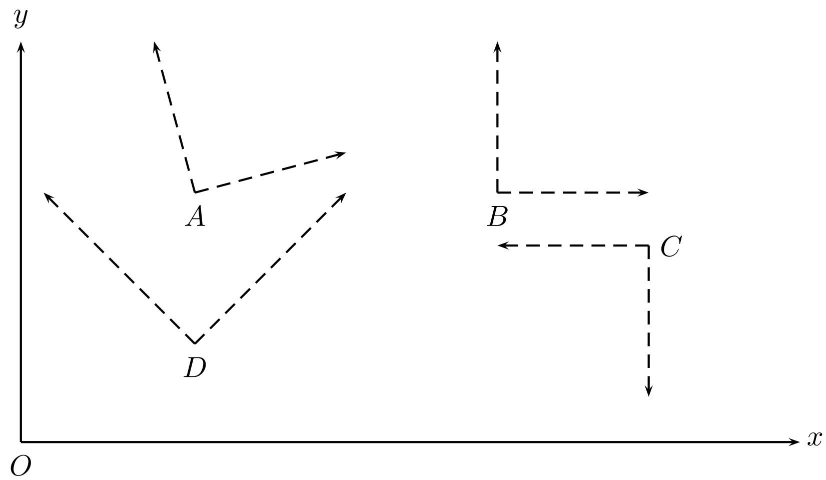


Figure 2: Examples of oriented sub-orthants

Clearly, $\mathcal{C}_{0^+}^{\frac{\pi}{4}}$, $\mathcal{C}_{0^+}^{\frac{3\pi}{4}}$, $\mathcal{C}_{0^+}^{\frac{5\pi}{4}}$ and $\mathcal{C}_{0^+}^{\frac{7\pi}{4}}$ are $(+,+)$; $(-,+)$; $(-,-)$; $(+,-)$ quadrants, respectively. In $\mathbb{R}^n$, if $\vec{u}_1 = \vec{e}$, $\vec{u}_2 = -\vec{e}'$ then $\mathcal{C}_0^{\vec{u}_1}$ and $\mathcal{C}_0^{\vec{u}_2}$ are, respectively, the nonnegative and nonpositive orthants, since $\mathcal{R}_{\vec{u_1}} = I_n$ and $\mathcal{R}_{\vec{u_2}} = -I_n$ (see equation (1)).

Proposition 2 *For any $\vec{u}$, if $x \in \mathcal{C}_y^{\vec{u}}$ then $\mathcal{C}_x^{\vec{u}} \subset \mathcal{C}_y^{\vec{u}}$.*

Proof. Suppose that $z \in \mathcal{C}_x^{\vec{u}}$. From Definition 4, $\mathcal{R}_{\vec{u}}(z-x) \geq 0$, and $\mathcal{R}_{\vec{u}}(x-y) \geq 0$ by hypothesis. Then, $\mathcal{R}_{\vec{u}}(z-y) = \mathcal{R}_{\vec{u}}(z-x) + \mathcal{R}_{\vec{u}}(x-y) \geq 0$ and, therefore, $z \in \mathcal{C}_y^{\vec{u}}$. ■

The following Proposition shows that there exists at least a transformation that allows to compare componentwise two points in $\mathbb{R}^n$.

Proposition 3 *If $x \neq y \in \mathbb{R}^n$ and $\vec{u} = \frac{(x-y)}{\|x-y\|}$, where $\|\cdot\|$ is the Euclidean norm then,*

i) $\mathcal{R}_{\vec{u}} y \leq \mathcal{R}_{\vec{u}} x$

ii) $\mathcal{C}_x^{\vec{u}} \subset \mathcal{C}_y^{\vec{u}}$.

Proof. *i)* According to transformation (1), for any unit vector $\vec{u}$, $\mathcal{R}_{\vec{u}}\vec{u} = \vec{e}$. In particular, for $\vec{u} = \frac{(x-y)}{\|x-y\|}$, then $\mathcal{R}_{\vec{u}}(x-y) \geq 0$ and therefore, $\mathcal{R}_{\vec{u}} y \leq \mathcal{R}_{\vec{u}} x$. *ii)* Since $\mathcal{R}_{\vec{u}}(x-y) \geq 0$, $x \in \mathcal{C}_y^{\vec{u}}$ and from Proposition 2 $\mathcal{C}_x^{\vec{u}} \subset \mathcal{C}_y^{\vec{u}}$. ■

3 Extremality Measure

Let $\mathfrak{F}$ be the class of distribution functions on $\mathbb{R}^n$, let X be a random vector with distribution function $F \in \mathfrak{F}$ and probability distribution P_F. Given a unit vector $\vec{u}$ and $x \in \mathbb{R}^n$, denote by $P_{x,\vec{u}}$ the measure P_F of $\mathcal{C}_x^{\vec{u}}$, that is,

$$P_{x,\vec{u}} = \int_{\mathcal{C}_x^{\vec{u}}} dP_F = P_F\left(\mathcal{C}_x^{\vec{u}}\right). \tag{6}$$

If P_F is absolutely continuous and the multivariate random vector X has joint density function f_X, then

$$P_{x,\vec{u}} = \int_{\mathcal{C}_x^{\vec{u}}} f_X(\mathbf{x})\mathbf{dx}. \tag{7}$$

Proposition 4 shows the general way to calculate $P_{x,\vec{u}}$ when $x \in \mathbb{R}^2$ for any orientation.

Proposition 4 *If (X,Y) is a random vector with joint density function f_{XY} and $\vec{u} = [\cos(\theta), \sin(\theta)]'$ as in the Example 2. Then,*

$$\begin{aligned}
P_F\left(\mathcal{C}_{(x,y)}^{\vec{u}}\right) = & \left(\int_x^\infty \int_{\frac{-x\sin(\theta-\frac{\pi}{4})+y\cos(\theta-\frac{\pi}{4})+t\sin(\theta-\frac{\pi}{4})}{\cos(\theta-\frac{\pi}{4})}}^{\frac{x\cos(\theta-\frac{\pi}{4})+y\sin(\theta-\frac{\pi}{4})-t\cos(\theta-\frac{\pi}{4})}{\sin(\theta-\frac{\pi}{4})}} f_{XY}(t,s)dsdt\right) 1_{\{\theta\in[0,\frac{\pi}{4})\cup(\frac{7\pi}{4},2\pi)\}} \\
& + \left(\int_y^\infty \int_{\frac{x\cos(\theta-\frac{\pi}{4})+y\sin(\theta-\frac{\pi}{4})-s\sin(\theta-\frac{\pi}{4})}{\cos(\theta-\frac{\pi}{4})}}^{\frac{x\sin(\theta-\frac{\pi}{4})-y\cos(\theta-\frac{\pi}{4})+s\cos(\theta-\frac{\pi}{4})}{\sin(\theta-\frac{\pi}{4})}} f_{XY}(t,s)dtds\right) 1_{\{\theta\in(\frac{\pi}{4},\frac{3\pi}{4})\}} \\
& + \left(\int_x^\infty \int_{\frac{x\cos(\theta-\frac{\pi}{4})+y\sin(\theta-\frac{\pi}{4})-t\cos(\theta-\frac{\pi}{4})}{\sin(\theta-\frac{\pi}{4})}}^{\frac{-x\sin(\theta-\frac{\pi}{4})+y\cos(\theta-\frac{\pi}{4})+t\sin(\theta-\frac{\pi}{4})}{\cos(\theta-\frac{\pi}{4})}} f_{XY}(t,s)dsdt\right) 1_{\{\theta\in(\frac{3\pi}{4},\frac{5\pi}{4})\}} \\
& + \left(\int_{-\infty}^y \int_{\frac{x\sin(\theta-\frac{\pi}{4})-y\cos(\theta-\frac{\pi}{4})+s\cos(\theta-\frac{\pi}{4})}{\sin(\theta-\frac{\pi}{4})}}^{\frac{x\cos(\theta-\frac{\pi}{4})+y\sin(\theta-\frac{\pi}{4})-s\sin(\theta-\frac{\pi}{4})}{\cos(\theta-\frac{\pi}{4})}} f_{XY}(t,s)dtds\right) 1_{\{\theta\in(\frac{5\pi}{4},\frac{7\pi}{4})\}} \\
& + \left(\int_x^\infty \int_y^\infty f_{XY}(t,s)dsdt\right) 1_{\{\theta=\frac{\pi}{4}\}} + \left(\int_{-\infty}^x \int_y^\infty f_{XY}(t,s)dsdt\right) 1_{\{\theta=\frac{3\pi}{4}\}} \\
& + \left(\int_{-\infty}^x \int_{-\infty}^y f_{XY}(t,s)dsdt\right) 1_{\{\theta=\frac{5\pi}{4}\}} + \left(\int_x^\infty \int_{-\infty}^y f_{XY}(t,s)dsdt\right) 1_{\{\theta=\frac{7\pi}{4}\}}.
\end{aligned}$$

In higher dimensions is more difficult to give a general expression for $P_F\left(\mathcal{C}_{\mathbf{x}}^{\vec{u}}\right)$ unless the unit vector $\vec{u}$ is given numerically.

Let $\mathbf{t} = \mathcal{R}_{\vec{u}}\mathbf{x}$ be the image of $\mathbf{x}$ under the transformation (1). Clearly, $\mathbf{x} = \mathcal{R}'_{\vec{u}}\mathbf{t}$ and the absolute value of the Jacobian is 1. If we write $D_{\mathbf{x}} = \{\mathbf{t} \in \mathbb{R}^n \mid \mathbf{t} \geq \mathbf{x}\}$, then (7) is equivalent to

$$P_{x,\vec{u}} = \int_{D_{\mathbf{x}}} f_X(\mathcal{R}_{\vec{u}}^{-1}\mathbf{t})\mathbf{dt}. \tag{8}$$

If $x_1, \ldots, x_m$ is a sample of the random vector X, the empirical version of $P_{x,\vec{u}}$ is given by

$$\hat{P}_{x,\vec{u}} = \frac{1}{m}\sum_{j=1}^m 1_{\left\{x_j\,\in\,\mathcal{C}_x^{\vec{u}}\right\}} = \frac{1}{m}\sum_{j=1}^m 1_{\{\mathcal{R}_{\vec{u}}(x_j-x)\geq 0\}}, \tag{9}$$

which is the proportion of the data cloud belonging to $\mathcal{C}_x^{\vec{u}}$.

As we have shown, $P_{x,\vec{u}}$ is the probability of an oriented sub-orthant. We can now formulate the extremality notion. This concept is the starting point for defining a new multivariate data order.

Definition 5 (Extremality Measure) *The extremality of $x \in \mathbb{R}^n$ with respect to a distribution function F in direction $\vec{u}$ is a mapping $\mathcal{E}_{\vec{u}}(x, F) : \mathbb{R}^n \times \mathfrak{F} \longrightarrow R^+ \cup \{0\}$, defined by*

$$\mathcal{E}_{\vec{u}}(x, F) = P_F\left(\overline{\mathcal{C}_x^{\vec{u}}}\right) = 1 - P_{x,\vec{u}}, \tag{10}$$

where $P_{x,\vec{u}}$ is given by (6).

The extremality of $x \in \mathbb{R}^n$ respect to a data cloud $\mathbb{X} = \{x_1, \ldots, x_n\}$ in direction $\vec{u}$, denoted by $\mathcal{E}_{\vec{u}}(x, \mathbb{X})$, is defined replacing $P_{x,\vec{u}}$ by $\hat{P}_{x,\vec{u}}$.

High extremality of a point x means that the convex cone $\mathcal{C}_x^{\vec{u}}$ contains a small part of the total mass and possibly x belongs to some tail of the distribution. Hence, high extremality can be interpreted as "farness" in the distribution.

Figure 3 presents the extremality curves of level 0.99; 0.95; 0.90; 0.85 when F is a bivariate distribution with independent marginal distributions $U(0\ 1)$. Left side in direction $\vec{u} = \frac{1}{\sqrt{2}}[1,\ 1]$ and right side in direction $[0,\ 1]$. Figure 4 shows the extremality surfaces of level 0.99; 0.95; 0.90; 0.85 in the direction $\vec{u} = \frac{1}{\sqrt{3}}[1,\ 1,\ 1]$ of a multivariate distribution F with three independent marginal distributions $U(0\ 1)$. All points on a particular curve or surface have the same extremality with respect to the distribution F.

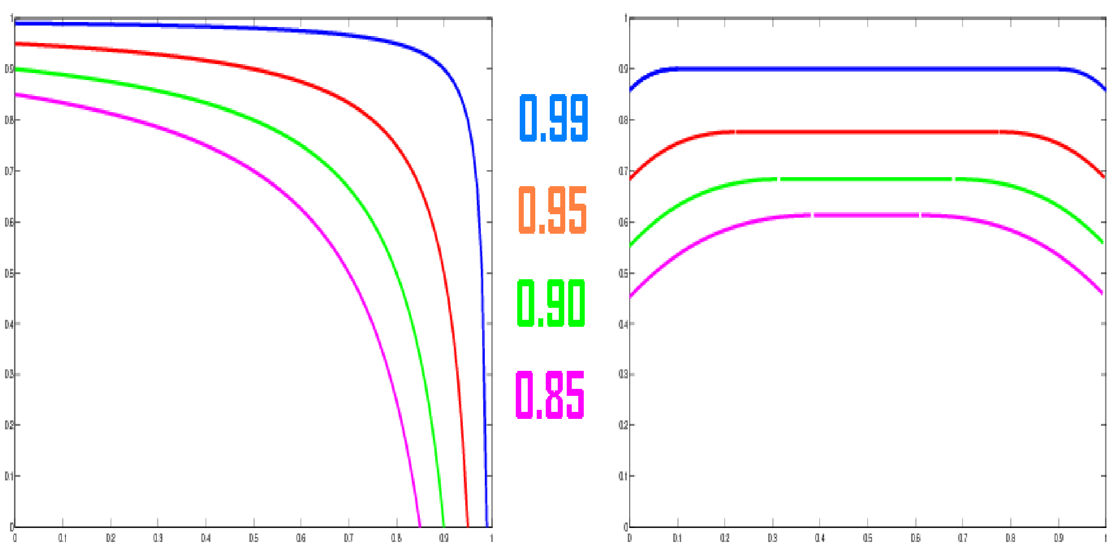


Figure 3: $\mathcal{E}_{\frac{1}{\sqrt{2}}[1,\ 1]}(x, F) = \alpha$ and $\mathcal{E}_{[0,\ 1]}(x, F) = \alpha$

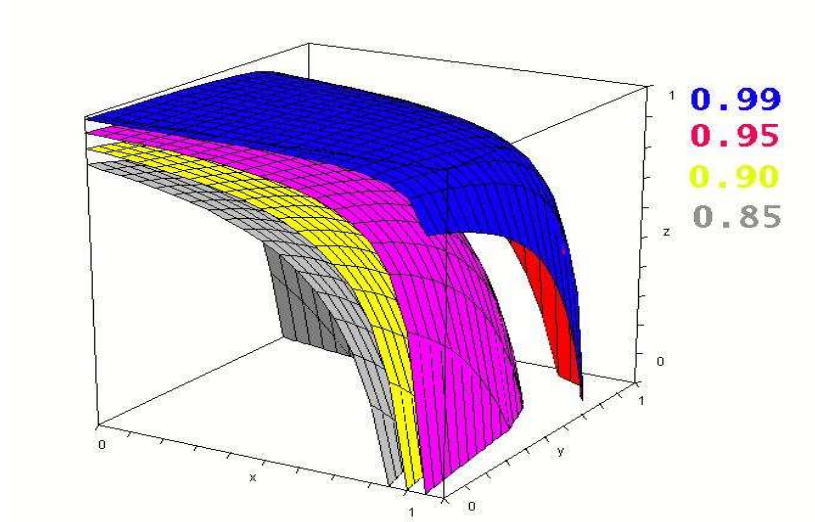


Figure 4: $\mathcal{E}_{\frac{1}{\sqrt{3}}[1,\ 1,\ 1]}(x, F) = \alpha$

Definition 5 induces a multivariate order as follows.

Definition 6 *Given $x, y \in \mathbb{R}^n$, y is said to be more extreme than x respect to F in direction $\vec{u}$, denoted $x \leq_{\mathcal{E}_{\vec{u}}} y$ if, and only if,*

$$\mathcal{E}_{\vec{u}}(x, F) \leq \mathcal{E}_{\vec{u}}(y, F).$$

For any $x, y \in \mathbb{R}^n$, any distribution function $F \in \mathfrak{F}$ and any $\vec{u}$ it holds that either $x \leq_{\mathcal{E}_{\vec{u}}} y$ or $y \leq_{\mathcal{E}_{\vec{u}}} x$. However, $\leq_{\mathcal{E}_{\vec{u}}}$ is not a partial order in $\mathbb{R}^n$, but a preorder. Because, although it satisfies reflexivity and transitivity properties, it does not satisfy antisymmetry. If F is an absolutely continuous distribution in the interval $[a,\ b]$, the extremality order with $\vec{u} = 1$ coincides with the usual order in $\mathbb{R}$.

4 Properties of extremality measure

The extremality measures are nonnegative and bounded. Next we establish its analytic properties that support the ordering proposed in the previous definition.

Property 1 *For any $x_0 \in \mathbb{R}^n$ and any absolutely continuous $F \in \mathfrak{F}$,*

$$\mathcal{E}_{\vec{u}}(x_0, F) \ \textit{is continuous in} \ \vec{u}.$$

Proof. Let f_X be the density function corresponding to F. From (8) and Definition 5, we have that

$$\mathcal{E}_{\vec{u}}(x_0, F) = 1 - \int_{D_{x_0}} f_X(\mathcal{R}_{\vec{u}}^{-1}\mathbf{t})\mathbf{dt},$$

which clearly is continuous in $\vec{u}$ since $\mathcal{R}_{\vec{u}}^{-1}$ is a linear transformation. ∎

The following property indicates that the vertex x has minimal extremality in the set $\mathcal{C}_x^{\vec{u}}$.

Property 2

$$\mathcal{E}_{\vec{u}}(x,F) \leq \mathcal{E}_{\vec{u}}(x^*,F), \quad \textit{for all } x^* \in \mathcal{C}_x^{\vec{u}}$$

Proof. Suppose that $x^* \in \mathcal{C}_x^{\vec{u}}$. From Proposition 2,

$$\mathcal{C}_{x^*}^{\vec{u}} \subset \mathcal{C}_x^{\vec{u}} \Longrightarrow P_F\left(\mathcal{C}_{x^*}^{\vec{u}}\right) \leq P_F\left(\mathcal{C}_x^{\vec{u}}\right) \Longrightarrow \mathcal{E}_{\vec{u}}(x,F) \leq \mathcal{E}_{\vec{u}}(x^*,F).$$

■

Property 3 *Let X be an $n-$ dimensional random variable with distribution function F. Let A be an orthogonal matrix and let $b \in \mathbb{R}^n$. Then*

$$\mathcal{E}_{A\vec{u}}(Ax+b, F_{AX+B}) = \mathcal{E}_{\vec{u}}(x, F_X)$$

Proof. Since A is an orthogonal matrix and $\vec{u}, \vec{e}$ are unit vectors, $A\vec{u}$ is also a unit vector. Using Proposition 1, their QR factorization are given by

$$\vec{e} = Q_1R_1,\ \vec{u} = Q_2R_1,\ A\vec{u} = Q_3R_1, \quad \text{where } R_1 = [1,0,\ldots,0]' \in \mathbb{R}^n. \tag{11}$$

Therefore, applying the transformations (1), we have

$$\mathcal{R}_{\vec{u}} = Q_1Q_2' \quad \text{and} \quad \mathcal{R}_{A\vec{u}} = Q_1Q_3'. \tag{12}$$

As R_1 has diagonal with non-negative entries, the QR factorization of $\vec{u}$ is unique. Therefore, from (11)

$$\vec{u} = Q_2R_1 = A'Q_3R_1, \text{ which implies that } Q_2 = A'Q_3,$$

and, from (12),

$$\mathcal{R}_{\vec{u}} = Q_1Q_2' = Q_1Q_3'A = \mathcal{R}_{A\vec{u}}A. \tag{13}$$

Then, using (13) in the last equality, we obtain

$$\begin{aligned}
\mathcal{E}_{A\vec{u}}(Ax+b, F_{AX+B}) &= 1 - P_{F_{Ax+b}}\left(\mathcal{C}_{Ax+b}^{A\vec{u}}\right) \\
&= 1 - P_F\left(\mathcal{R}_{A\vec{u}}(AX+b-Ax-b) \geq 0\right) \\
&= 1 - P_F\left(\mathcal{R}_{A\vec{u}}A(X-x) \geq 0\right), \quad \text{(from (13))} \\
&= 1 - P_F\left(\mathcal{R}_{\vec{u}}(X-x) \geq 0\right) = 1 - P_F\left(\mathcal{C}_x^{\vec{u}}\right) = \mathcal{E}_{\vec{u}}(x, F_X).
\end{aligned}$$

■

Property 4 *Let $x \in \mathbb{R}^n\{0\}$ and $\vec{u} = \frac{x}{\|x\|}$, where $\|\cdot\|$ is the Euclidean norm. If*

$$\|x\| \longrightarrow \infty \Longrightarrow \mathcal{E}_{\vec{u}}(x,F) \longrightarrow 1.$$

Proof. Let $B = \{b \in \mathbb{R}^n : \|b\| \geq \|x\|\}$. Suppose that $z \in \mathcal{C}_x^{\vec{u}}$. Then $\mathcal{R}_{\vec{u}}(z-x) \geq 0$ (see equation (2)) and $\mathcal{R}_{\vec{u}}z \geq \mathcal{R}_{\vec{u}}x$. Using the transformation defined in (1),

$$\mathcal{R}_{\vec{u}}\vec{u} = \mathcal{R}_{\vec{u}}\frac{x}{\|x\|} = \vec{e},$$

and then,

$$\mathcal{R}_{\vec{u}}z \geq \|x\|\vec{e} > 0.$$

Therefore, $\mathcal{C}_x^{\vec{u}} \subset B$ since

$$\|z\|^2 = z'z = z'\mathcal{R}'_{\vec{u}}\mathcal{R}_{\vec{u}}z = (\mathcal{R}_{\vec{u}}z)'\,\mathcal{R}_{\vec{u}}z = \|\mathcal{R}_{\vec{u}}z\|^2 \geq \|x\|^2\|\vec{e}\|^2 = \|x\|^2.$$

It follows that

$$0 \leq P_F\left(\mathcal{C}_x^{\vec{u}}\right) = P_F\left(\mathcal{R}_{\vec{u}}(X-x) \geq 0\right) \leq P_F\left(\|X\| \geq \|x\|\right).$$

And the proof is complete letting $\|x\| \longrightarrow \infty$ ■

4.1 Convergence Analysis

In this section we propose an estimator of the extremality measure and establish its consistency. Let $X_1, \ldots, X_m$ be independent random variables with a common distribution function F_X. The empirical distribution is

$$P_{F,m}(B) = \frac{\sharp\{k\ :\ X_k \in B,\ 1 \leq k \leq m\}}{m}, \quad B \in \mathcal{B}. \tag{14}$$

Following (10), a natural empirical counterpart of $\mathcal{E}_{\vec{u}}(x,F)$ (that is, the extremality measure) is

$$\mathcal{E}_{\vec{u}}(x,\hat{F}_m) = 1 - P_{F,m}\left(\mathcal{C}_x^{\vec{u}}\right) = 1 - \hat{P}_{x,\vec{u}},$$

where $\hat{F}_m$ denotes the empirical distribution function.

In Theroem 1 below, we show that $\mathcal{E}_{\vec{u}}(x,\hat{F}_m)$ is a strongly consistent estimator of $\mathcal{E}_{\vec{u}}(x,F)$ and obtain its asymptotic distribution.

Theorem 1 *Let X be a random vector with distribution function F. Then,*

i) $\mathcal{E}_{\vec{u}}(x,\hat{F}_m) \longrightarrow \mathcal{E}_{\vec{u}}(x,F)$ *a.s., as* $m \longrightarrow \infty$

ii) $\sup_{x,\vec{u}\in\mathbb{R}^n} \mid \mathcal{E}_{\vec{u}}(x,\hat{F}_m) - \mathcal{E}_{\vec{u}}(x,F) \mid \longrightarrow 0$ *a.s, as* $m \longrightarrow \infty$

iii) $m^{\frac{1}{2}} \frac{\mathcal{E}_{\vec{u}}(x,\hat{F}_m) - \mathcal{E}_{\vec{u}}(x,F)}{\sqrt{\mathcal{E}_{\vec{u}}(x,F)(1-\mathcal{E}_{\vec{u}}(x,F))}} \longrightarrow Z$*, weakly as* $m \longrightarrow \infty$*, where Z is a standard normal random variable.*

.

Proof. Let $I_A(X)$ be the indicator function of set A. Since $X_1, \ldots, X_m$ are *i.i.d* random vectors, $I_{\left(\mathcal{C}_x^{\vec{u}}\right)}(X_i)$, $i = 1, 2, \ldots$ are also *i.i.d* random variables with mean $P_F\left(\mathcal{C}_x^{\vec{u}}\right)$ and variance $P_F(\mathcal{C}_x^{\vec{u}})(1 - P_F(\mathcal{C}_x^{\vec{u}}))$. Let $\mathcal{C}$ be the class of all oriented sub-orthant $\mathcal{C}_x^{\vec{u}}$ with $x, \vec{u} \in \mathbb{R}^n$. From equation (14),

$$P_{F,m}\left(\mathcal{C}_x^{\vec{u}}\right) = \frac{1}{m}\sum_{i=1}^{m} I_{\left(\mathcal{C}_x^{\vec{u}}\right)}(X_i).$$

i) From Strong Law of Large Numbers, (see, e. g., Gaenssler [13], page 2) for each $\mathcal{C}_x^{\vec{u}} \in \mathcal{C}$

$$P_{F,m}\left(\mathcal{C}_x^{\vec{u}}\right) \longrightarrow P_F\left(\mathcal{C}_x^{\vec{u}}\right) \quad a.s., \text{ as } m \longrightarrow \infty.$$

Therefore, $\hat{P}_{x,\vec{u}} \longrightarrow P_{x,\vec{u}}$ *a.s.*, and from Definition 5,

$$\mathcal{E}_{\vec{u}}(x,\hat{F}_m) \longrightarrow \mathcal{E}_{\vec{u}}(x,F) \quad a.s., \text{ as } m \longrightarrow \infty.$$

ii) Using a Glivenko - Cantelli Theorem (see, e. g., Gaenssler[13] page 16), we have that

$$\sup_{\mathcal{C}_x^{\vec{u}} \in \mathcal{C}} \left| P_{F,m}\left(\mathcal{C}_x^{\vec{u}}\right) - P_F\left(\mathcal{C}_x^{\vec{u}}\right) \right| \longrightarrow 0 \quad a.s.$$

and, from Definition 5,

$$\sup_{x,\vec{u} \in \mathbb{R}^2} \left| \mathcal{E}_{\vec{u}}(x, \hat{F}_m) - \mathcal{E}_{\vec{u}}(x, F) \right| \longrightarrow 0 \quad a.s.$$

iii) From Central Limit Theorem , for each $\mathcal{C}_x^{\vec{u}} \in \mathcal{C}$

$$m^{\frac{1}{2}} \frac{P_{F,m}\left(\mathcal{C}_x^{\vec{u}}\right) - P_F\left(\mathcal{C}_x^{\vec{u}}\right)}{\sqrt{P_F\left(\mathcal{C}_x^{\vec{u}}\right)\left(1 - P_F\left(\mathcal{C}_x^{\vec{u}}\right)\right)}} \longrightarrow Z, \text{ as } m \longrightarrow \infty,$$

where Z is a random variable with standard normal distribution. According to Definition 5 the previous expression can be rewritten as

$$m^{\frac{1}{2}} \frac{\mathcal{E}_{\vec{u}}(x, \hat{F}_m) - \mathcal{E}_{\vec{u}}(x, F)}{\sqrt{\mathcal{E}_{\vec{u}}(x, F)\left(1 - \mathcal{E}_{\vec{u}}(x, F)\right)}} \longrightarrow Z, \text{ weakly, as } m \longrightarrow \infty.$$

■

5 Financial applications: Multivariate VaR

An important goal for a risk manager is to find the maximum aggregate loss that occur with probability α. Value at Risk (VaR) is the risk measure most used in the univariate case. VaR is the $\alpha-$ quantile of the loss distribution function. If F is the loss distribution and $\alpha \in [0, 1]$ then

$$VaR_\alpha(X) := \inf\{x \in \mathbb{R} \mid F(x) \geq \alpha\}. \tag{15}$$

A natural idea to study risk for portfolio vectors $X = (X_1, \ldots, X_n)$ is to consider a function $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ and a one-dimensional risk measure on $f(X)$. Thus, the VaR of the joint portfolio is that associated to $f(X)$. Examples can be found in Burgert and Rüschendorf [4] where,

$$f(X) = \sum_{i=1}^{n} X_i \text{ or } f(X) = \max_{i \leq n} X_i.$$

The multivariate VaR analogue of univariate VaR is discussed in Embrechts and Pucceti [12] and Cascos and Molchanov [5]. In [12], the VaR is defined through the α- level sets of the joint loss distribution function and the joint loss tail function, while in [5] it is defined through the level sets of depth functions called depth-trimmed regions; more specifically, the multivariate VaR notion is constructed from halfspace trimming regions.

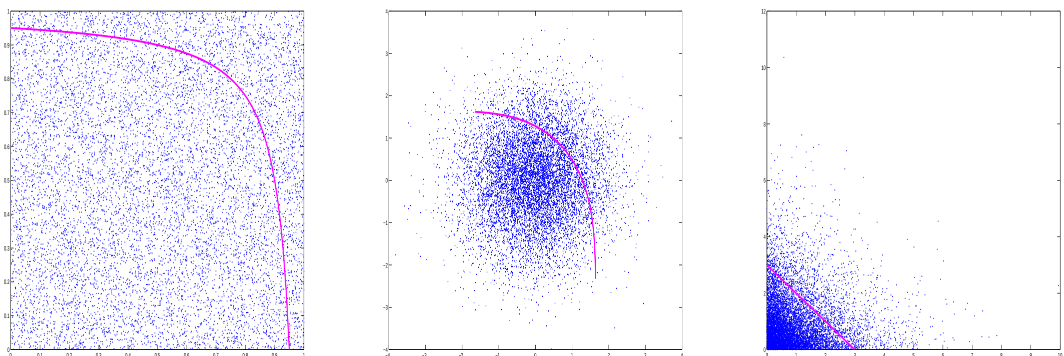
In this section, we introduce a multivariate VaR based on the extremality measure previously defined; indeed, the VaR can be seen as its level sets. If F is a multivariate distribution function, consider the sets

$$A_\alpha^{\vec{u}}(F) = \left\{x \in \mathbb{R}^n : \mathcal{E}_{\vec{u}}(x, F) \geq 1 - \alpha\right\}.$$

The boundary of the set $A_\alpha^{\vec{u}}(F)$ can be interpreted as an *oriented multivariate value at risk* in the level α and denoted by

$$VaR_\alpha^{\vec{u}}(X) = \partial A_\alpha^{\vec{u}}(F). \tag{16}$$

In particular, for $\vec{u} = \vec{e}$ and $\vec{u} = -\vec{e}$, $VaR_\alpha^{\vec{u}}(X)$ is the upper-orthant value at risk and the lower-orthant value at risk, respectively, discussed in Embrechts and Puccetti [12]. However, directions as $\vec{u} = \frac{1}{\sqrt{n]}}[\pm 1, \ldots, \pm 1]'$ and principal components can be interesting in financial applications. Figure 5 shows the VaR of level 0.05 in the direction $\vec{u} = \vec{e}$ for three cases of bivariate distributions with marginals independent and identical distributed as $U(0,1), N(0,1)$ and $Exp(1)$.



Uniforms **Normals** **Exponentials**

Figure 5: ■ Theoretical curve $VaR_{0.05}^{\vec{e}}(X)$ for independent bivariate distributions.

The $VaR_\alpha^{\vec{u}}(X)$ can be estimated nonparametrically by using a multivariate sample $\{x_1, \ldots, x_m\}$ and fitting a surface on the set

$$S_\alpha^{\vec{u}}(F_m) = \left\{ x_i : \mathcal{E}_{\vec{u}}\left(x_i, \hat{F_m}\right) = 1 - \alpha \right\}.$$

However, it may be possible that $S_\alpha^{\vec{u}}(F_m) = \emptyset$ or there are few elements that satisfy the strict equality. To solve this problem, we consider the set

$$S_{\alpha,h}^{\vec{u}}(F_m) = \left\{ x_i : \left| \mathcal{E}_{\vec{u}}\left(x_i, \hat{F_m}\right) - 1 + \alpha \right| \leq h \right\},$$

where h is a slack. Since $S_\alpha^{\vec{u}}(F_m) \subset S_{\alpha,h}^{\vec{u}}(F_m)$, a more accurate estimation of the boundary can be made. The direction given by $\vec{u}$ may have influence in the estimation of $S_{\alpha,h}^{\vec{u}}(F_m)$. Indeed, the classical methods used to smooth functions may fail because the surface of interest is not a function in all the cases. Therefore, to estimate $VaR_\alpha^{\vec{u}}(X)$, we propose to change the original coordinates as follows.

Suppose that $S_{\alpha,h}^{\vec{u}}(F) = \{x_1, x_2 \ldots, x_k\}$. Transforming the set according to (1), we get

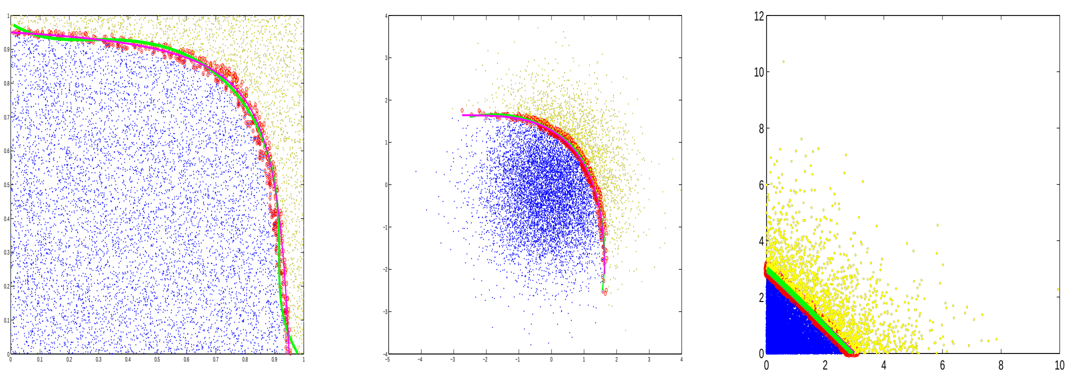
$$\mathcal{R}_{\vec{u}} S_{\alpha,h}^{\vec{u}}(F) = \{\mathcal{R}_{\vec{u}} x_1, \mathcal{R}_{\vec{u}} x_2 \ldots, \mathcal{R}_{\vec{u}} x_k\}. \tag{17}$$

Now, the smoothing of the points in (17) is done by usual methods, and the resulting surface is transformed back to the original system. This is summarized in the following algorithm:

Input:
$\vec{u}$, α, h, and the multivariate sample $\mathbb{X} = (x_1, \dots, x_m)$
for $i = 1$ to m
$\mathcal{E}_i = \mathcal{E}_{\vec{u}}(x_i, \hat{F_m})$
if $|\mathcal{E}_i - 1 + \alpha| \leq h$
$x_i \in S^{\vec{u}}_{\alpha,h}(\hat{F_m})$
end
end
Fitting a function f on $\mathcal{R}_{\vec{u}} S^{\vec{u}}_{\alpha,h}(\hat{F_m})$
$VaR^{\vec{u}}_{\alpha}(X) = \mathcal{R}^{-1}_{\vec{u}} f$

For smoothing we have implemented **gridfit**, a surface modelling tool available in ($http://www.mathworks.com/matlabcentral/fileexchange/8998$).

Figure 6 shows the estimation of the respective theoretical curves $VaR^{\vec{e}}_{0.05}$ drawn in Figure 5. We have considered $h = 0.01$. Here the VaR is calculated in the direction $\vec{u} = \vec{e}$ so $\mathcal{R}_{\vec{u}}$ is the identity matrix. An oriented sub-orthant in direction $\vec{u} = \vec{e}$ and vertex in any yellow point contains a mass less or equal to 0.05. Therefore the extremality of any of those points is greater than 0,95.



Uniform **Normal** **Exponential**

Figure 6: $\mathcal{E}_{\vec{e}}\left(x_i, \hat{F}\right) > 0.95$ $S^{\vec{u}}_{\alpha,h}(F)$ Estimated $VaR^{\vec{e}}_{0.05}$ $VaR^{\vec{e}}_{0.05}$

Figures 7 and 8 show daily negative returns of two leading companies in Spain, since 29/10/2001 until 08/01/2008. In Figure 7 the VaR is estimated in the classical direction $\vec{u} = \vec{e}$; while in Figure 8 the VaR is estimated using $\vec{u} = \vec{pc}$ where $\vec{pc}$ is direction of maximum variability, that is, the direction given by the first principal component of the data. $VaR_{\alpha}^{\vec{pc}}$ can be interpreted a more conservative risk measure.

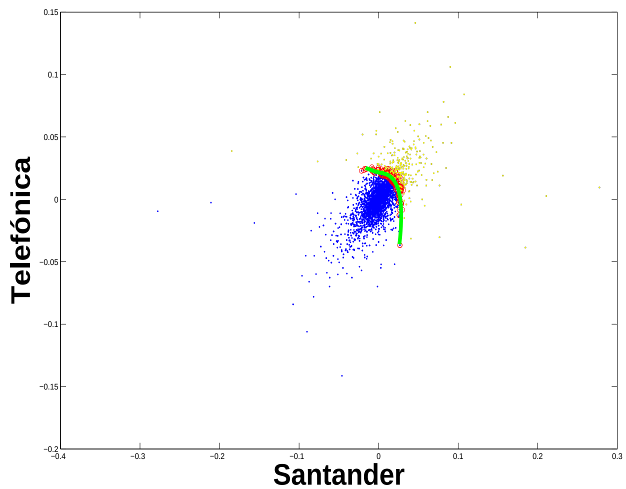


Figure 7: $\mathcal{E}_{\vec{e}}\left(x_i, \hat{F}\right) > 0.95$ Estimated $VaR_{0.05}^{\vec{e}}$

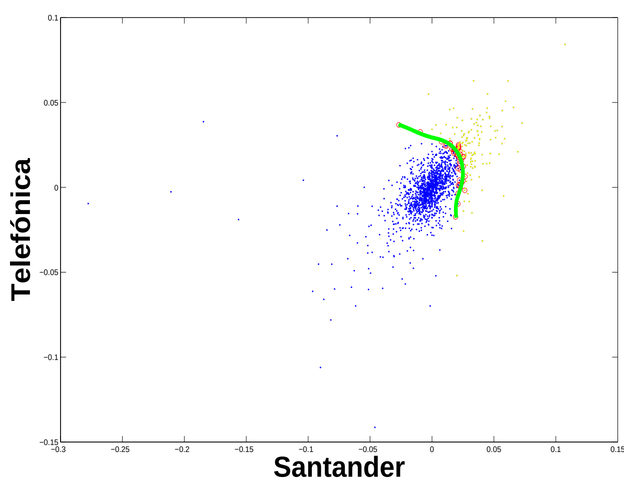


Figure 8: $\mathcal{E}_{\vec{e}}\left(x_i, \hat{F}\right) > 0.95$ Estimated $VaR_{0.05}^{\vec{pc}}$

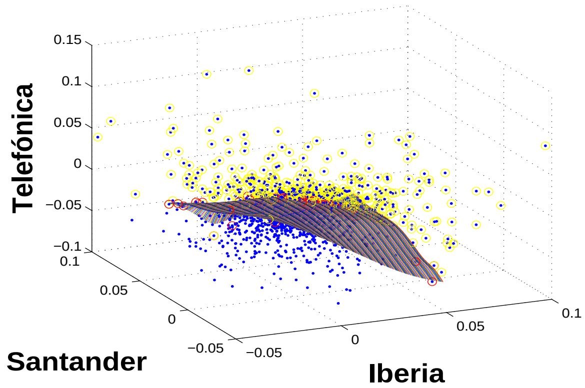


Figure 9: ■ $\mathcal{E}_{\vec{e}}\left(x_i, \hat{F}\right) > 0.95$. |||Estimated $VaR^{\vec{e}}_{0.05}$.

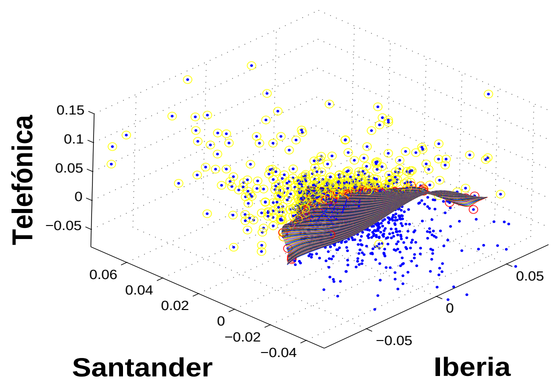


Figure 10: ■ $\mathcal{E}_{\vec{e}}\left(x_i, \hat{F}\right) > 0.95$. |||Estimated $VaR^{pc}_{0.o5}$.

Figures 9 and 10 show estimations of $VaR^{\vec{e}}_{0.05}$ and $VaR^{\vec{pc}}_{0.05}$, respectively, for daily negative returns of three leading companies in Spain, since 29/10/2001 until 08/01/2008. Note that the risk measure depends heavily on the selected direction.

6 Financial applications: ordering portfolios

Next, we propose a strategy based on extremality for selecting portfolios. Instead of the usual optimization techniques, we work with the order induced by extremality in the direction $\vec{u} = \frac{1}{\sqrt{2}}[1, -1]'$.

Consider the general portfolio optimization problem

$$\min_{w} \left[\hat{\rho}(Rw) - \frac{1}{\gamma} f(Rw) \right], \quad s.t. \quad \sum_{i=1}^{N} w_i = 1, \tag{18}$$

where $w \in \mathbb{R}^N$ is the vector of portfolio weights and R is a $M \times N$ data matrix (M denotes number of returns of N assets). $\hat{\rho}$ is a risk measure based on the data; for instance, a risk statistic or a natural risk statistic (Heyde [16] and Shabbir [26]). $f : \mathbb{R}^m \longrightarrow \mathbb{R}$ is a function that quantifies the returns and γ is the risk-aversion parameter. When $\hat{\rho}$ is the variance and $f(x) = \frac{x' 1_M}{M}$, we have the classical portfolio model mean-variance discussed in Markowitz[21]. In this case, problem (18) is reduced to

$$\min_{w} \left[w' \hat{\Sigma} w - \frac{1}{\gamma} \hat{\mu} w \right], \quad s.t. \quad \sum_{i=1}^{N} w_i = 1. \tag{19}$$

The model proposed in Markowitz [21] is relevant in modern portfolio theory which tries to maximize return and minimize risk. Its philosophy is that investors decide portfolio weights based on the trade-off between expected return and risk. Markowitz [21] showed that an investor should hold a portfolio belonging to the intersection of the set of portfolios with minimum variance and the set of portfolios with maximum return. The set of possible options is usually called the efficient frontier, which contains portfolios for which one cannot improve risk and return at same time. When $\gamma \longrightarrow \infty$, the problem (19) corresponds to the minimum-variance portfolio. Note that for different values of γ, we obtain the different mean-variance portfolios on the efficient frontier.

Each w, in the hyperplane $\sum_{i=1}^{N} w_i = 1$ generates a couple $(\hat{\rho}(Rw), f(Rw))$ of feasible portfolios. Figure 11 presents possible combinations of assets plotted in risk-return space. We see that the efficient frontier are the maximum return portfolios for a given level of risk. Conversely, for a given amount of risk, the portfolio lying on the efficient frontier represents the combination offering the best possible return. Portfolios in the A square are better than portfolios in the B square because they have higher returns with less risk. Similarly, portfolios in C are better than portfolios in B, but those in A, C are not comparable in terms of return and risk simultaneously. We propose a method based on the extremality notion that permits a total comparison between feasible portfolios. With our approach, portfolios in A can be compared with portfolios in C.

We now formalize this idea. Let Ω be the set of possible weights for a collection of N assets and let $\mathcal{P}$ given by,

$$\begin{aligned} \mathcal{P} : \Omega &\longrightarrow \mathbb{R}^2 \\ w &\longrightarrow \mathcal{P}_w = (\rho_w, r_w). \end{aligned}$$

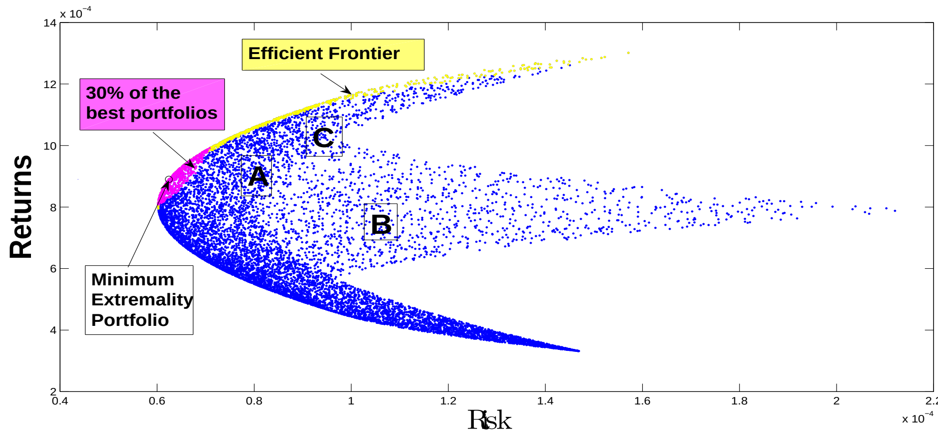


Figure 11: Feasible Portfolios

Let $\mathcal{S} = \mathcal{P}(\Omega) \subset \mathbb{R}^2$ be the set of possible values of (ρ, r). The following definition shows a way to compare portfolios in terms of (ρ, r) values.

Definition 7 *Let w and $w' \in \Omega$ and $\vec{u} = \frac{1}{\sqrt{2}}[1, -1]'$. We say that w dominates w' if, and only if, $\mathcal{P}_w \leq_{\mathcal{E}_{\vec{u}}} \mathcal{P}_{w'}$.*

Several criteria for portfolio selection have been studied in the literature (see DeMiguel et al. [10] and references therein). The simpler strategy to implement is given by $w = \frac{1}{N}1_N$. Almost all models can be expressed as (18). In this case, the solution is a portfolio in S that minimizes the scalar projection on the vector $[1, \frac{1}{\gamma}]'$, where γ is the risk-aversion parameter. The main difference between models is how risk is measured and estimated. It is usual to consider the standard deviation; however, other estimators less unstable can be used for this purpose. For example, DeMiguel and Nogales [9] propose portfolio policies that are based on robust estimators. Constraints on weights are also a difference between models (see Jagannathan and Ma [17], DeMiguel et al [9]). We propose to sort the portfolios according to Definition 7 and to choose the smallest as the optimal portfolio. The following Proposition shows that our strategy can be linked with the Markowitz's solution since the optimal portfolio belongs to the efficient frontier.

Proposition 5 *If $\mathcal{P}_w \leq_{\mathcal{E}_{\vec{u}}} \mathcal{P}_{w'}$ for all $\mathcal{P}_{w'} \in \mathcal{S}$ then $\mathcal{P}_w$ belongs to the efficient frontier.*

Proof. Suppose that $\mathcal{P}_w$ does not belong to the efficient frontier. Then, there is a portfolio $\mathcal{P}_v = (\rho_v, r_v) \in \mathcal{S}$, such that $\rho_v < \rho_w$ and $r_v > r_w$, this means that, for $\vec{u} = \frac{1}{\sqrt{2}}[1, -1]'$,

$$\mathcal{P}_w \in \mathcal{C}^{\vec{u}}_{\mathcal{P}_v} \quad \text{and, from Property 2,} \quad \mathcal{E}_{\vec{u}}(\mathcal{P}_v, \hat{F}_{\mathcal{P}}) < \mathcal{E}_{\vec{u}}(\mathcal{P}_w, \hat{F}_{\mathcal{P}}), \tag{20}$$

which contradicts the fact that $\mathcal{P}_w \leq_{\mathcal{E}_{\vec{u}}} \mathcal{P}_{w'}$. ∎

In order to evaluate the performance of the proposed Minimum- Extremality portfolio selection (MEP), we consider three classical strategies: Mean-variance portfolio with risk aversion parameter $\gamma = 1$ (MEANVAR), Minimum- Variance portfolio with shortsales unconstrained (MINU), and Equality-weighted Portfolio ($\frac{1}{N}$). The comparison is carried out using out-of-sample portfolio mean, out-of-sample portfolio risk, out-of-sample portfolio Sharpe ratio, and portfolio turnovers. We use the technique "rolling-horizon" implemented in DeMiguel and Nogales[8], which depends on a window τ to perform the estimation. For our case of monthly data, $\tau = 120$ which correspond to 10 years. Thus, using the monthly data over the estimation window, we estimate the feasible portfolios set S through the simulation of the five thousand points of the hyperplane $\sum_{i=1}^{N} w_i = 1$ with $w \geq 0$ and we choose the minimum extremality portfolio (MEP) acording to the Definition 7. In each estimation window, we compute (MEANVAR), (MINU) and ($\frac{1}{N}$). This procedure is repeated month to month including the next month and dropping the earliest month until the end of the data set is reached.

Comparison criteria are calculated as follows (see DeMiguel [8], [9]):

- Out-of-sample portfolio mean (Mean)

$$\hat{\mu}^i = \frac{1}{M-\tau} \sum_{t=\tau}^{M-1} (w_t^i)' r_{t+1}, \quad i \in \left\{ \text{MEP, MEANVAR, MINU, } \frac{1}{N} \right\}$$

  where w_t^i denotes the portfolio weight vector chosen at time t under strategy i, r_{t+1} denotes the asset returns at time $t+1$ and M is the sample size

- Out-of-sample portfolio risk (Risk)

$$\hat{\sigma}^i = \left(\frac{1}{M-\tau-1} \sum_{t=\tau}^{M-1} \left((w_t^i)' r_{t+1} - \hat{\mu}^i \right)^2 \right)^{\frac{1}{2}}$$

- Out-of-sample portfolio Sharpe ratio (SR)

$$\widehat{SR}^i = \frac{\hat{\mu}^i}{\hat{\sigma}^i}$$

- Turnovers

$$\text{Turnover1} = \frac{1}{M-\tau-1} \sum_{t=\tau}^{M-1} \sum_{j=1}^{N} \left(|w_{j,t+1}^i - w_{j,t}^i| \right)$$

$$\text{Turnover2} = \frac{1}{M-\tau-1} \sum_{t=\tau}^{M-1} \sum_{j=1}^{N} \left(|w_{j,t+1}^i - w_{j,t^+}^i| \right),$$

  where $w_{j,t}$ is the portfolio weight of the asset j chosen at time t. The Turnover1 can be interpreted as a measure of the stability between weights in each period for each asset. For example, in the case of $\frac{1}{N}$ strategy, the Turnover1= 0. The smallest Turnover1 gives the idea of more stable strategy and this strategy has the advantage of being a credible strategy for the investor. In the Turnover2, w_{j,t^+}^i is the portfolio weight before rebalancing, but an $t+1$ and $w_{j,t+1}^i$ the desired portfolio weight at time $t+1$, after rebalancing. Unlike Turnover1 with

$\frac{1}{N}$ strategy, Turnover2 may be different to zero due to changes in asset prices between t and $t+1$. Turnover2 can be interpreted as the average percentage of wealth traded period to period and is related to transaction costs.

Regarding the data sets, we have used some portfolios that are quite popular among practitioners and have been selected from Kenneth French web-site *http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html*

These data are monthly asset returns and are presented in the Table 1, with the abbreviation used to refer to each data set, the number of assets in each data set and the time period.

Data Sets	**Abrev.**	**N**	**Period**
5 industry portfolio representing the U.S. stock market	5Ind	5	07/1963-12/2004
6 Fama and French portfolios sorted by size and book-to-market	6FF	6	07/1963-12/2004
10 industry portfolio representing the U.S. stock market	10Ind	10	07/1963-12/2004
25 Fama and French portfolios sorted by size and book-to-market	25FF	25	07/1963-12/2004
30 industry portfolio representing the U.S. stock market	30Ind	30	07/1963-12/2004
48 industry portfolio representing the U.S. stock market	48Ind	48	07/1963-12/2004

Table 1: Data sets of monthly asset returns.

Tabla 2 reports the out-of-sample criteria for the portfolios considered and the strategies analyzed. Some points can be addressed

- **Out-of-sample portfolio mean** We see that the out-of-sample portfolio mean of the sample based in minimum extremality strategy (MEP) is often higher than other strategies. This is attractive to investors, but it is not the main criterion for choosing a good strategy, since investors want to take into account the risk.

- **Out-of-sample portfolio risk** We have measured the risk with the standard deviation and we find that out-of-sample risk of the minimum extremality portfolio (MEP) strategy is much lower than that mean-variance strategy for all portfolios considered, even in the $\frac{1}{N}$ strategy for (5Ind, 6FF, 30Ind, 48Ind) our strategy is not improved.

- **Out-of-sample portfolio Sharpe ratio** Comparing the Sharpe ratios of the MEP satretgy, we see that it has higher Sharpe ratio than the equally weight ($\frac{1}{N}$) for all data sets, has higher Sharpe ratio than mean-variance strategy except in (6FF) and it has higher Sharpe ratio for the two largest portfolios. To have a Sharpe ratio better than ($\frac{1}{N}$) is considered a good benchmark.

- **Turnovers** Obviously $\frac{1}{N}$ has a Turnover1$=0$ since portfolio weights are equal period to period. Turnover1 of (MEP) strategy is much lower than that mean-variance strategy for all portfolios considered; this is not surprising because as explained in DeMiguel and Nogales [9] the mean-variance strategy generates portfolio weights extremely unstable. However the turnover1 of (MINU) and (MEP) strategies reflects that the portfolio weights obtained by (MEP) fluctuates less in the portfolios of larger number of assets. Also the best portfolio

5Ind Portfolio					
Strategy	**Mean**	**Risk**	**S.R.**	**Turnover1**	**Turnover2**
MEP	0.0112	0.0455	0.2454	0.2077	0.2122
MINU	0.0107	0.0432	0.2486	0.0870	0.1282
MEANVAR	0.0077	0.0884	0.0868	0.4543	0.4987
$\frac{1}{N}$	0.0111	0.0460	0.2411	0	0.0364

6FF Portfolio					
Strategy	**Mean**	**Risk**	**S.R.**	**Turnover1**	**Turnover2**
MEP	0.0139	0.0478	0.2908	0.4722	0.4781
MINU	0.0144	0.0406	0.3537	0.1722	0.2257
MEANVAR	0.0191	0.0603	0.3165	0.6724	0.8429
$\frac{1}{N}$	0.0130	0.0490	0.2641	0	0.0389

10Ind Portfolio					
Strategy	**Mean**	**Risk**	**S.R.**	**Turnover1**	**Turnover2**
MEP	0.0117	0.0448	0.2612	0.3959	0.4020
MINU	0.0108	0.0376	0.2861	0.1450	0.1708
MEANVAR	0.0065	0.0846	0.0760	0.6394	0.7013
$\frac{1}{N}$	0.011	0.0435	0.2560	0	0.0344

25FF Portfolio					
Strategy	**Mean**	**Risk**	**S.R.**	**Turnover1**	**Turnover2**
MEP	0.0161	0.0560	0.2878	0.5188	0.5295
MINU	0.0159	0.0390	0.4076	0.7314	0.7923
MEANVAR	0.0215	0.1113	0.1932	3.4864	4.1988
$\frac{1}{N}$	0.0135	0.0509	0.2660	0	0.0402

30Ind Portfolio					
Strategy	**Mean**	**Risk**	**S.R.**	**Turnover1**	**Turnover2**
MEP	0.0127	0.0479	0.2644	0.2808	0.2881
MINU	0.0097	0.0404	0.2410	0.4521	0.4821
MEANVAR	0.0060	0.1061	0.0568	1.9354	2.2062
$\frac{1}{N}$	0.0116	0.0479	0.2433	0	0.0374

48Ind Portfolio					
Strategy	**Mean**	**Risk**	**S.R.**	**Turnover1**	**Turnover2**
MEP	0.0120	0.0482	0.2495	0.4341	0.4399
MINU	0.0086	0.0442	0.1944	0.7779	0.8124
MEANVAR	0.0056	0.2401	0.0233	3.5479	4.7163
$\frac{1}{N}$	0.0117	0.0488	0.2390	0	0.0383

Table 2: Out-of-sample performance for portfolios considered

in terms of Turnover2 are $\frac{1}{N}$. We can see in Table 2 that our proposal (MEP) has lower transaction costs than (MEANVAR) for all portfolios considered and lower transaction costs than (MINU) for (25FF, 30Ind, 30Ind, 48Ind).

7 Conclusions

This paper introduces an extremality measure in multivariate problems analysis which induces a natural order in $\mathbb{R}^n$, allowing an extension of the quantiles studied in Tibiletti [27]. The multivariate Value at Risk defined in this paper generalizes those given in Embrechts and Pucceti [12] and Tibiletti [27] by the inclusion of directions $\vec{u}$. We give a portfolio selection strategy based on order induced for extremality. Generally is quite difficult to find a strategy that delivers both a high Sharpe ratio and low Turnover. Nevertheless, from Table 2, (MEP) strategy offers a good behaviour in both criteria respect to (MINU) and (MEANVAR) strategies in the portfolios of larger number of assets.

References

[1] Arbia, G. (2002). Bivariate value at risk. *Statistica, anno LXII.* 2, $231 - 247$.

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

[3] Barnett, V. (1976). The Ordering of multivariate data. (with comments), *Journal of the Royal Statistical Society,* Ser. A, 139, 318-354.

[4] Burgert, C. and Rüschendorf, L. (2006). Consistent risk measures for portfolio vectors. *Insurance: Mathematics and Economics.* 38, 289-297.

[5] Cascos, I. and Molchanov, I. (2007). Multivariate risk and depth-trimmed regions. *Finance and Stochastic.* 11, 373-397.

[6] Cerioli A. and Riani M. (1999). The ordering of spatial data and the detection of multiple outliers. *Journal of Computational and Graphical Satistics.* pp. 1-20 ISSN: 1061-8600.

[7] Chaudhuri, P. (1996). On a geometric notion of quantiles for multivariate data. *Journal of the American Statistical Association.* 91, 862 872.

[8] DeMiguel, V., Garlappi, L., Nogales, F.J. and Uppal, R. (2009). A Generalized approach to portfolio optimization: improving performance by constraining portfolio norms. *Management Science.* 55, 798-812.

[9] DeMiguel, V. and Nogales, F.J. (2009). Portfolio selection with robust estimation *Operations Research.* 57, 560-577.

[10] DeMiguel, V., Garlappi, L. and Uppal, R. (2009.) Optimal versus naive Diversification: How Inneficient is the $1/N$ portfolio Strategy?. *The Review of Financial Studies.* 22, 1915-1953.

[11] Einmahl, J. H. and Mason, D. M. (1992). Generalized quantile processes. *Annals of Statistics.* 20, 1062-1078.

[12] Embrechts, P., Pucceti, G. (2006). Bounds for functions of multivariate risks. *Journal of Multivariate Analysis* 97, 526-547.

[13] Gaenssler, P. (1983) *Empirical processes.* Institute of Mathematical Statistics, Hayward, California.

[14] Gentle, J.E. (1998). *Numerical linear algebra for applications in statistics.* Statistics and Computing Springer.

[15] Hallin, M. Paindaveine, D. and Šiman, M. (2010). Multivariate quantiles and multiple-output regression quantiles: from L_1 optimization to halfspace depth. *Annals of Statistics.* 38, 635-669.

[16] Heyde, C.C., Kou, S.G. and Peng, X.H. (2007). What is a good external risk measure: bridging the gaps between robustness, Subadditivity, and Insurance Risk Measure (Preprint).

[17] Jagannathan, R. and Ma, T. (2003). Risk reduction in large portfolios: why imposing the wrong constraints helps. *Journal of Finance.* 58, 1651-1684.

[18] Kong, L. and Mizera, I. (2008). Quantile tomography: Using quantiles with multivariate data. preprint.

[19] Liu, R. Y., Parelius, J.M. and Singh, K. (1999). Multivariate analysis by data depth: descriptive statistics, graphics and inference. *Annals of Statistics.* $27, 783 - 858$.

[20] Li, J. and Liu, R.Y. (2008). Multivariate spacing based on data: construccion of nonparametric multivariate tolerance regions. *Annals of statistics.* 36, $1299 - 1323$.

[21] Markowitz, H. M. (1952). Mean-variance analysis in portfolio choice and capital markets. *Journal of Finance.* 7, 77-91.

[22] Marshall, A.W. and Olkin, I. (1979). *Inequalities: theory of majorization and its applications.* New York: Academic Press.

[23] Nappo,G. and Spizzichino, F. (2009). Kendall distributions and level sets in bivariate exchangeable survival models (in press)

[24] Rockafellar, R.T. (1997). *Convex analysis.* princenton University Press.

[25] Serfling, R. (2002). Generalized quantile processes based on multivariate depth function, with applications in nonparametric multivariate analysis. *Journal of multivariate analysis.* vol 83, $232 - 247$.

[26] Shabir, A., Filipovic, D. and Svindland, G. (2008). A Note on Natural Risk Statistics. *Operations Research Letters* 36, 662-664

[27] Tibiletti, L. (1993). On a new notion of multidimensional quantile. *Metron* vol. LI n. $3 - 4$, $77 - 83$.

[28] Tibiletti, L. (2001). Incremental value at risk: traps and misinterpretations, in trends in Mathematics, (M. Kohlmann ed.) Birkhauser Verlag, Basel (Switzerland), 355-364.

[29] Zani, S., Riani M. and Corbellini A. (1999). New methods for ordering multivariate data: an application to the performance of investment funds. *Applied Stochastic Models and Data Analysis.* 15, 485-493.

[30] Zuo, Y. Serfling, R. (2000). General notions of statistical depth function *Annals of statistics* 28, $461 - 482$.