Universidad Carlos III de Madrid

TESIS DOCTORAL

Extremality in Multivariate Statistics

Autor:
Henry Laniado Rodas

Director/es:
Rosa E. Lillo y Juan Romo

DEPARTAMENTO DE ESTADÍSTICA

Leganés, junio de 2012

TESIS DOCTORAL

Extremality in Multivariate Statistics

Autor: Henry Laniado Rodas

Directores: Rosa E. Lillo y Juan Romo

Firma del Tribunal Calificador:

Firma

Presidente: (Nombre y apellidos)

Vocal: (Nombre y apellidos)

Secretario: (Nombre y apellidos)

Calificación:

Leganés, de de

UNIVERSIDAD CARLOS III DE MADRID

Ph.D. Dissertation

Extremality in Multivariate Statistics

Henry Laniado Rodas

Advisors:

Rosa E. Lillo
Juan Romo

Department of Statistics

Leganés, June 2012

This thesis was performed in the Statistics Department of the Universidad Carlos III de Madrid Campus de Leganés under the direction of the Professors Rosa E. Lillo Rodríguez and Juan Romo. The author was supported by a scholarship for master studies (UC3M) and was subsequently hired as a teaching and research staff (PDI). Besides he had the partial support of the following research projects: COMUNIDAD DE MADRID-UC3M 2009/00147/001, COMUNIDAD DE MADRID-UC3M 2011/00068/001, MINISTERIO DE CIENCIA E INNOVACIÓN 2009/00035/001, MINISTERIO DE CIENCIA E INNOVACIÓN 2012/00084/001.

Writing and Editing of Thesis

with $\LaTeX\,2_\varepsilon$, *WinEdt*

and Operating System

Windows 7

♡ ♡ ♡ ♡

A HEDANIFAHE

mi familia numerosa

♡ ♡ ♡ ♡

Si la gente no cree que las matemáticas son simples, es sólo porque no se dan cuenta de lo complicada que es la vida.

Von Neumann

La cuestión en la vida no es saber mucho, sino olvidarse de poco.

Homero Expósito

Las ciencias tienen las raíces amargas, pero muy dulces las frutas.

Aristóteles

Lo único difícil para mí es lo imposible.

Nicolás Laniado Valencia

Agradecimientos

Quiero agradecer profundamente a mis directores, los profesores Rosa E. Lillo y Juan Romo por la confianza que me depositaron desde el primer momento para realizar la propuesta, el desarrollo y la conclusión de este trabajo. Gracias por su apoyo, sus enseñanzas, sus sabios consejos y seguimiento que me dieron durante mi formación, así como la libertad que me permitieron tomar para la elaboración de este trabajo de investigación. Algo fundamental para mi formación como investigador ha sido su orientación que me ha permitido cumplir con las publicaciones y participaciones académicas requeridas para alcanzar el grado de doctor. Muchas gracias por haberme compartido de corazón su gran experiencia que a lo largo de varios años han acumulado como investigadores.

A Dios y a mis padres Sirley Rodas y Mayer Laniado†, ellos tres comenzaron este proyecto poco más de 36 años. Especialmente a mi madre agradecerle por sus oraciones y cuidados de los cuales, afortunadamente aun gozo y por su educación que ha rendido frutos en mi formación profesional y en mi formación como ser humano. Esta tesis es insignificante a lo que tengo que hacer para merecerla.

Agradecer a las 4 personas más importantes de mi vida, mi Esposa Dalia que a lo largo de casi 15 años ha sido un pilar de amor para mí, gracias por tu paciencia y soportar que este proyecto usurpara parte de mi tiempo que era para tí. Me alegra saber que tu compañía se extenderá mucho más allá de este periodo. A mis Hijos: Nicolás, Fabiana y Helena, ellos son la motivación de mi lucha y mi inspiración, sin ellos, este documento no hubiese llevado su ejemplo, de algo bien hecho y agradable.

Para iniciar mis estudios doctorales ha sido determinante mi hermano Edwin y su esposa Yanitza, ellos han sido los primeros en motivarme y sugerirme hacer una carrera como investigador y en ese sentido me animaron para realizar mi doctorado. A mi hermano Marcel y su esposa Érika gracias por la compañía en la recta final de mi tesis.

Mi gratitud al Departamento de Estadística por el apoyo financiero y el inmejorable ambiente de trabajo. A su planta de profesores y también al personal administrativo: Gema, Susana y Paco por su apoyo que ha ido desde el préstamo de material, dotación, insumos y asesoría en los trámites hasta depositar la Tesis.

Quiero agradecer a mis compañeros de batalla, Sofía y Ana Paula, y también a los que ya ganaron la batalla, Alba, Lee, Miguel ("Mingue"), Nuria y Ángel, de quienes aprendí muchos secretos de como ganar la mía.

Quiero expresar especial agradecimiento al Profesor Franco Pellerey del Politecnico di Torino por su excelente asesoría, dedicación y hospitalidad en mi visita a Turín. Al Profesor Froilan Martínez por su valiosa tutoría sobre rotaciones, a Ignacio Cascos por sus oportunas tutorias sobre riesgos.

A todas las personas nombradas muchas gracias.

Contents

List of Figures

Indice of tables

Extremality in Multivariate Statistics

Ph.D. Dissertation

Abstract

Henry Laniado Rodas

Department of Statistics

Universidad Carlos III de Madrid

Multivariate order is a valuable tool for analyzing data properties and for extending univariate concepts based on order such as median, range, extremes, quantiles or order statistics to multivariate data.

Generalizing such concepts to the multivariate case is not straightforward. While different ways of generalizing quantiles have been studied by Chaudhuri [10], a description of extensions of concepts such a median, range and quantiles to the multivariate framework has been provided by Barnett [3]. The key problem, however, in generalizing these concepts to several dimensions is the lack of a unique criterion for ordering multivariate observations.

Over the last few decades, multivariate stochastic orders have also become a powerful means of comparing random vectors, especially in situations where the distributions are partially known. In particular, multivariate stochastic orders have a wide range of applications in portfolio theory.

The thesis is motivated by the aspects mentioned above and its purpose is threefold. Firstly it introduces the multivariate extremality as a methodology that measures the farness of a point $\mathbf{x}$ with respect to a data cloud or a distribution function. We study the main properties of this new concept, as well as asymptotic results, and define a new multivariate data order based on rotations. This data order allows us to introduce a new version of the multivariate quantile which can be seen as a generalization of definitions previously studied in the literature. As a consequence of this ordering, we are able to develop an application in finance by defining a new version of the multivariate Value at Risk. Secondly, we develop a new multivariate stochastic order based on directions for generalizing the well- known orthant orders. Some examples are presented of applications in portfolio comparisons. Particular attention is paid to applications in which the use of directions is well justified in determining optimal allocations of wealth among risks in single period portfolio problems. Thirdly, the thesis aims to investigate an alternative methodology for selecting the portfolio weights in a data set that represents returns of n assets for

investing. We also define new concepts of efficient frontiers based on the initial idea of Markowitz. We apply the extremality multivariate data order to order feasible portfolios in a direction that depends on specific indexes; these may be chosen by the investor and may be different from the classical variance-return in Markowitz's model.

The thesis is organized as follows: in Chapter 1 we provide a brief review of some multivariate data orders introduced in the literature in order to extend univariate statistical concepts to the multivariate setting. Following some multivariate stochastic orderings, we examine different means of comparing random vectors based on their survival and distribution functions. Finally, the chapter presents a brief introduction to the portfolio selection problem.

In Chapter 2, we propose a new approach for analyzing multivariate extremes. It provides a multivariate data order based on a concept that we will call "extremality". We establish the most relevant properties of this extremality measure and we give the theoretical basis for its nonparametric estimation. Finally, we include an application in Finance, we define an oriented multivariate Value at Risk (VaR) with level sets built through extremality which is computationally feasible in high dimensions.

The results of Chapter 3 concern a new multivariate stochastic order that compares random vectors in a direction which is determined by a unit vector, generalizing the well-known upper and lower orthant orders. The main properties of this new order, together with its relationships with other multivariate stochastic orders are investigated. We also present some examples of application in the determination of optimal allocations of wealth among risks in single period portfolio problems.

In Chapter 4, we introduce new concepts of efficient frontier that depend on some indexes that may be chosen by the investor and that are different from the classical variance- return in Markowitz's model. Feasible portfolios are built with MonteCarlo simulations and the new efficient frontiers are estimated by using an extremality order to sort portfolios. The performance of the selection method is illustrated with real data.

Finally, in Chapter 5, we present some general conclusions and summarize the main contributions of the thesis.

Extremalidad en Estadística Multivariante

Tesis Doctoral

Resumen

Henry Laniado Rodas

Departamento de Estadística

Universidad Carlos III de Madrid

Un orden multivariante es una valiosa herramienta para analizar propiedades de un conjunto de datos y extender algunos conceptos estadísticos como mediana, rango, atípicos, quantiles y estadísticos de orden al escenario multivariante.

Generalizar dichos conceptos al caso multivariante no ha sido sencillo. Mientras diferentes formas de generalizar quantiles han sido estudiadas por Chaudhuri [10], una descripción de los conceptos tales como mediana, rango y quantiles para el escenario multivariante ha sido proporcionado por Barnett [3]. El principal problema en la generalización de estos conceptos a mayores dimensiones es la dificultad de definir un orden total para observaciones multivariantes.

En las últimas décadas los órdenes estocásticos multivariantes también se han convertido en una sofisticada metodología para comparar vectores aleatorios, especialmente en situaciones donde sus distribuciones son parcialmente conocidas. En particular, los órdenes estocásticos multivariantes han tenido un amplio campo de aplicaciones en la teoría de portfolios.

La tesis es motivada por los aspectos mencionados anteriormente y propone tres objetivos principales. Primero se introduce el concepto de extremalidad como una metodología que permite medir la lejanía de un punto $\mathbf{x}$ con respecto a una nube de puntos o a una función de distribución. Se estudian tanto las propiedades principales de este nuevo concepto como resultados asintóticos y se define un nuevo orden para datos multivariantes basado en rotaciones. El orden de datos permite introducir una nueva versión de quantil multivariante que puede ser visto como una generalización de definiciones anteriormente estudiadas en la literatura. Como una consecuencia de este orden para datos, se desarrolla una interesante aplicación en finanzas al definir una nueva versión del Valor en Riesgo multivariante. Como segundo objetivo se desarrolla un nuevo orden estocástico multivariante basado en direcciones que generaliza los bien conocidos orthant orders. Algunos ejemplos de aplicación son presentados sobretodo en comparasión de portfolios y en la determinación de las proporciones de cierto capital disponible el cual es destinado a ser invertido entre

activos con riesgo. El tercer objetivo de la tesis es el desarrollo de una metodología alternativa para seleccionar los pesos en un portfolio. Se defienen también nuevos conceptos de fronteras eficientes basadas en la idea inicial de Markowitz. Se aplica el orden de extremalidad multivariante a fin de ordenar portfolios factibles en una dirección la cual depende de algunos criterios elegidos por el inversor y que pueden ser diferentes de los criterios clásicos de varianza-media del modelo de Markowitz.

La tesis es organizada de la siguiente manera: en el Capítulo 1 se proporciona una breve revisión de algunos órdenes multivariantes introducidos en la literatura para extender conceptos estadísticos univariantes al escenario multivariante. También se hace una revisión de algunos órdenes estocásticos univariantes y multivariantes. Finalmente, el capítulo presenta una introducción al problema de selección de portfolios.

En el Capítulo 2, se propone un nuevo enfoque para analizar extremos multivariantes. Se introduce un orden para datos multivariantes basado en un concepto que es llamado "extremalidad". Se establecen las propiedades más importantes de la medida de extremalidad y se proporciona la base teórica para su estimación no paramétrica. Finalmente, se incluye una aplicación en finanzas definiendo un Valor en Riesgo Multivariante Orientado el cual es computacionalmente factible en altas dimensiones.

Los resultados del Capítulo 3 se refieren a un nuevo orden estocástico multivariante que permite comparar vectores aleatorios en una dirección determinada por un vector unitario. Este orden estocástico puede ser visto como una generalización de los bien conocidos Upper and Lower orthant orders. La principales propiedades de este nuevo orden y su relación con otros órdenes multivariantes son investigadas. Se presentan algunos ejemplos de aplicación en comparación y selección de portfolios.

En el Capítulo 4 se introduce nuevos conceptos de fronteras efficientes que dependen de algunos índices elegidos por el inversor y que pueden ser diferentes a los clásicos de varianza-media del modelo de Markowitz. Se construyen portfolios factibles a través de simulación MonteCarlo y las nuevas fronteras efficientes son estimadas mediante un orden de extremalidad que permite ordenar portfolios factibles. El método es evaluado con datos reales y se comparan los resultados con otras estrategiás ya introducidas en la literatura.

Finalmente, en el Capítulo 5, se presentan las conclusiones y se resumen las principales contribuciones de la tesis.

1 Introduction and background

In this dissertation, we study a multivariate data order and a multivariate stochastic order that provide useful new alternatives for comparing multivariate observations and multivariate random vectors, respectively. The comparison proposals studied are designed for the multivariate framework and the applications are focused on the financial setting.

We have investigated new ways of comparing data and random vectors through directions. In the univariate case just two directions are allowed for comparison (left and right), whereas in higher dimensions there is an infinite number of directions that can be determined by a unit vector. This work takes advantage of the different directions that can be defined on the Euclidean space in $\mathbb{R}^n$, thus enabling us to introduce a novel way of ordering multivariate data and to compare multivariate distributions from a directional approach. We provide some applications in finance where the use of a directional order is clearly justified.

This Chapter gives an overview of the basic concepts, terminology and related work regarding the topics of this dissertation. The thesis has developed a threefold approach:

- Firstly, to introduce the concept of extremality in order to sort multivariate data from different directions, to study multivariate extremes from a directional perspective and, as a consequence, to generalize some concepts of multivariate quantiles, and finally, to provide a new version of the multivariate Value at Risk.

- Secondly, to define a new multivariate stochastic order for generalizing the upper and lower orthant order through the inclusion of directions, and to study its main properties, its relationship with other orders from the literature and to present some of its applications in a portfolio selection.

- Thirdly, to propose a strategy based on the extremality order for selecting portfolios instead of using strategies based on the usual optimization techniques, and to compare portfolios through the ordering induced by the extremality when a direction is chosen by the economic agent.

Next, we briefly introduce some multivariate data orders which have previously been studied in the literature as a basic concept that expresses intrinsic natural features of a multivariate data set.

1.1 Multivariate data orders

The univariate order has several good properties that facilitate the study of statistical features for univariate data sets. It is well known, for example, that the usual order relation on the real line is a total order relation and this property simplifies the task of performing statistical analysis in one-dimensional data. In the multivariate case, however, the total order property does not hold and the lack of an objective basis for ordering multivariate observation has been a significant problem in attempts made to extend some one-dimensional statistical concepts to the multidimensional case.

Despite the absence of total order relations for multivariate data, there are many works in the traditional literature which attempt to extend univariate order and statistical concepts such as medians, extremes and ranges to the higher dimensional situations.

1.1.1 Barnett ordering

Barnett [3] describes extensions of concepts such as median, range and quantiles to the multivariate framework. He states that although there is no natural order for multivariate data, most univariate statistics that are naturally defined as transformations of order statistics, can be extended to higher dimensions. He has also investigated the possible types of multidimensional orderings and has retained four variants: the marginal ordering, which is equivalent to the separated ordering of each component, the reduced ordering, which orders vectors according to some scalars computed from the components of each vector, the partial ordering which uses the convex hull and the conditional ordering, which is the scalar ordering of a single component. Here we briefly comment upon three of these variants.

- ***Marginal ordering***

  This type of ordering depends on the ranking with respect to one or more of the marginal samples. The multivariate sample mean is based on this type of ordering since it only depends on the marginal sample means. The same does not occur with the variance where the correlations among the marginals have to be considered.

  The marginal median vector can be taken as a first approximation to the multivariate median and even multivariate quantiles may be constructed through marginal quantiles. However, there are more sophisticated techniques for defining multivariate quantiles. See for instance, Einmahl and Mason [15], Tibiletti [58], Chaudhuri [10], Serfling [54], Fernández-Ponce and Suárez-Llorens [19], Belzunce et al. [4] and Hallin et al. [25] for a thorough treatment of multivariate quantiles.

  As its name suggest, Marginal ordering permits statistical concepts in a multivariate sample to be extended via the univariate concepts of the marginal samples.

- ***Reduced ordering***

  This type of ordering consists of reducing each multivariate observation to a single value by means of some combination of the components of the sample values. A simple example is ordering the multivariate sample through its distance to origin. Frequently the metric employed to sort data is the generalized distance represented by a quadratic function $(\mathbf{x}-\alpha)\Gamma^{-1}(\mathbf{x}-\alpha)$ for some convenient choice of α and Γ. The value α can be the origin, the mean, the marginal median vector, or some interesting observation and Γ can be the identical matrix whose orders will be established by the Euclidean distance to α. Γ also may be the variance and covariance matrix; in this case, if α is the mean vector, we recover the Mahalanobis distance. Both situations are displayed in Figure 1.1 where two types of reduced ordering are illustrated. Order on the the left side is established by Euclidean distance while order on the right side is induced by Mahalanobis distance. The concentric circles or ellipses, would be the corresponding bases on which the order in each data would be evaluated, i.e, the points on the wider circles or ellipses should have a higher order and therefore should be more extremes.

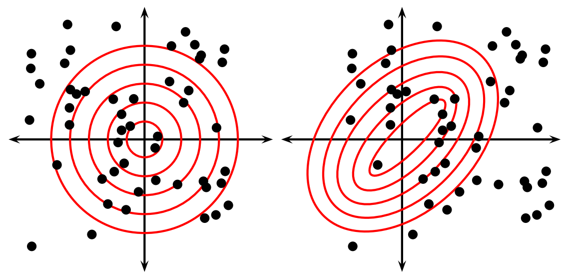


Figure 1.1: Reduced ordering

Most examples of ordering after the data has been transformed also are considered to be in the reduced order class. This is because such examples are not intended to represent marginal behavior but to express overall characteristics of the multivariate data. For instance, via either the maximum or the mean of each observation.

- ***Partial ordering***

  Here the multivariate sample is divided into distinct groups of different order. These type of orders neither consider the marginal samples nor the individual multivariate observations by considering the joint properties in the total sample set in which all individual observations are ranked in relation to other observations. One way in which the multivariate observations may be ranked is by defining different regions of the sample space where such partitioning may be based on the convex hull (one of the most important and intuitive methods used for distinct groups of observations).

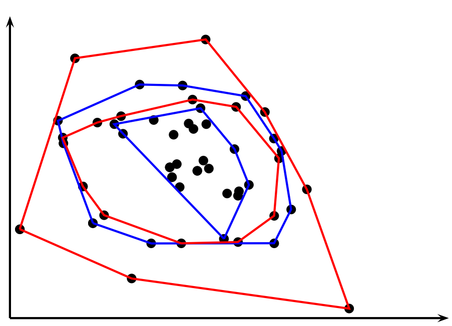


Figure 1.2: Partial ordering

Figure 1.2 shows a multivariate sample in which the convex hull is constructed by drawing a minimum convex set which covers all sample points. While those points on the perimeter are discarded, the convex hull is formed out of the remaining points, and the second perimeter is deleted again. The process is repeated, thus providing a method of dividing the data into ordered groups. Population version of the convex hull may not be too tractable, therefore some analogous methodologies as depth functions with distributional results have been proposed. Some depth functions are mentioned in the following section.

1.1.2 Order through depth functions

A depth notion can be used to measure the centrality or outlyingness of a point from sample with respect to the multivariate sample or to its underlying distribution. As a consequence, this leads a natural center-outward ordering of the sample points. Based on this ordering, a wide range of statistical univariate techniques such as multivariate goodness-of-fit, location measure, scatter estimates and risk measurement can be extended to the multivariate setting. A review on depth functions and its applications can be found in Cascos [7].

A depth function can be defined by a mapping $D : \mathbb{R}^n \longmapsto [0,1]$ that satisfies the properties of *affine invariance, vanishing at infinity, monotonicity with respect to the deepest point and maximality at center*. Here, we describe briefly two classical depth functions named *halfspace depth and simplicial depth*. The first of these was proposed by Tukey [60] in a data analysis context. Given a multivariate sample, the halfspace depth of a point x is the smallest fraction of data points in a closed halfspace containing x, or also the smallest fraction of data points that should be deleted so that x lies outside the convex hull of the remaining data points.

The second of these functions was introduced by Liu [40] and it is based on random simplices. The simplicial depth of a point x is given by the probability that the point x is contained inside a random simplex whose vertices are $p+1$ independent observations.

Figure 1.3 shows some level sets of the halfspace depth and a grey scale for the simplicial depth. The data represent the results of athletes that competed in Olympic Games, in Barcelona in 1992 in long jump (in meters, axis Y) and in the 200m race (in seconds, axis Y). Figure 1.3 is taken from Figure 1 in Cascos et al. [8].

An empirical version of those depth functions provides a natural ordering of the data points from the center outward. The ordering thus obtained leads to the introduction of multivariate generalizations of the univariate sample median and also L-statistics. However, they have the drawback of not being computationally feasible

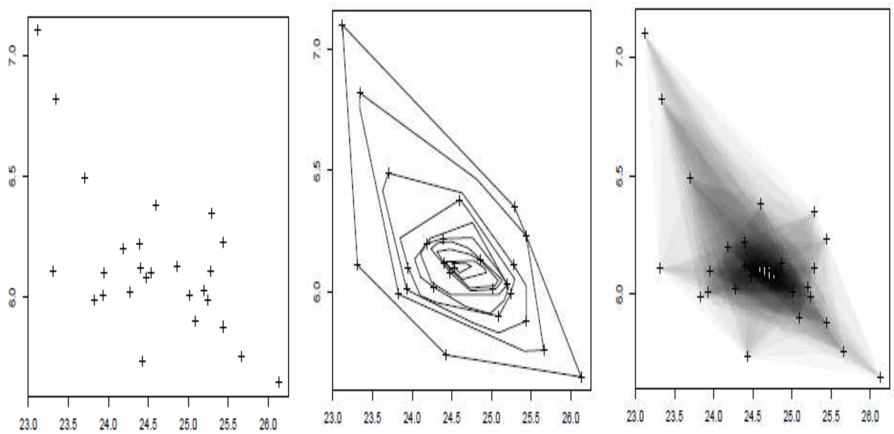


Figure 1.3: Decathlon data, halfspace and simplicial depth. Cascos et al. [8]

in high dimensions, hence the multivariate order induced through depth functions is quite limited for more than three dimensions.

Nevertheless, we want to emphasize that one of the proposals introduced in Chapter 2 of this work is a multivariate data order that works well in terms of computational times in high dimensions and for large samples. This enables new versions of multivariate quantiles to be built in high dimensions and means that new alternatives which are computationally feasible can be introduced which will be useful in order to define new version of a multivariate Value at Risk.

1.2 Multivariate stochastic orders

Stochastic orders are partial order relations on a set of distribution functions. They have become valuable tools in finance, economic, queueing theory, reliability, statistics and insurance, among many other areas. The first attempt to define a stochastic order was made by using means and variances to compare random variables. This, however, did not prove to be exhaustive since comparison between two single numbers is often not very informative. What is more, many distributions need more information to be characterized and compared. Notwithstanding, several effective ways of comparing random variables and even random vectors have appeared over the last thirty years.

Indeed, a second area of concern of this thesis is a discussion of a new alternative to compare random vectors. For ease of reference, we briefly recall some traditional concepts in this area. Concerning the stochastic comparisons, we first provide the definitions of the orders which are usually considered in a univariate setting.

Definition 1.2.1 *Given two random variables X and Y we say that X is smaller than Y in:*

- usual stochastic order *(denoted by $X \leq_{st} Y$) if, and only if, $\bar{F}_X(t) \leq \bar{F}_Y(t)$ for all t, where $\bar{F}$ denotes the survival function.*
- convex order *(denoted by $X \leq_{cx} Y$) if, and only if, $E[\phi(X)] \leq E[\phi(Y)]$ for all convex function ϕ for which the expectations exist.*
- increasing convex order *(denoted by $X \leq_{icx} Y$) if, and only if, $E[\phi(X)] \leq E[\phi(Y)]$ for all increasing convex function ϕ for which the expectations exist.*
- concave *(denoted by $X \leq_{cv} Y$) if, and only if, $E[\phi(X)] \leq E[\phi(Y)]$ for all concave function ϕ for which the expectations exist.*
- concave *(denoted by $X \leq_{icv} Y$) if, and only if, $E[\phi(X)] \leq E[\phi(Y)]$ for all increasing concave function ϕ for which the expectations exist.*
- Laplace transform order *(denoted by $X \leq_{Lt} Y$) if, and only if, $E[e^{-aX}] \leq E[e^{-aY}]$, where a is any positive number.*

These stochastic orders have had a major impact on the areas of economy and insurance, where comparisons among expected utilities and risk measures are commonly considered. For example, Bäuerle and Müller [5] have produced leading studies regarding the problem of the preservation of risk measures with respect to usual stochastic order and convex order. These results are used to derive bounds for the risk measures of portfolios, i.e., of the joint financial position $X_1, \ldots, X_n$. This is of practical importance, since in many situations only the marginal distributions of $X_1, \ldots, X_n$ are known and not the dependence relation between them.

The usual stochastic order and the increasing concave order are also referred to as *First order Stochastic Dominance* (FSD) and *Second order Stochastic Dominance* (SSD), respectively.

In the multivariate setting, there also are many forms of comparing random vectors. We address the following stochastic orders that have been defined as a multivariate generalization of the usual stochastic order.

Definition 1.2.2 *Given two random vectors $\mathbf{X}$ and $\mathbf{Y}$, $\mathbf{X}$ is said to be smaller than $\mathbf{Y}$ in:*

- usual stochastic order *(denoted by $\mathbf{X} \leq_{st} \mathbf{Y}$) if $\mathrm{E}[\phi(\mathbf{X})] \leq \mathrm{E}[\phi(\mathbf{Y})]$ for any increasing function ϕ with finite expectations.*

- upper orthant order *(denoted by* $\mathbf{X} \leq_{uo} \mathbf{Y}$*) if* $\bar{F}_{\mathbf{X}}(\mathbf{t}) \leq \bar{F}_{\mathbf{Y}}(\mathbf{t})$ *for all* $\mathbf{t}$*.*
- lower orthant order *(denoted by* $\mathbf{X} \leq_{lo} \mathbf{Y}$*) if* $F_{\mathbf{X}}(\mathbf{t}) \geq F_{\mathbf{Y}}(\mathbf{t})$ *for all* $\mathbf{t}$*.*

In the upper orthant order and the lower orthant order $\bar{F}$ and F denote the survival and distribution function, respectively. Hence, an equivalent version of those multivariate orders can be defined by using the indicator function as follows

$$\left.\begin{array}{c} \mathbf{X} \leq_{uo} \mathbf{Y} \\ \\ \mathbf{X} \leq_{lo} \mathbf{Y} \end{array}\right| \equiv \left| \begin{array}{c} E\left[\mathbf{1}_{(\mathbf{t},\infty)}(\mathbf{x})\right] \leq E\left[\mathbf{1}_{(\mathbf{t},\infty)}(\mathbf{y})\right] \\ \\ E\left[\mathbf{1}_{(-\infty,\mathbf{t})}(\mathbf{x})\right] \geq E\left[\mathbf{1}_{(-\infty,\mathbf{t})}(\mathbf{y})\right] \end{array}\right. \quad \text{for all } \mathbf{t} \in \mathbb{R}^n$$

Observe that both orders are defined by ordering the probability of a set determined by any translation of the nonnegative and nonpositive orthant, respectively. In the univariate case they are equivalent since $\bar{F} = 1 - F$ but in the multivariate case this relationship does not hold.

Further details, properties and applications of all the stochastic orders defined above may be found, for example, in Shaked and Shanthikumar [55], Müller and Stoyan [46], or Denuit et al. [14].

Chapter 3 of this dissertation provides a generalization of those orthant orders. This is done by including translations and also rotations of the nonnegative orthant. The rotations of this orthant are determined by a unit vector and for two particular directions we recover the upper and lower orthant orders.

We also study stochastic comparison among the univariate random variables defined by the projection of a random vector on that unit vector that defines the rotations of the nonnegative orthant. In particular, we present some examples of application, in portfolio comparisons, in the determination of optimal allocations of wealth among risks in single period portfolio problems. We show that there are other directions which may be more interesting than those used to define the upper and lower orthant orders.

A brief review of the portfolio selection problem is introduced in the following Section.

1.3 The portfolio selection problem

In portfolio theory, the portfolio selection problem basically consists of finding the optimum way of investing in a given set of assets. The problem can be described as follows: consider an investor who has the possibility of investing in n different stocks. Investing one unit of money into stock i yields a random return X_i. Thus there is a vector of returns $\mathbf{X} = (X_1, \dots, X_n)'$. The investor has to allocate his/her budget, which without loss of generality can be 1 euro, to the different stocks in order to maximize an expected utility. Denote by $\mathbf{w} = (w_1, \dots, w_n)'$ the weights vector of the portfolio whose components represents the proportion of budget that the investor assigns to each stock X_i. If the investor has the utility function U then he/she faces the optimization problem

$$\max_{\mathbf{w}} EU\left(\sum_{i=1}^{n} w_i X_i\right) \quad \text{subject to} \quad \sum_{i=1}^{n} w_i = 1. \tag{1.3.1}$$

In most cases the utility function is a subjective function that measures the "happiness" of the investor according to increments in wealth. However, sometimes already partial knowledge of the utility function is sufficient to find the optimal allocation. For example, Hadar and Russel [24] proved that for independent and identically distributed random variables $X_1, \dots, X_n$ and a risk averse investor with a concave utility function, the solution to the problem (1.3.1) is given by the maximal diversification, i.e, $\frac{1}{n}$- rule that assigns the same proportion of budget to each stock. This result was generalized in Ma [42] in which the assumption of independence was replaced with the assumption of exchangeability. Related results also have been provided in Pellerey and Semeraro [48].

However, exchangeability still can lead to problems in the practical application of the results. Therefore, Chapter 3 contains a discussion regarding the treatment of the portfolio selection problem when assuming that the assets vector follows a multivariate elliptical distribution, a common assumption in the theory of portfolio selection. We show that the hypothesis of exchangeability can be verified through rotations of the distribution for the 2-dimensional case. The same result holds for higher dimensions under some constraints.

Consider now the case in which an investor only cares about the mean and variance of static portfolio returns. This case is the well-known Markowitz's model of mean-variance and the problem (1.3.1) becomes:

$$\min_{\mathbf{w}} \left[\mathbf{w}'\Sigma\mathbf{w} - \frac{1}{\gamma}\mu'\mathbf{w}\right] \quad \text{subject to} \quad \sum_{i=1}^{n} w_i = 1, \tag{1.3.2}$$

where Σ is the variance and covariance matrix of the stock vector $\mathbf{X}$, μ is its means vector and γ is a risk aversion parameter. The problem (1.3.2) is known as the mean-variance portfolio and when the risk aversion parameter tends to infinity we recover

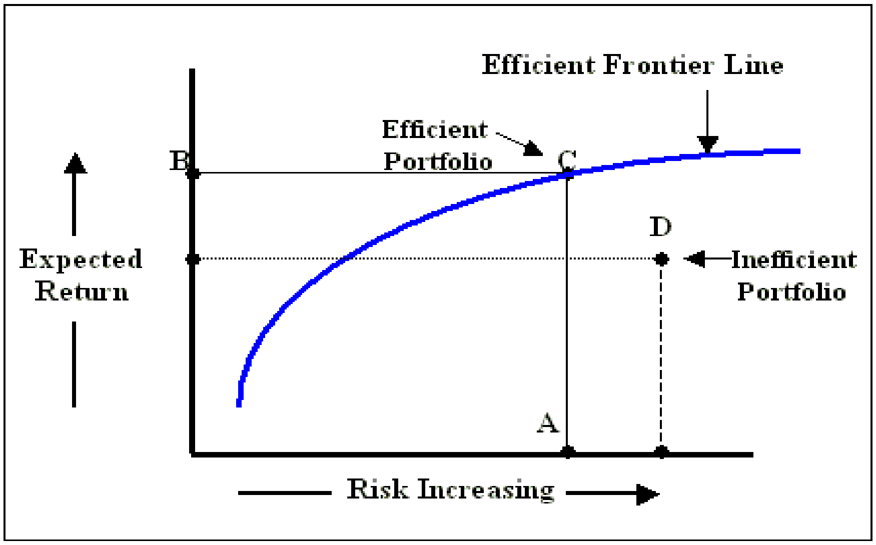


Figure 1.4: Efficient frontier

the minimum-variance portfolio. To implement these portfolios in practice, one has to estimate the mean and the covariance matrix of asset returns. Traditionally, the sample means and covariances have been used for this purpose.

The model proposed in Markowitz [43] has been relevant in modern portfolio theory where the main goal is to maximize return and minimize risk. Its philosophy is that the investors decide a portfolio weights $\mathbf{w}$ based on the trade-off between expected return and variance. Markowitz showed that an investor should hold a portfolio on the intersection of the set of portfolios with minimum risk and the set of portfolios with maximum return. That set is commonly called efficient frontier, an example of which is displayed in Figure 1.4[1].

Observe that the portfolio **D** is an inefficient portfolio since another portfolio can be found on the blue line which has the same risk and a better return. Following Markowitz's criteria, **C** is an efficient portfolio since it can not be improved in terms of the risk and returns at the same time. The set of all portfolios with the features of the portfolio **C** define the efficient frontier. The selection of the portfolio on this set depends on the risk aversion parameter since for different values of the risk aversion parameter γ we obtain the different mean-variance portfolios on the efficient frontier.

In this aspect, Chapter 4 of this thesis defines new versions of efficient frontiers by considering other criteria which are different from those of mean and variance used

[1]Figure taken from Hatton consulting, INC web page

in the Markowitz model. In fact, our approach also allows a definition of surface efficient when more than two criteria are considered by the investor. We have also designed a methodology in Chapter 4 based on order induced by extremality for selecting the best portfolio, when a direction chosen by the investor depends on the criteria that most interest him or her.

1.4 Structure of the thesis

This thesis contains five chapters. The current Chapter 1 presents a brief review of the classic concepts, emphasizing in the topics where the thesis develops its main contributions.

First, there is a discussion of various examples in the literature in which multivariate data orders have been introduced to extend some univariate statistical concepts to the multivariate setting. This is then followed by a brief review of some multivariate stochastic orderings so as to examine different ways of comparing random vectors based on their survival and distribution functions. Finally, there is a brief introduction to the portfolio selection problem.

The contributions of this thesis are developed in Chapters 2, 3 and 4. The first part of Chapter 2 introduces multivariate extremality as a methodology based on directions which measures the farness of a point $\mathbf{x}$ with respect to the data cloud or the distribution function. The main properties, as well as asymptotic results, are discussed. A new multivariate data order is also introduced which, taking its inspiration from extremality, is based on rotations of the nonnegative orthant. We discuss the fact that the multivariate quantiles introduced in Tibiletti [58] and Fernández-Ponce and Suárez-Llorens [19] can be generalized by extremality through its level set. In the second part of this Chapter, we develop an application of extremality in finance and, we introduce a new version of the multivariate Value at Risk. It will be observed that this last generalizes versions introduced in Tibiletti [59] and Embrechts and Puccetti [17] by considering other directions which might be of more interest to the risk manager.

Chapter 3, firstly defines a new multivariate stochastic order, called *extremality order*. This stochastic order is introduced as a generalization of the upper and lower orthant order discussed in Shaked and Shanthikumar [55] and Marshall and Olkin [44]. We provide a new way to compare random vectors in different directions. The closure and characterization properties also are discussed as well as the sufficient and necessary conditions for two random vectors to be ordered according to the extremality stochastic order. In second place, we present some examples of applications, in portfolio comparisons, in the determining optimal allocations of wealth among risks in single period portfolio problems. We also show that other directions can be more interesting than those used to define the upper and lower orthant orders. The results of Chapter 3 are mainly based on the published paper of Laniado et al. [36].

In Chapter 4, we develop an alternative methodology for selecting the portfolio weights in a data set that represents returns of n assets for investing. Markowitz [43] defined the efficient frontier as the set of the feasible portfolios which cannot be improved in terms of risk and return, simultaneously. Following Markowitz's idea, the first part of this Chapter introduces new concepts of efficient frontier that depend on indexes which may be chosen by the investor and which may be different to the classics of variance and return in Markowitz's model. Feasible portfolios are built with Montecarlo simulations and the new efficient frontiers are estimated by using an extremality order previously introduced in Chapter 2. In the second part of Chapter 4, we use real data to test the method introduced for selecting the best portfolio. We also compare the strategies developed in this thesis with some more traditional approaches studied in the literature.

Finally, in Chapter 5, we present some general conclusions and summarize the main contributions of the thesis.

2

Multivariate extremes: a directional approach

2.1 Introduction

The analysis of multivariate extreme outcomes is becoming very relevant in different fields. Moreover, a multivariate order is a valuable tool to analyze data properties and to obtain direct analogues for multivariate data of univariate concepts based on order such as median, range, extremes, quantiles or order statistics. Generalization of these concepts to the multivariate case is not straightforward. Chaudhuri [10] and references therein have 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 [15].

Several extensions of usual orders from $\mathbb{R}$ to $\mathbb{R}^n$, such as the Pareto-dominance types and the componentwise order, have the drawback of not being total orders. For facilitating total comparisons in the multivariate case the antisymmetry property is waived and therefore, preorders are obtained instead of orders. A preorder can be defined through a function of interest $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ that compares data according to its f- value, i.e., $\mathbf{x} \leq \mathbf{y} \equiv f(\mathbf{x}) \leq f(\mathbf{y})$. Examples of this type of ordering are those based on comparing data by their norms or projections onto some vector $\mathbf{u}$ such as the order by average or weighted average (see Barnett [3]).

Another alternative for multivariate orders are those based on depth functions, which assign to each point in $\mathbb{R}^n$ a measure of centrality with respect to the data cloud or probability distribution. This depth function decreases from the center outward (see, e.g., Zuo and Serfling [64] and Liu et al. [41]) and thus it provides a multivariate order that allows to define multivariate versions of median, order statistics, multivariate spacing and tolerance regions (Li and Liu [39]). However, they have the drawback of not being computationally feasible in high dimensions, hence the multivariate order induced through depth functions is quite limited for more than three dimensions.

Another option is the majorization order (Marshall and Olkin [44]) that is based on the idea of homogeneity between the components of a vector in $\mathbb{R}^n$ and is used in economics to compare the distribution of wealth in different populations.

Some orders can be characterized by a Euclidean convex cone C; for instance, for $\mathbf{x},\ \mathbf{y} \in \mathbb{R}^n$ $\ \mathbf{x} \leq \mathbf{y} \equiv \mathbf{y} - \mathbf{x} \in C$. This is the case of the componentwise order, where $C = \mathbb{R}^n_+ \cup \{\mathbf{0}\}$ or $C = \mathbb{R}^n_+$. These are two of the most important convex cones: the *non-negative orthant* and the *positive orthant* which are a basic concepts in the theory of inequalities. It is customary to write $\mathbf{x} \leq \mathbf{y}$, if $\mathbf{y} - \mathbf{x}$ belongs to the non-negative orthant (see, e.g., Rockafellar [51], page 13).

We propose in this Chapter a new multivariate data order based on the idea of extremality. Given a probability distribution in $\mathbb{R}^n$, the extremality of $\mathbf{x} \in \mathbb{R}^n$ in the direction $\mathbf{u}$ is one minus the probability of an oriented convex cone with vertex in $\mathbf{x}$. An important step in considering directions for multivariate data analysis is due to Kong and Mizera [32], who adopted a definition of quantiles through projections on unit vector. More recently Hallin et al. [25] proposed a new multivariate quantile based on a directional version of traditional regression quantiles, which also are associated with a vector $\mathbf{u}$. 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, unlike Hallin et al. [25], where $\mathbf{u}$ is the direction of the "vertical" axis in the regression, in this Chapter $\mathbf{u}$ is "bisectrix" of the oriented cone. For example, if $\mathbf{u} = \frac{1}{\sqrt{n}} 1_n$ then $\mathcal{C}^{\mathbf{u}}_{\mathbf{x}} = \mathbf{x} + \mathbb{R}^n_+$.

As a consequence of this extremality definition, we propose an order for multivariate data that allows to establish the "farness" of $\mathbf{x}$ with respect to a data cloud or to a distribution. We want to emphasize that our proposal introduced in this Chapter is a multivariate data order that works well in terms of computational times in high dimensions and for large samples. Besides the extremality measure provides a statistical methodology for segmenting a multivariate data sample, because the set of $\mathbf{x}^* \in \mathbb{R}^n$ such that $\mathbb{P}\left(\mathcal{C}^{\mathbf{u}}_{\mathbf{x}^*}\right) = q$ can be interpreted as a multidimensional quantile in the same way as in Tibiletti [58] and Fernández-Ponce and Suárez-Llorens

[19], where only the direction $\mathbf{u} = \frac{1}{\sqrt{n}}[\pm 1, \ldots, \pm 1]'$ was considered. Therefore, our proposal of extremality is a starting point to study segmentation by considering alternative directions that can be more interesting from a financial point of view such as, e.g., principal components or the direction of the weights vector in a portfolio. An interesting application that we also discuss is the definition of a new version of multivariate Value at Risk.

We also establish the most relevant properties of this extremality measure and give the theoretical basis for its nonparametric estimation. We include in this work an application in finance using the new definition of an Oriented Multivariate Value at Risk whose level sets are built through the order extremality. From Oriented Multivariate Value at Risk introduced in this Chapter, we present a procedure to bound the univariate Value at Risk in reasonable computational times. One advantage of the Value at Risk introduced is that it is fast to compute and applicable to high-dimensional data.

This Chapter is is organized as follows. Section 2.2 introduces the definition and properties of the oriented orthant and how it is constructed. This definitions and construction are necessary to present the extremality in Section 2.3. The main properties and consistency results are discussed in Section 2.4. An Oriented multivariate VaR is proposed in Section 2.5. Finally, in Section 2.6 we summarize the main conclusions.

2.2 Preliminaries

In this Section we introduce the main tools that we will use throughout the Chapter. As the unit vectors play a special role in this Chapter, we start by defining some mathematical formulation related to that will be relevant in the results given in the next Sections.

Definition 2.2.1 (Factorization QR) *Let A be an $m \times n$ 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 [23], 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 (from now on, we will assume nonnegative elements in R). The next result establishes that, in the QR factorization of any unit vector, R will be the first element of the canonical basis in $\mathbb{R}^n$

Proposition 2.2.2 *Let* $\mathbf{u} = [u_1, \ldots, u_n]'$ *be a vector with Euclidean norm* $\|\mathbf{u}\| = 1$. *If* $\mathbf{u} = QR$, *then* $R = [1, 0, \ldots, 0]'$.

Proof. We have that

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

Therefore, R has to be $[1, 0, \ldots, 0]'$ according to Definition 2.2.1. $\square$

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

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

$R_2 = R_1 = [1, 0, \ldots, 0]'$, from Proposition 2.2.2.

Hence, $Q_2'\mathbf{u} = Q_1'\mathbf{e}$ and $Q_1 Q_2'\mathbf{u} = \mathbf{e}$. Thus,

$$\mathcal{R}_{\mathbf{u}} = Q_1 Q_2' \tag{2.2.1}$$

is an orthogonal matrix transforming $\mathbf{u}$ into a unit vector with identical components. This transformation will send each vector $\mathbf{x}$ to a new orthogonal coordinates system, where $\mathbf{u}$ has all its angles equal with respect to the new nonnegative axis coordinates, that is, $\mathcal{R}_{\mathbf{u}}\mathbf{u} = \mathbf{e}$. The transformation (2.2.1) is used to define an oriented orthant as follows.

Definition 2.2.3 *Given a unit vector* $\mathbf{u} \in \mathbb{R}^n$ *and* $\mathbf{x} \in \mathbb{R}^n$, *the oriented orthant with vertex* $\mathbf{x}$ *and direction* $\mathbf{u}$ *is a cone convex given by*

$$\mathcal{C}_{\mathbf{x}}^{\mathbf{u}} = \{\mathbf{z} \in \mathbb{R}^n \mid \mathcal{R}_{\mathbf{u}}(\mathbf{z} - \mathbf{x}) \geq \mathbf{0}\}, \tag{2.2.2}$$

where the inequality is componentwise and $\mathcal{R}_{\mathbf{u}}$ *is the orthogonal matrix transforming* $\mathbf{u}$ *into a unit vector with identical components.*

$\mathcal{C}_{\mathbf{x}}^{\mathbf{u}}$ is the convex cone obtained by rotating the nonnegative orthant according to $\mathbf{u}$ and moving the origin to $\mathbf{x}$. Moreover, it has a single *extreme point* in $\mathbf{x}$, the semi-line

$$l = \{\mathbf{z} \in \mathbb{R}^n \mid \mathbf{z} = \mathbf{x} + \lambda \mathbf{u}, \quad \lambda \geq 0\} \tag{2.2.3}$$

is totally contained in $\mathcal{C}_{\mathbf{x}}^{\mathbf{u}}$ and its angles with respect to the new nonnegative semi-axis coordinates are $\cos^{-1}\left(\frac{1}{\sqrt{n}}\right)$. Note that for $\mathbf{u} = \frac{1}{\sqrt{n}}[\pm 1 \cdots \pm 1]'$ and $\mathbf{x} = \mathbf{0}$, $\mathcal{C}_{\mathbf{x}}^{\mathbf{u}}$ coincides with the 2^n orthants in $\mathbb{R}^n$. Specifically in $\mathbb{R}^2$, $\mathcal{C}_{\mathbf{x}}^{\mathbf{u}}$ is characterized by

$$\mathcal{C}_{\mathbf{x}}^{\mathbf{u}} = \left\{ \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \in \mathbb{R}^2 \; : \; \frac{\sqrt{2}}{2} \begin{pmatrix} u_1 + u_2 & u_2 - u_1 \\ u_1 - u_2 & u_1 + u_2 \end{pmatrix} \begin{pmatrix} z_1 - x_1 \\ z_2 - x_2 \end{pmatrix} \geq \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\},$$

where $\mathbf{u} = [u_1, u_2]'$ and $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$. Also,

$$\mathcal{C}_{\mathbf{x}}^{\mathbf{u}} = \left\{ \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \in \mathbb{R}^2 : \begin{pmatrix} \cos(\theta - \frac{\pi}{4}) & \sin(\theta - \frac{\pi}{4}) \\ -\sin(\theta - \frac{\pi}{4}) & \cos(\theta - \frac{\pi}{4}) \end{pmatrix} \begin{pmatrix} z_1 - x_1 \\ z_2 - x_2 \end{pmatrix} \geq \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\},$$

for $\mathbf{u} = [\cos\theta, \sin\theta]'$. Thus, $\mathcal{C}_{\mathbf{x}}^{\mathbf{u}}$ is a convex cone obtained by 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 (2.2.3) will be bisectrix of $\mathcal{C}_{\mathbf{x}}^{\mathbf{u}}$ with angles $\cos^{-1}\left(\frac{1}{\sqrt{2}}\right)$ with respect to the rotated nonnegative semi-axis.

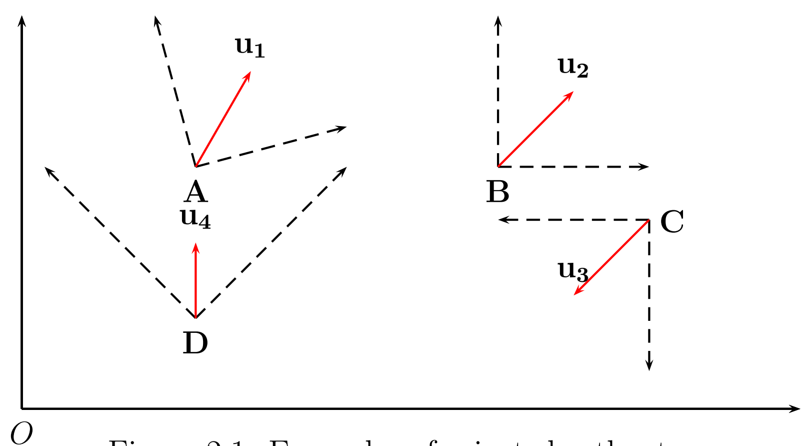


Figure 2.1: Examples of oriented orthant

Figure 2.1 shows some examples of *oriented orthants*. $\mathcal{C}_{\mathbf{A}}^{\mathbf{u_1}}, \mathcal{C}_{\mathbf{B}}^{\mathbf{u_2}}, \mathcal{C}_{\mathbf{C}}^{\mathbf{u_3}}$ and $\mathcal{C}_{\mathbf{D}}^{\mathbf{u_4}}$ with vertices in $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$ and $\mathbf{u_i} = [\cos\theta_i, \sin\theta_i]$, for $\theta_i = \frac{\pi}{3}, \frac{\pi}{4}, \frac{5\pi}{4}, \frac{\pi}{2}$, respectively. Hence,

$\mathcal{C}_{\mathbf{0}^+}^{\frac{\pi}{4}}$, $\mathcal{C}_{\mathbf{0}^+}^{\frac{3\pi}{4}}$, $\mathcal{C}_{\mathbf{0}^+}^{\frac{5\pi}{4}}$ and $\mathcal{C}_{\mathbf{0}^+}^{\frac{7\pi}{4}}$ can be seen as the $(+,+)$; $(-,+)$; $(-,-)$; $(+,-)$ quadrants, respectively.

From now on, let $\mathbf{e} = \frac{1}{\sqrt{n}}[1,\ldots,1]'$. If $\mathbf{u_1} = \mathbf{e}$, $\mathbf{u_2} = -\mathbf{e}$, then $\mathcal{C}_{\mathbf{0}}^{\mathbf{u_1}}$ and $\mathcal{C}_{\mathbf{0}}^{\mathbf{u_2}}$ are, respectively, the nonnegative and nonpositive orthants, since $\mathcal{R}_{\mathbf{u_1}} = I_n$ and $\mathcal{R}_{\mathbf{u_2}} = -I_n$ (see equation (2.2.1))

Proposition 2.2.4 *For any* $\mathbf{u}$, *if* $\mathbf{x} \in \mathcal{C}_{\mathbf{y}}^{\mathbf{u}}$, *then* $\mathcal{C}_{\mathbf{x}}^{\mathbf{u}} \subset \mathcal{C}_{\mathbf{y}}^{\mathbf{u}}$.

Proof. Suppose that $\mathbf{z} \in \mathcal{C}_{\mathbf{x}}^{\mathbf{u}}$. Definition 2.2.3 implies that $\mathcal{R}_{\mathbf{u}}(\mathbf{z}-\mathbf{x}) \geq \mathbf{0}$, and $\mathcal{R}_{\mathbf{u}}(\mathbf{x}-\mathbf{y}) \geq \mathbf{0}$ by hypothesis. Then as $\mathcal{R}_{\mathbf{u}}(\mathbf{z}-\mathbf{y}) = \mathcal{R}_{\mathbf{u}}(\mathbf{z}-\mathbf{x}) + \mathcal{R}_{\mathbf{u}}(\mathbf{x}-\mathbf{y}) \geq \mathbf{0}$, $\mathbf{z} \in \mathcal{C}_{\mathbf{y}}^{\mathbf{u}}$. □

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

Proposition 2.2.5 *If* $\mathbf{x} \neq \mathbf{y} \in \mathbb{R}^n$ *and* $\mathbf{u} = \frac{(\mathbf{x}-\mathbf{y})}{\|\mathbf{x}-\mathbf{y}\|}$, *where* $\|\cdot\|$ *is the Euclidean norm, then*

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

ii) $\mathcal{C}_{\mathbf{x}}^{\mathbf{u}} \subset \mathcal{C}_{\mathbf{y}}^{\mathbf{u}}$.

Proof. *i*) According to transformation (2.2.1), for any unit vector $\mathbf{u}$, $\mathcal{R}_{\mathbf{u}}\mathbf{u} = \mathbf{e}$. In particular, for $\mathbf{u} = \frac{(\mathbf{x}-\mathbf{y})}{\|\mathbf{x}-\mathbf{y}\|}$, clearly we see that $\mathcal{R}_{\mathbf{u}}(\mathbf{x}-\mathbf{y}) \geq \mathbf{0}$, therefore $\mathcal{R}_{\mathbf{u}}\mathbf{y} \leq \mathcal{R}_{\mathbf{u}}\mathbf{x}$. *ii*) Since for any unit vector $\mathbf{u}$, $\mathcal{R}_{\mathbf{u}}\mathbf{u} = \mathbf{e}$, then $\mathcal{R}_{\mathbf{u}}(\mathbf{x}-\mathbf{y}) \geq \mathbf{0}$, which means $\mathbf{x} \in \mathcal{C}_{\mathbf{y}}^{\mathbf{u}}$ and so, from Proposition 2.2.4, $\mathcal{C}_{\mathbf{x}}^{\mathbf{u}} \subset \mathcal{C}_{\mathbf{y}}^{\mathbf{u}}$. □

2.3 A directional extremality

Let $\mathbf{X}$ be a random vector with associated probability distribution P, cumulative distribution function F and joint density function f. Given a unit vector $\mathbf{u}$, let $P_{\mathbf{x},\mathbf{u}} = P_F\left(\mathcal{C}_{\mathbf{x}}^{\mathbf{u}}\right)$ be the probability that $\mathbf{X}$ belongs to $\mathcal{C}_{\mathbf{x}}^{\mathbf{u}}$. Note that $P_F\left(\mathcal{C}_{\mathbf{x}}^{-\mathbf{e}}\right) = F(\mathbf{x})$. With respect to calculation of $P_{\mathbf{x},\mathbf{u}}$, we can show now that $P_{\mathbf{x},\mathbf{u}}$ can be explicitly calculated in dimension two as follows.

$$
\begin{aligned}
P\left(\mathcal{C}_{\mathbf{x}}^{\mathbf{u}}\right) = & \left(\int_{x_1}^{\infty}\int_{\frac{-x_1\sin(\theta-\frac{\pi}{4})+x_2\cos(\theta-\frac{\pi}{4})+t\sin(\theta-\frac{\pi}{4})}{\cos(\theta-\frac{\pi}{4})}}^{\frac{x_1\cos(\theta-\frac{\pi}{4})+x_2\sin(\theta-\frac{\pi}{4})-t\cos(\theta-\frac{\pi}{4})}{\sin(\theta-\frac{\pi}{4})}} f(t,s)dsdt\right)\mathbf{1}_{\left\{\theta\in\left[0,\frac{\pi}{4}\right)\cup\left(\frac{7\pi}{4},2\pi\right)\right\}} \\
& + \left(\int_{x_2}^{\infty}\int_{\frac{x_1\cos(\theta-\frac{\pi}{4})+x_2\sin(\theta-\frac{\pi}{4})-s\sin(\theta-\frac{\pi}{4})}{\cos(\theta-\frac{\pi}{4})}}^{\frac{x_1\sin(\theta-\frac{\pi}{4})-x_2\cos(\theta-\frac{\pi}{4})+s\cos(\theta-\frac{\pi}{4})}{\sin(\theta-\frac{\pi}{4})}} f(t,s)dtds\right)\mathbf{1}_{\left\{\theta\in\left(\frac{\pi}{4},\frac{3\pi}{4}\right)\right\}} \\
& + \left(\int_{x_1}^{\infty}\int_{\frac{x_1\cos(\theta-\frac{\pi}{4})+x_2\sin(\theta-\frac{\pi}{4})-t\cos(\theta-\frac{\pi}{4})}{\sin(\theta-\frac{\pi}{4})}}^{\frac{-x_1\sin(\theta-\frac{\pi}{4})+x_2\cos(\theta-\frac{\pi}{4})+t\sin(\theta-\frac{\pi}{4})}{\cos(\theta-\frac{\pi}{4})}} f(t,s)dsdt\right)\mathbf{1}_{\left\{\theta\in\left(\frac{3\pi}{4},\frac{5\pi}{4}\right)\right\}} \\
& + \left(\int_{-\infty}^{x_2}\int_{\frac{x_1\sin(\theta-\frac{\pi}{4})-x_2\cos(\theta-\frac{\pi}{4})+s\cos(\theta-\frac{\pi}{4})}{\sin(\theta-\frac{\pi}{4})}}^{\frac{x_1\cos(\theta-\frac{\pi}{4})+x_2\sin(\theta-\frac{\pi}{4})-s\sin(\theta-\frac{\pi}{4})}{\cos(\theta-\frac{\pi}{4})}} (t,s)dtds\right)\mathbf{1}_{\left\{\theta\in\left(\frac{5\pi}{4},\frac{7\pi}{4}\right)\right\}} \\
& + \left(\int_{x_1}^{\infty}\int_{x_2}^{\infty} f(t,s)dsdt\right)\mathbf{1}_{\left\{\theta=\frac{\pi}{4}\right\}} + \left(\int_{-\infty}^{x_1}\int_{x_2}^{\infty} f(t,s)dsdt\right)\mathbf{1}_{\left\{\theta=\frac{3\pi}{4}\right\}} \\
& + \left(\int_{-\infty}^{x_1}\int_{-\infty}^{x_2} f(t,s)dsdt\right)\mathbf{1}_{\left\{\theta=\frac{5\pi}{4}\right\}} + \left(\int_{x_1}^{\infty}\int_{-\infty}^{x_2} f(t,s)dsdt\right)\mathbf{1}_{\left\{\theta=\frac{7\pi}{4}\right\}}.
\end{aligned}
$$

However, in higher dimensions is more difficult to give a general expression for $P\left(\mathcal{C}_{\mathbf{x}}^{\mathbf{u}}\right)$ unless the unit vector $\mathbf{u}$ is given numerically. It can be show that $P_{\mathbf{x},\mathbf{u}}$ is related to the transformation $\mathcal{R}_{\mathbf{u}}$, that is, if $D_{\mathbf{x}} = \{\mathbf{t} \in \mathbb{R}^n \mid \mathbf{t} \geq \mathbf{x}\}$ and $\mathbf{t} = \mathcal{R}_{\mathbf{u}}\mathbf{x}$, then clearly $\mathbf{x} = \mathcal{R}'_{\mathbf{u}}\mathbf{t}$ and

$$
P_{\mathbf{x},\mathbf{u}} = \int_{D_{\mathbf{x}}} f(\mathcal{R}_{\mathbf{u}}^{-1}\mathbf{t})\mathbf{dt}. \tag{2.3.1}
$$

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

$$
\hat{P}_{\mathbf{x},\mathbf{u}} = \frac{1}{m}\sum_{j=1}^{m}\mathbf{1}_{\{\mathbf{x}_j \,\in\, \mathcal{C}_{\mathbf{x}}^{\mathbf{u}}\}} = \frac{1}{m}\sum_{j=1}^{m}\mathbf{1}_{\{\mathcal{R}_{\mathbf{u}}(\mathbf{x}_j-\mathbf{x})\geq\mathbf{0}\}}, \tag{2.3.2}
$$

$\hat{P}_{\mathbf{x},\mathbf{u}}$ denotes the proportion of points belonging to $\mathcal{C}_{\mathbf{x}}^{\mathbf{u}}$.

We now formulate the extremality notion, which is the starting point to define a new multivariate data order.

Definition 2.3.1 (Directional Extremality) *The extremality of* $\mathbf{x}$ *in the direction* $\mathbf{u}$ *with respect to a distribution function* F *is given by*

$$
\mathcal{E}_{\mathbf{u}}(\mathbf{x}, F) = P_F\left(\overline{\mathcal{C}_{\mathbf{x}}^{\mathbf{u}}}\right) = 1 - P_{\mathbf{x},\mathbf{u}}. \tag{2.3.3}
$$

The extremality $\mathcal{E}_{\mathbf{u}}(\mathbf{x}, \mathbb{X})$ of $\mathbf{x} \in \mathbb{R}^n$ in direction $\mathbf{u}$ with respect to the data

$$\mathbb{X} = \{\mathbf{x}_1, \ldots, \mathbf{x}_m\}$$

is defined by replacing $P_{\mathbf{x},\mathbf{u}}$ by $\hat{P}_{\mathbf{x},\mathbf{u}}$, that is the proportion of points belonging to $\mathcal{C}_{\mathbf{x}}^{\mathbf{u}}$.

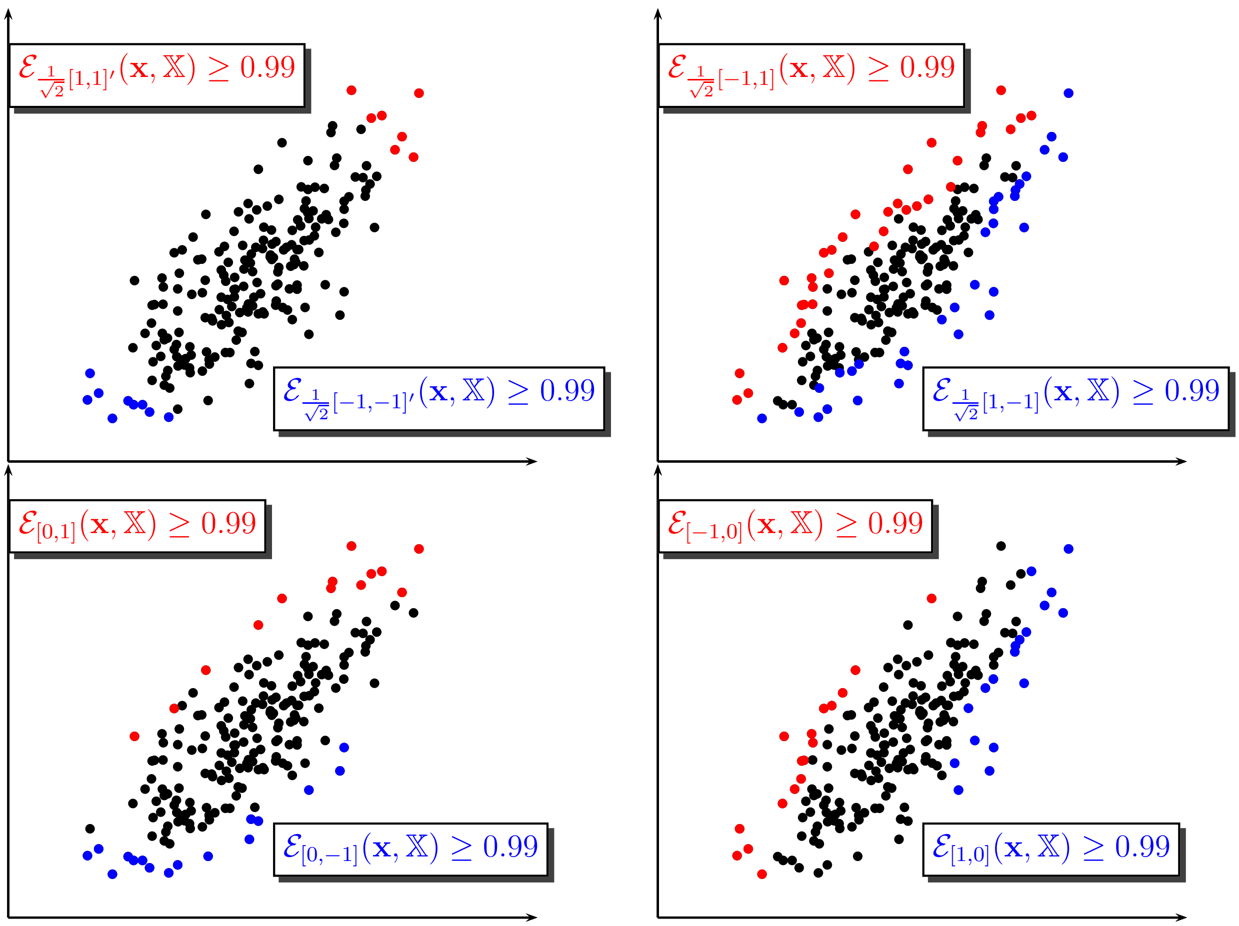


Figure 2.2: Points with high extremality

In Figure 2.2 it is displayed points with a sample extremality greater than 0.99 for different directions in a sample coming from a normal distribution with high positive correlation. From Figure 2.2 can be observed that high extremality of a point $\mathbf{x}$ means that the convex cone $\mathcal{C}_{\mathbf{x}}^{\mathbf{u}}$ contains a small proportion of the points of the sample or of the total mass of probability and possibly $\mathbf{x}$ belongs to some tail of the distribution. Hence, high extremality can be interpreted as "farness" with respect to the data cloud.

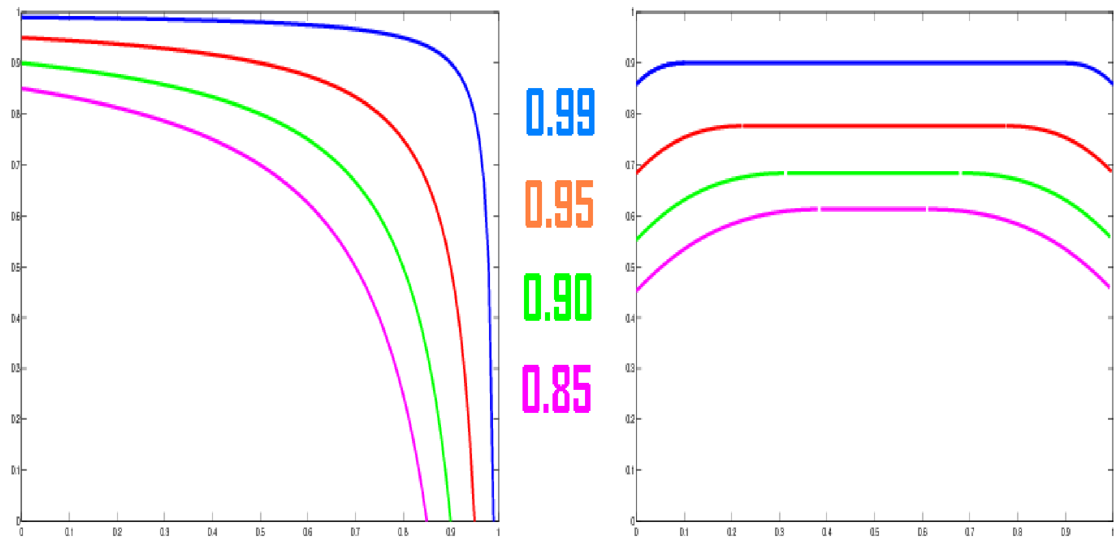


Figure 2.3: $\mathcal{E}_{\frac{1}{\sqrt{2}}[1,\ 1]'}(\mathbf{x}, F) = \alpha$ and $\mathcal{E}_{[0,\ 1]'}(\mathbf{x}, F) = \alpha$

Observe that those points labeled with high extremality can also be considered as extremes since an oriented orthant with vertex on any of them contains less of 1% of data. Therefore, the frontier-curve of those points may be seen as a new versions of multidimensional quantiles. We illustrate this idea in Figure 2.3 where extremality level curves corresponding to extremality values of 0.99; 0.95; 0.9 and 0.85 are displayed when F is a bivariate distribution with independent marginal distributions $U(0,\ 1)$. The left side for the direction $\mathbf{u} = \frac{1}{\sqrt{2}}[1,\ 1]'$ and the right one for the direction $[0,\ 1]'$, so that each election of a unit vector provide different way of segmenting the distribution or the multivariate sample.

Figure 2.4 provides extremality level surface corresponding to extremality values of 0.99; 0.95; 0.90; 0.85 in the direction $\mathbf{u} = \frac{1}{\sqrt{3}}[1,\ 1,\ 1]'$ of a multivariate distribution F with three independent marginal distributions $U(0,\ 1)$. Both Figure 2.3 and Figure 2.4 show a segmentation of the distribution and the curves or surfaces can be seen as quantile-curves or quantiles-surfaces. In contrast with the one-dimensional case where we only have two directions -1 or 1 originating the traditional α-quantile through $F^{-1}(\alpha)$ or $\bar{F}^{-1}(\alpha)$, but infinity directions can be used in higher dimensions. For particular cases of these directions we get the multidimensional quantiles studied in Tibiletti [58] and Belzunce et al. [4] where provide the classical orthants. Hence, our proposal is more flexible and allows to define new version of quantiles for a multivariate distribution in different directions. We will introduce in Section 2.5 other directions that can be useful in applications.

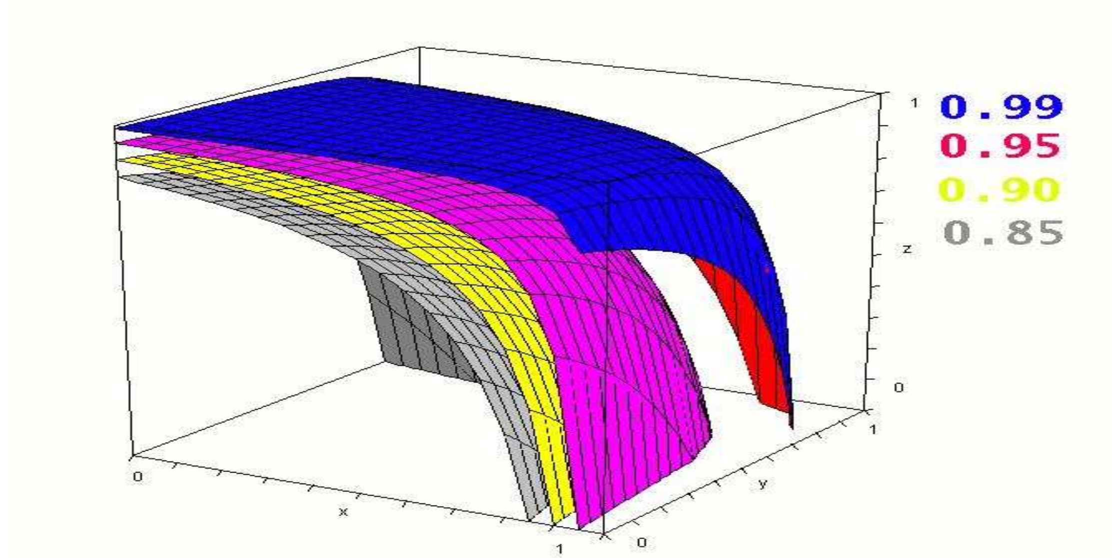


Figure 2.4: $\mathcal{E}_{\frac{1}{\sqrt{3}}[1,\,1,\,1]'}(\mathbf{x}, F) = \alpha$

The data order induced by $\mathbf{u}$ is based on comparing the extremality of the corresponding points as follows.

Definition 2.3.2 $\mathbf{y}$ *is said to be more extreme than* $\mathbf{x}$ *with respect to* F *in direction* $\mathbf{u}$*, denoted* $\mathbf{x} \leq_{\mathcal{E}_{\mathbf{u}}} \mathbf{y}$*, if* $\mathcal{E}_{\mathbf{u}}(\mathbf{x}, F) \leq \mathcal{E}_{\mathbf{u}}(\mathbf{y}, F)$.

For any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, any distribution function F and any direction $\mathbf{u}$, it holds that either $\mathbf{x} \leq_{\mathcal{E}_{\mathbf{u}}} \mathbf{y}$ or $\mathbf{y} \leq_{\mathcal{E}_{\mathbf{u}}} \mathbf{x}$. However, $\leq_{\mathcal{E}_{\mathbf{u}}}$ is not a partial order in $\mathbb{R}^n$, but a preorder; although it satisfies reflexivity and transitivity properties, it does not satisfy antisymmetry. The extremality order in the direction $\mathbf{u}$ allows to sort elements in $\mathbb{R}^n$ according to its extremality value. Elements with larger extremality value are extreme points and candidates to be outliers. Observe also that the set of $\mathbf{x} \in \mathbb{R}^2$ such that $\mathcal{E}_{\mathbf{u}}(\mathbf{x}, F) = q$, for each $q \in [0, 1]$, generalizes those quantiles curves introduced in Belzunce et al. [4] by inclusion of more directions. For greater dimensions, $\mathcal{E}_{\mathbf{u}}(\mathbf{x}, F) = q$, generalizes the multidimensional quantiles discussed in Tibiletti [58].

Given a direction, the extremality data order allows to segment a multivariate sample by using the ranks provided by the extremality measure.

Figure 2.5 shows a segmentation of a sample of a bivariate normal with means zero and covariance matrix

$$\Sigma = \begin{pmatrix} 1 & 0.7 \\ 0.7 & 2 \end{pmatrix}.$$

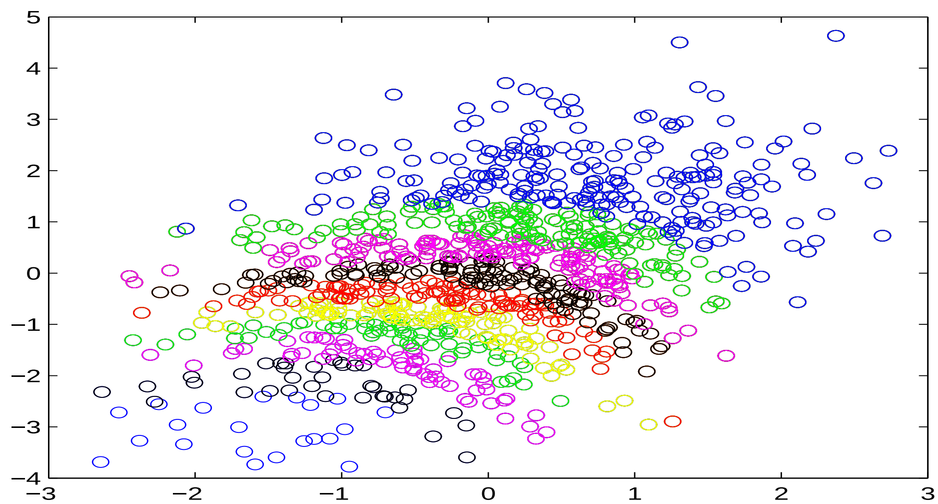


Figure 2.5: Segmentation

The partitions are given by the points $\mathbf{x}$ such that $\mathcal{E}_{\mathbf{u}}(\mathbf{x}, F) \geq 0.9, 0.8, \ldots, 0.1$, where $\mathbf{u}$ is the a unit vector in the direction of the first principal component.

2.4 Properties of the directional extremality

In this Section, we show some properties of the directional extremality defined in the previous Section which indicate us that our definition support intuitive and desirable properties associated to the extremality measure.

Property 2.4.1 *For any* $\mathbf{x}$ *and any absolutely continuous distribution function* F, $\mathcal{E}_{\mathbf{u}}(\mathbf{x}_0, F)$ *is continuous in* $\mathbf{u}$.

Proof. Let f be the density corresponding to F. Let $D_{\mathbf{x}} = \{\mathbf{t} \in \mathbb{R}^n \mid \mathbf{t} \geq \mathbf{x}\}$. Then from the inverse transformed Theorem and Definition 2.3.1, $\mathcal{E}_{\mathbf{u}}(\mathbf{x}_0, F)$ can be written as

$$\mathcal{E}_{\mathbf{u}}(\mathbf{x}_0, F) = 1 - \int_{D_{\mathbf{x}}} f(\mathcal{R}_{\mathbf{u}}^{-1}\mathbf{t})\mathbf{dt},$$

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

The following property shows that the vertex $\mathbf{x}$ has minimal extremality on the set $\mathcal{C}_{\mathbf{x}}^{\mathbf{u}}$.

Property 2.4.2 $\mathcal{E}_{\mathbf{u}}(\mathbf{x}, F) \leq \mathcal{E}_{\mathbf{u}}(\mathbf{y}, F)$, *for all* $\mathbf{y} \in \mathcal{C}_{\mathbf{x}}^{\mathbf{u}}$.

Proof. If $\mathbf{y} \in \mathcal{C}_{\mathbf{x}}^{\mathbf{u}}$ then $\mathcal{C}_{\mathbf{y}}^{\mathbf{u}} \subset \mathcal{C}_{\mathbf{x}}^{\mathbf{u}}$. Therefore,

$$P_F\left(\mathcal{C}_{\mathbf{y}}^{\mathbf{u}}\right) \leq P_F\left(\mathcal{C}_{\mathbf{x}}^{\mathbf{u}}\right) \quad \text{and} \quad \mathcal{E}_{\mathbf{u}}(\mathbf{x}, F) \leq \mathcal{E}_{\mathbf{u}}(\mathbf{y}, F).$$

□

Extremality also is invariant by orthogonal transformations as it is proven in the following result.

Property 2.4.3 *Let* $\mathbf{X}$ *be an* $n-$*dimensional random vector with distribution function* F*. Let* A *be an orthogonal matrix and let* $\mathbf{b} \in \mathbb{R}^n$*. Then*

$$\mathcal{E}_{A\mathbf{u}}(A\mathbf{x} + \mathbf{b}, F_{A\mathbf{X}+\mathbf{b}}) = \mathcal{E}_{\mathbf{u}}(\mathbf{x}, F_{\mathbf{X}}).$$

Proof. Since A is an orthogonal matrix and $\mathbf{u}, \mathbf{e}$ are unit vectors, $A\mathbf{u}$ is also a unit vector. Using Proposition 2.2.2, the QR factorization is given by

$$\mathbf{e} = Q_1 R_1, \ \mathbf{u} = Q_2 R_1, \ A\mathbf{u} = Q_3 R_1, \quad \text{where} \ \ R_1 = [1, 0, \ldots, 0]' \in \mathbb{R}^n. \tag{2.4.1}$$

Therefore, applying transformation (2.2.1), we have

$$\mathcal{R}_{\mathbf{u}} = Q_1 Q_2' \quad \text{and} \quad \mathcal{R}_{A\mathbf{u}} = Q_1 Q_3'. \tag{2.4.2}$$

Since R_1 is diagonal with non-negative entries, the QR factorization of $\mathbf{u}$ is unique. Therefore, from (2.4.1),

$$\mathbf{u} = Q_2 R_1 = A' Q_3 R_1, \ \ \text{which implies that} \ \ Q_2 = A' Q_3,$$

and, from (2.4.2),

$$\mathcal{R}_{\mathbf{u}} = Q_1 Q_2' = Q_1 Q_3' A = \mathcal{R}_{A\mathbf{u}} A. \tag{2.4.3}$$

Then, using (2.4.3) in the last equality, we obtain

$$\begin{aligned}
\mathcal{E}_{A\mathbf{u}}(A\mathbf{x} + \mathbf{b}, F_{A\mathbf{X}+\mathbf{b}}) &= 1 - P_{F_{A\mathbf{X}+\mathbf{b}}}\left(\mathcal{C}_{A\mathbf{x}+\mathbf{b}}^{A\mathbf{u}}\right) \\
&= 1 - P_F\left(\mathcal{R}_{A\mathbf{u}}(A\mathbf{X} + \mathbf{b} - A\mathbf{x} - \mathbf{b}) \geq \mathbf{0}\right) \\
&= 1 - P_F\left(\mathcal{R}_{A\mathbf{u}} A(\mathbf{X} - \mathbf{x}) \geq \mathbf{0}\right), \quad \text{(from (2.4.3))} \\
&= 1 - P_F\left(\mathcal{R}_{\mathbf{u}}(\mathbf{X} - \mathbf{x}) \geq \mathbf{0}\right) = 1 - P_F\left(\mathcal{C}_{\mathbf{x}}^{\mathbf{u}}\right) = \mathcal{E}_u(\mathbf{x}, F_{\mathbf{X}}).
\end{aligned}$$

□

Property 2.4.4 *Let* $\mathbf{x} \in \mathbb{R}^n - \{0\}$ *and* $\mathbf{u} = \frac{\mathbf{x}}{\|\mathbf{x}\|}$*, where* $\|\cdot\|$ *is the Euclidean norm. If* $\|\mathbf{x}\| \longrightarrow \infty$*, then* $\mathcal{E}_{\mathbf{u}}(\mathbf{x}, F) \longrightarrow 1$.

Proof. Let $B = \{\mathbf{b} \in \mathbb{R}^n : \|\mathbf{b}\| \geq \|\mathbf{x}\|\}$. Suppose that $\mathbf{z} \in \mathcal{C}_{\mathbf{x}}^{\mathbf{u}}$. Then $\mathcal{R}_{\mathbf{u}}(\mathbf{z}-\mathbf{x}) \geq \mathbf{0}$ (see equation (2.2.2)) and $\mathcal{R}_{\mathbf{u}}\mathbf{z} \geq \mathcal{R}_{\mathbf{u}}\mathbf{x}$. Using the transformation defined in (2.2.1),

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

and then,

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

Therefore, $\mathcal{C}_{\mathbf{x}}^{\mathbf{u}} \subset B$ since

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

It follows that

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

And the proof is complete letting $\|\mathbf{x}\| \longrightarrow \infty$. □

Let $\mathcal{E}_u(\mathbf{x}, F_m)$ as in (2.3.2) be the empirical version of $\mathcal{E}_u(\mathbf{x}, F)$. In Theorem 2.4.5 below, we show that $\mathcal{E}_u(\mathbf{x}, F_m)$ is a strongly consistent estimator of $\mathcal{E}_{\mathbf{u}}(\mathbf{x}, F)$ and we obtain its asymptotic distribution.

Theorem 2.4.5 *Let* $\mathbf{X}$ *be a random vector with distribution function* F. *Then,*

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

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

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

Proof. Let $\mathbf{X}_1, \ldots, \mathbf{X}_m$ be independent random vectors with a common distribution F. Let $\mathbf{1}_A(X)$ be the indicator function of A. Since $\mathbf{X}_1, \ldots, \mathbf{X}_m$ are *i.i.d* random vectors, $\mathbf{1}_{(\mathcal{C}_{\mathbf{x}}^{\mathbf{u}})}(\mathbf{X}_i)$, $i = 1, \ldots, m$, also are *i.i.d* random variables such that, for all $i = 1, \ldots, m$,

$$E\left[\mathbf{1}_{(\mathcal{C}_{\mathbf{x}}^{\mathbf{u}})}(\mathbf{X}_i)\right] = P_F\left(\mathcal{C}_{\mathbf{x}}^{\mathbf{u}}\right) \quad \text{and} \quad V\left[\mathbf{1}_{(\mathcal{C}_{\mathbf{x}}^{\mathbf{u}})}(\mathbf{X}_i)\right] = P_F(\mathcal{C}_{\mathbf{x}}^{\mathbf{u}})(1 - P_F(\mathcal{C}_{\mathbf{x}}^{\mathbf{u}})).$$

Let $\mathcal{C}$ be the class of all oriented orthants $\mathcal{C}_{\mathbf{x}}^{\mathbf{u}}$ with $\mathbf{x}, \mathbf{u} \in \mathbb{R}^n$ and define $P_{F_m}\left(\mathcal{C}_{\mathbf{x}}^{\mathbf{u}}\right)$ as

$$P_{F_m}\left(\mathcal{C}_{\mathbf{x}}^{\mathbf{u}}\right) = \frac{1}{m}\sum_{i=1}^{m}\mathbf{1}_{(\mathcal{C}_{\mathbf{x}}^{\mathbf{u}})}(\mathbf{X}_i).$$

i) From Strong Law of Large Numbers, (see, e.g., Gaenssler [22], page 2), we have that

$$P_{F_m}\left(\mathcal{C}_{\mathbf{x}}^{\mathbf{u}}\right) \longrightarrow P_F\left(\mathcal{C}_{\mathbf{x}}^{\mathbf{u}}\right) \quad a.s., \text{ as } m \longrightarrow \infty.$$

And thus,

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

ii) Using the Glivenko - Cantelli Theorem (see, e. g., Gaenssler[22] page 16), we have that

$$\sup_{\mathcal{C}_{\mathbf{x}}^{\mathbf{u}} \in \mathcal{C}} \left| P_{F_m}\left(\mathcal{C}_{\mathbf{x}}^{\mathbf{u}}\right) - P_F\left(\mathcal{C}_{\mathbf{x}}^{\mathbf{u}}\right) \right| \longrightarrow 0 \quad a.s.,$$

which implies that

$$\sup_{\mathbf{x}, \mathbf{u} \in \mathbb{R}^n} \left| \mathcal{E}_{\mathbf{u}}(\mathbf{x}, F_m) - \mathcal{E}_{\mathbf{u}}(\mathbf{x}, F_{\mathbf{X}}) \right| \longrightarrow 0 \quad a.s.$$

iii) Applying the Central Limit Theorem, it is easy to see

$$m^{\frac{1}{2}} \frac{P_{F_m}\left(\mathcal{C}_{\mathbf{x}}^{\mathbf{u}}\right) - P_F\left(\mathcal{C}_{\mathbf{x}}^{\mathbf{u}}\right)}{\sqrt{P_F\left(\mathcal{C}_{\mathbf{x}}^{\mathbf{u}}\right)\left(1 - P_F\left(\mathcal{C}_{\mathbf{x}}^{\mathbf{u}}\right)\right)}} \longrightarrow Z, \text{ as } m \longrightarrow \infty,$$

for each $\mathcal{C}_{\mathbf{x}}^{\mathbf{u}} \in \mathcal{C}$, where Z is a random variable with standard normal distribution. According to Definition 2.3.1 the previous expression can be rewritten as

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

and the result follows.

□

2.5 Financial application: a multivariate VaR

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

$$VaR_\alpha(X) \equiv \inf\{x \in \mathbb{R} \mid F(x) \geq \alpha\}. \tag{2.5.1}$$

In the multivariate context, it is usual to have a portfolio vector $\mathbf{X} = (X_1, \ldots, X_n)'$ for which it is necessary associate a risk measure. A natural idea is to consider a function $g : \mathbb{R}^n \longrightarrow \mathbb{R}$ and an one-dimensional risk measure on $g(\mathbf{X})$. Thus, the VaR of the joint portfolio is that associated to $g(\mathbf{X})$. Examples of this idea can be found in Burgert and Rüschendorf [6], where

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

The multivariate VaR analogue of univariate VaR is discussed in Embrechts and Puccetti [17] who define the VaR through the α-level sets of the joint loss distribution function and the joint loss tail function. Cascos and Molchanov [9] also define a multivariate VaR as level sets of halfspace trimming regions. Making use of the definition of the directional extremality, we introduce a multivariate VaR as level set of the extremality. If F is a multivariate distribution function, consider the sets

$$A_{\alpha}^{\mathbf{u}}(F) = \left\{ \mathbf{x} \in \mathbb{R}^n : \mathcal{E}_{\mathbf{u}}(\mathbf{x}, F) \geq 1 - \alpha \right\}, \text{ for all } 0 \leq \alpha \leq 1.$$

Its boundary $\partial A_{\alpha}^{\mathbf{u}}(F)$ can be interpreted as an *oriented multivariate value at risk* at the level α and we will denoted it by $VaR_{\alpha}^{\mathbf{u}}(\mathbf{X})$. In particular, for $\mathbf{u} = \mathbf{e}$ and $\mathbf{u} = -\mathbf{e}$, $VaR_{\alpha}^{\mathbf{u}}(\mathbf{X})$ are respectively the upper-orthant and the lower-orthant value at risk, discussed in Embrechts and Puccetti [17]. However, directions as $\mathbf{u} = \frac{1}{\sqrt{n]}}[\pm 1, \ldots, \pm 1]'$, $\mathbf{u} = \frac{\mathbf{w}}{\|\mathbf{w}\|}$ where $\mathbf{w}$ can be a vector of portfolio weights to invest in n assets, or even the direction of principal components can also be interesting in financial applications.

The $VaR_{\alpha}^{\mathbf{u}}(\mathbf{X})$ can be nonparametrically estimated by using a multivariate sample $\{\mathbf{x}_1, \ldots, \mathbf{x}_m\}$ and fitting a surface on the set $S_{\alpha}^{\mathbf{u}}(F_m) = \{\mathbf{x}_i : \mathcal{E}_{\mathbf{u}}(\mathbf{x}_i, F_m) = 1 - \alpha\}$. It may occur that $S_{\alpha}^{\mathbf{u}}(F_m) = \emptyset$ or that there were few elements satisfying the equality. To overcome this problem, we consider the set $S_{\alpha,h}^{\mathbf{u}}(F_m) = \{\mathbf{x}_i : |\mathcal{E}_{\mathbf{u}}(\mathbf{x}_i, F_m) - 1 + \alpha| \leq h\}$, with slack h. Since $S_{\alpha}^{\mathbf{u}}(F_m) \subset S_{\alpha,h}^{\mathbf{u}}(F_m)$, a more accurate estimation of the boundary can be made. The direction given by $\mathbf{u}$ may have influence in the estimation of $S_{\alpha,h}^{\mathbf{u}}(F_m)$. Indeed, the classical methods used to smooth functions may fail because the surface of interest is not a function in all cases, since the set $S_{\alpha,h}^{u}(F_m)$ can have couples $(\mathbf{x}, y_i), (\mathbf{x}, y_j)$ such that $\mathbf{x} \in \mathbb{R}^{n-1}$, $y_i \neq y_j$. Thus, we need to change the original coordinates to estimate the $VaR_{\alpha}^{\mathbf{u}}(\mathbf{X})$ as follows. Suppose that $S_{\alpha,h}^{\mathbf{u}}(F_m) = \{\mathbf{x}_1, \mathbf{x}_2 \ldots, \mathbf{x}_k\}$. Transforming the set according to (2.2.1), we get

$$\mathcal{R}_{\mathbf{u}} S_{\alpha,h}^{\mathbf{u}}(F_m) = \left\{ \mathcal{R}_{\mathbf{u}}\mathbf{x}_1, \mathcal{R}_{\mathbf{u}}\mathbf{x}_2 \ldots, \mathcal{R}_{\mathbf{u}}\mathbf{x}_k \right\}. \tag{2.5.2}$$

Now, the smoothing of the points in (2.5.2) is done by usual methods, and the resulting surface is transformed back to the original system to obtain finally an estimation for $VaR^{\mathbf{u}}_{\alpha}(\mathbf{X})$. The above is summarized in the following algorithm.

Input:
  $\mathbf{u}$, α, h, and the multivariate sample $\mathbb{X} = \mathbf{x}_1, \ldots, \mathbf{x}_m$
    for $i = 1$ to m
      $\mathcal{E}_i = \mathcal{E}_{\mathbf{u}}(\mathbf{x}_i, F_m)$
      if $|\mathcal{E}_i - 1 + \alpha| \leq h$
        $\mathbf{x}_i \in S^{\mathbf{u}}_{\alpha,h}(F_m)$
      end
    end
Fitting a function f on $\mathcal{R}_{\mathbf{u}} S^{\mathbf{u}}_{\alpha,h}(F_m)$
  $VaR^{\mathbf{u}}_{\alpha}(\mathbf{X}) = \mathcal{R}^{-1}_{\mathbf{u}} f$

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

We illustrate graphically this approach. Figure 2.6 shows the Theoretical VaR of level 0.05 in the direction $\mathbf{u} = \mathbf{e}$ for three bivariate distributions with independent marginals identically distributed as $U(0,1), N(0,1)$ and $Exp(1)$, respectively.

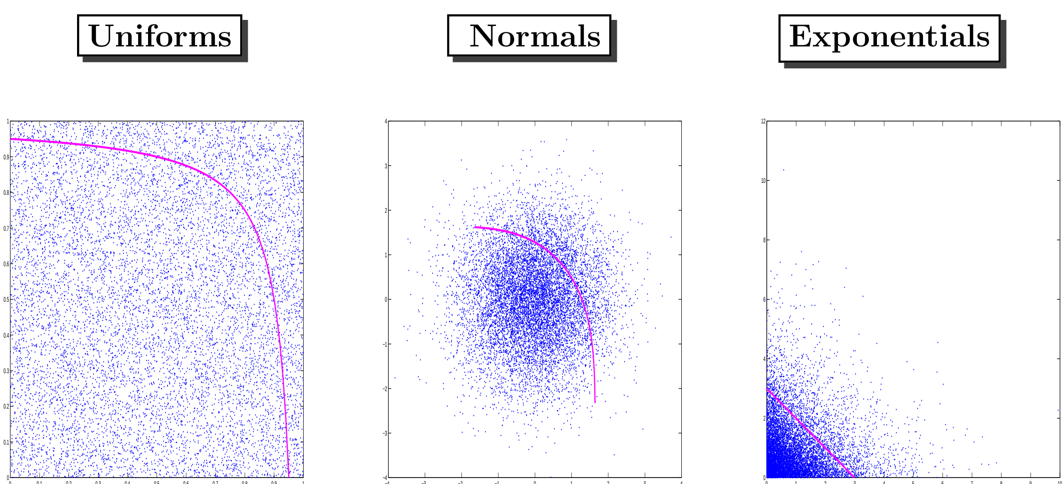


Figure 2.6: ■ Theoretical curve $VaR^{\mathbf{e}}_{0.05}(\mathbf{X})$

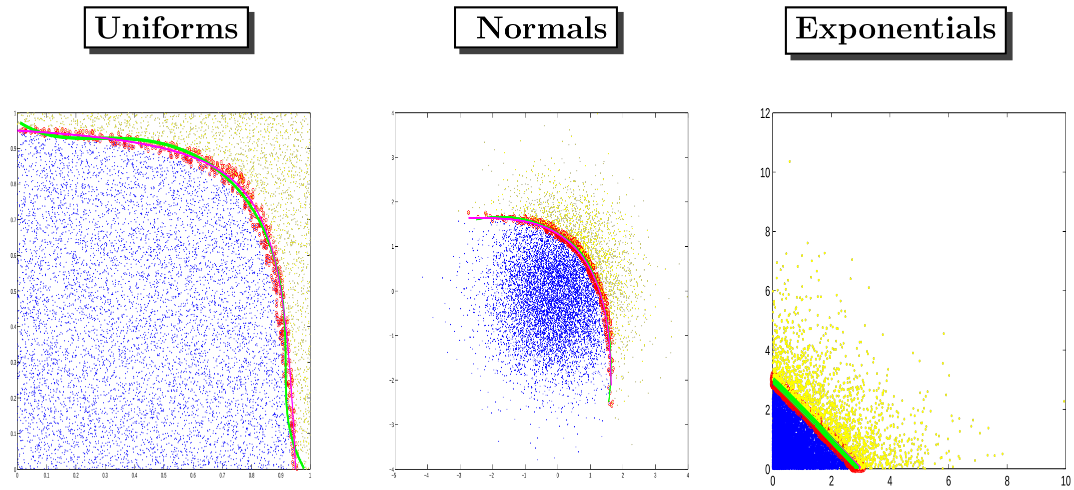


Figure 2.7: $\mathcal{E}_{\mathbf{e}}(\mathbf{x}_i, F_m) > 0.95$ $S^{\mathbf{u}}_{\alpha,h}(F_m)$ Estimated $VaR^{\mathbf{e}}_{0.05}$ $VaR^{\mathbf{e}}_{0.05}$

The estimation of the respective theoretical curves $VaR^{\mathbf{e}}_{0.05}$ is shown in Figure 2.7. We have considered $h = 0.01$. Here the VaR is calculated in the direction $\mathbf{u} = \mathbf{e}$ and $\mathcal{R}_{\mathbf{u}}$ is the identity matrix. Observe that an oriented orthant in direction $\mathbf{u} = \mathbf{e}$ and vertex at any yellow point contains a mass less than or equal to 0.05. Therefore the extremality of any of those points is greater than 0.95.

Figures 2.8 and 2.9 show daily negative returns of two leading companies, Microsoft and Google, from 19/08/2004 to 04/11/2010.

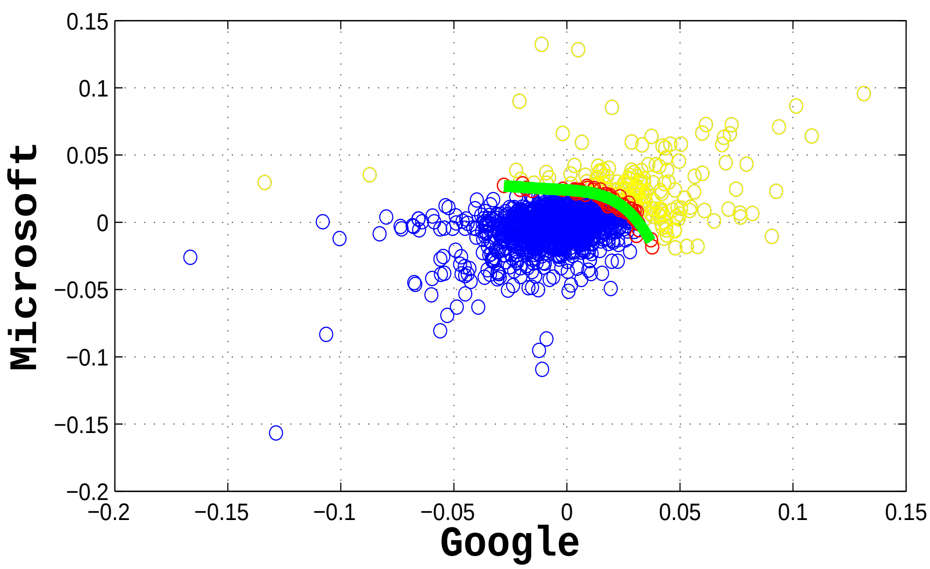


Figure 2.8: $\mathcal{E}_{\mathbf{e}}(\mathbf{x}_i, F_m) > 0.95$ Estimated $VaR^{\mathbf{e}}_{0.05}$

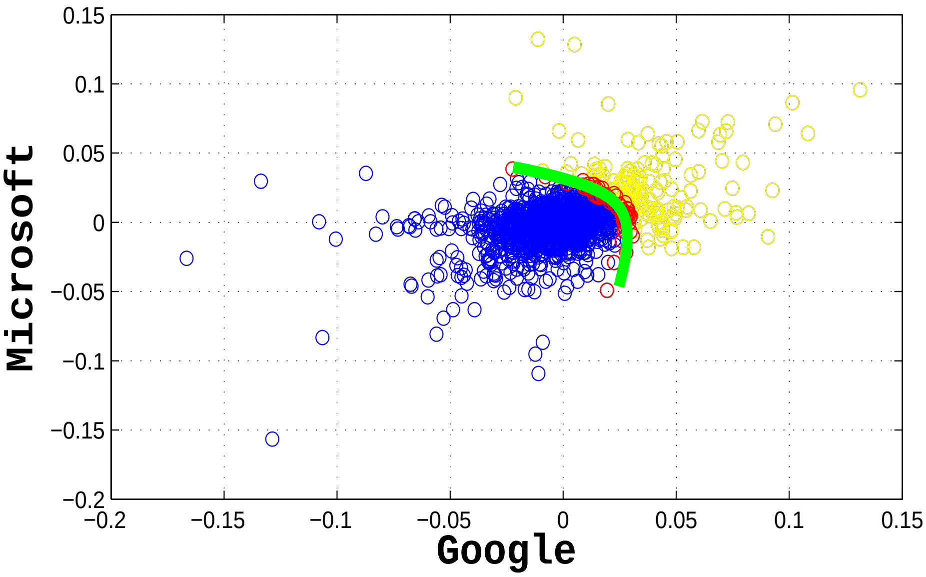


Figure 2.9: $\mathcal{E}_{\mathbf{pc}}\left(\mathbf{x}_i, F_m\right) > 0.95$ Estimated $VaR^{\mathbf{pc}}_{0.05}$

In Figure, 2.8 the VaR is estimated in the usual direction $\mathbf{u} = \mathbf{e}$; while in Figure 2.9 it is estimated using $\mathbf{u} = \mathbf{pc}$ where $\mathbf{pc}$ is the direction of maximum variability, that is, the direction given by the first principal component of the data. This $VaR^{\mathbf{pc}}_{\alpha}$ can be interpreted as a more conservative risk measure.

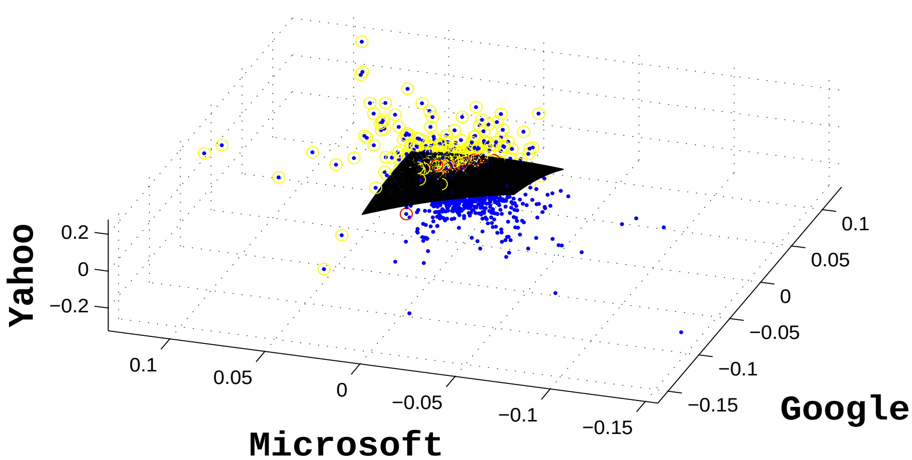


Figure 2.10: $\mathcal{E}_{\mathbf{e}}\left(\mathbf{x}_i, F_m\right) > 0.95$. Estimated $VaR^{\mathbf{e}}_{0.05}$.

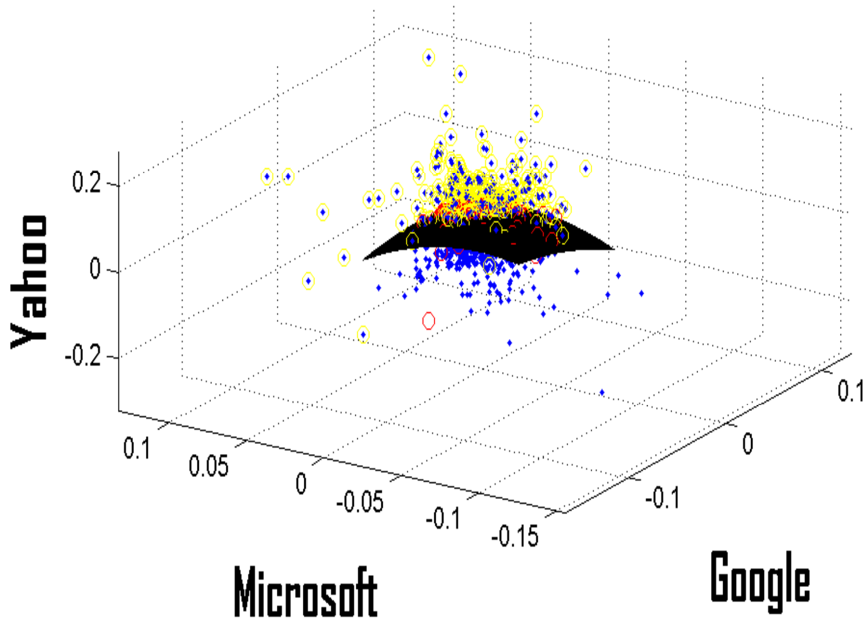


Figure 2.11: $\mathcal{E}_{\mathbf{pc}}(\mathbf{x}_i, F_m) > 0.95$. Estimated $VaR^{\mathbf{pc}}_{0.05}$.

Figures 2.10 and 2.11 show estimations of $VaR^{\mathbf{e}}_{0.05}$ and $VaR^{\mathbf{pc}}_{0.05}$, respectively, for daily negative returns of three leading companies, Microsoft, Google and Yahoo from 19/08/2004 to 04/11/2010. Note that the risk measure depends heavily on the selected direction. Observe that in all case the estimator obtained following our procedure seems to fit well the theoretical VaR.

An important aspect in financial risk is to obtain the risk measure in a reasonable time. An advantage of our approach is that can be easily computed for large values of n and sample size m. In order to show the computational times, we consider a sample of size m of $\mathbf{X} = (X_1, \ldots, X_n)'$ which follows a multivariate Normal distribution with parameters $\mu_i = 0$, $\sigma^2_{ii} = 1$ and $\sigma^2_{ij} = 0.8$. The following Table displays the elapsed time for calculating the points of the sample belonging to $VaR^{\mathbf{u}}_{\alpha}(\mathbf{X})$ for different values of the dimension n and of the sample size m. The times of Table 2.1 are independent of the direction $\mathbf{u}$.

Tables 2.1: Elapsed time for different dimensions and different sample size.

n	**2**	**5**	**10**	**20**	**50**
m	**Elapsed**	**time**	**in**	**seconds**	
1000	**5**	**6**	**6**	**6**	**8**
3000	**46**	**48**	**50**	**60**	**75**
5000	**128**	**135**	**138**	**143**	**205**
8000	**329**	**344**	**355**	**380**	**529**
10000	**509**	**533**	**551**	**643**	**830**

To our knowledge, the versions of the multivariate Value at Risk introduced in the literature have been from theoretical point of view. The computational aspect has not been treated in detail. Therefore, observing Table 2.1, we highlight that the main advantage of our proposal relative to other multivariate VaRs is that it is fast to compute and applicable to high dimensional data. We also want to emphasize that the computational cost to calculate the points que belongs to $VaR^{\mathbf{u}}_{\alpha}(\mathbf{X})$ for a sample n-dimensional of size m has a complexity $\mathcal{O}(m.n)$, where $n << m$.

In the next result, we give an interesting relationship between the Oriented Multivariate Value at Risk and the usual Univariate Value at Risk given in (2.5.1). Let $\mathbf{X} = (X_1, \ldots, X_n)'$ be the rate of returns of n different assets in a portfolio and suppose that $\mathbf{w} = (\omega_1, \ldots, \omega_n)'$ denote the portfolio weights. The total portfolio rate of return is given by $Z = \mathbf{w}'\mathbf{X}$, which is a sum of non-independent random variables.

The problem of obtaining the distribution function of Z has received considerable attention in the literature, since it has relevant applications in quantitative risk management. Embrechts and Puccetti [18] discuss bounds for the Value at Risk of Z, since one should be interested in bounding from above the probability that the aggregate loss amount will exceed some given threshold. Through the following proposition, we provide an interesting link between $VaR^{\mathbf{u}}_{\alpha}(\mathbf{X})$ and $VaR_{\alpha}(Z)$.

Proposition 2.5.1 *Let* $\mathbf{u} = \frac{\mathbf{w}}{\|\mathbf{w}\|}$ *be a unit vector in the direction of the portfolio weights vector. Consider* $\mathbf{X}$ *and* Z *as before. If* $\mathbf{x} \in VaR^{-\mathbf{u}}_{\alpha}(\mathbf{X})$, *then* $\mathbf{w}'\mathbf{x} \geq VaR_{\alpha}(Z)$.

Proof. Note that $\mathbf{x} \in VaR^{-\mathbf{u}}_{\alpha}(\mathbf{X})$ implies that $\mathcal{E}_{-\mathbf{u}}(\mathbf{x}, F) = 1 - \alpha$. By definition of extremality and from (3.3.3), we have that $P\left[\mathcal{R}_{-\mathbf{u}}(\mathbf{X} - \mathbf{x}) \geq 0\right] = \alpha$. Therefore,

$$P\left[\mathbf{1}'\mathcal{R}_{-\mathbf{u}}(\mathbf{X} - \mathbf{x}) \geq 0\right] \geq \alpha, \quad \text{where } \mathbf{1} = [1, \ldots, 1]'. \tag{2.5.3}$$

From (2.2.1) we can conclude that $\mathcal{R}_{-\mathbf{u}}(-\mathbf{u}) = \frac{1}{\sqrt{n}}\mathbf{1}$. Now, note that

$$\begin{aligned}
P\left[\mathbf{1}'\mathcal{R}_{-\mathbf{u}}(\mathbf{X}-\mathbf{x}) \geq 0\right] &= P\left[\sqrt{n}\left(\mathcal{R}_{-\mathbf{u}}(-\mathbf{u})\right)'\mathcal{R}_{-\mathbf{u}}(\mathbf{X}-\mathbf{x}) \geq 0\right] \\
&= P\left[\sqrt{n}(-\mathbf{u})'\mathcal{R}'_{-\mathbf{u}}\mathcal{R}_{-\mathbf{u}}(\mathbf{X}-\mathbf{x}) \geq 0\right] \\
&= P\left[\sqrt{n}(-\mathbf{u})'I_n(\mathbf{X}-\mathbf{x}) \geq 0\right] \\
&= P\left[\sqrt{n}\left(-\frac{\mathbf{w}}{\|\mathbf{w}\|}\right)'(\mathbf{X}-\mathbf{x}) \geq 0\right] \\
&= P\left[-\mathbf{w}'\mathbf{X} \geq -\mathbf{w}'\mathbf{x})\right] = P\left[Z \leq \mathbf{w}'\mathbf{x})\right].
\end{aligned}$$

Thus, (2.5.3) can be also written as

$$P\left[Z \leq \mathbf{w}'\mathbf{x})\right] \geq \alpha, \tag{2.5.4}$$

and by (2.5.1), $\mathbf{w}'\mathbf{x} \geq VaR_\alpha(Z)$. □

As consequence of Proposition 2.5.1, a bound easily computable in high dimension for the univariate $VaR_\alpha(Z)$ can be obtained by solving the following problem

$$\min \mathbf{w}'\mathbf{x} \quad s.t. \quad \mathbf{x} \in VaR_\alpha^{-\frac{\mathbf{w}}{\|\mathbf{w}\|}}(\mathbf{X}).$$

This result indicates how to consider a directional approach of the multivariate VaR can be useful and interesting in financial applications.

As an example of Proposition 2.5.1, we have simulated a sample of size 10000 of $\mathbf{X} = (X_1, \ldots, X_{10})'$ from a Normal distribution with parameters $\mu_i = 0$, $\sigma_{ii}^2 = 1$ and $\sigma_{ij}^2 = 0.8$ and we have obtained 2.31 as a bound for $VaR_{0.95}(Z)$, by using the direction $\mathbf{w} = \frac{1}{10}\mathbf{1}$. If we increase the dependence with $\sigma_{ij}^2 = 0.9$, the bound for $VaR_{0.95}(Z)$ was 2.11. We have also simulated comonotonic risks $\mathbf{X} = (X_1, \ldots, X_{20})'$ where $X_i = U^{\frac{i}{20}}$ being $U \sim$ Uniform(0,1) and the bound for $VaR_{0.95}(Z)$ was 0.9750. To establish a comparison with the real $VaR_\alpha(Z)$, we have estimated $VaR_\alpha(Z)$ through an empirical 0.95-quantile from the sample in the previous cases and the results have been: 1.5, 1.6, and 0.9750 respectively.

Calculation times in a simple laptop with a dual-core processor was approximately 535 seconds for each case. It is interesting to observe that our procedure is feasible in high dimension and if $\mathbf{X}$ presents strong dependence, the bound is more reliable. We highlight that after of having those points belong to $VaR_\alpha^{-\mathbf{u}}(\mathbf{X})$, the time to obtain the bound on the different values for the dimension n y for the sample size m is negligible. Hence, we emphasize that our procedure, as for obtaining $VaR_\alpha^{-\mathbf{u}}(\mathbf{X})$ as for finding the bound is quite feasible in high dimensions since the elapsed time depends weakly of the dimension (see Table 2.1)

2.6 Conclusions

In this Chapter we have introduced a new extremality notion in the multivariate context that induces a natural data order in $\mathbb{R}^n$ when a direction is chosen. As consequence of the extremality measure, it is possible to extend the idea of the multivariate quantiles studied in Tibiletti [58] and Fernández-Ponce and Suárez-Llorens [19] by considering other directions those that define the classical orthant.

We give a new version of the multivariate Value at Risk by including direction that generalizes those given in Embrechts and Puccetti [17] and Tibiletti [59]. This directional approach adds versatility to compare multivariate data, providing different versions of the multivariate Value at Risk.

We have studied an interesting link between the Oriented Multivariate Value at Risk introduced in this Chapter and the Classical univariate Value at Risk. In particular, in portfolio selection, we observe that the direction of the portfolio weights vector can be interesting in order to find a bound for the risk of the total value of the portfolio. From a computational point of view the bound is obtained in reasonable times even for portfolios with a large number of assets.

Various concepts introduced in this paper can be useful for future research line. As an example, the good of fitness data in $\mathbb{R}^2$. We note that if two random vectors have the same distribution function, they will have the same extremality for all $\mathbf{x}$ and any direction $\mathbf{u}$. Thus, we could determinate if two samples came from the same population testing a Extremality-Extremality-plot, in several direction.

Clearly the plots begins at $(0,0)$ and ends at $(1,1)$, and if the plot follows a straight line, then the random vectors should come from the same distribution. We illustrate the situation in Figure 2.12.

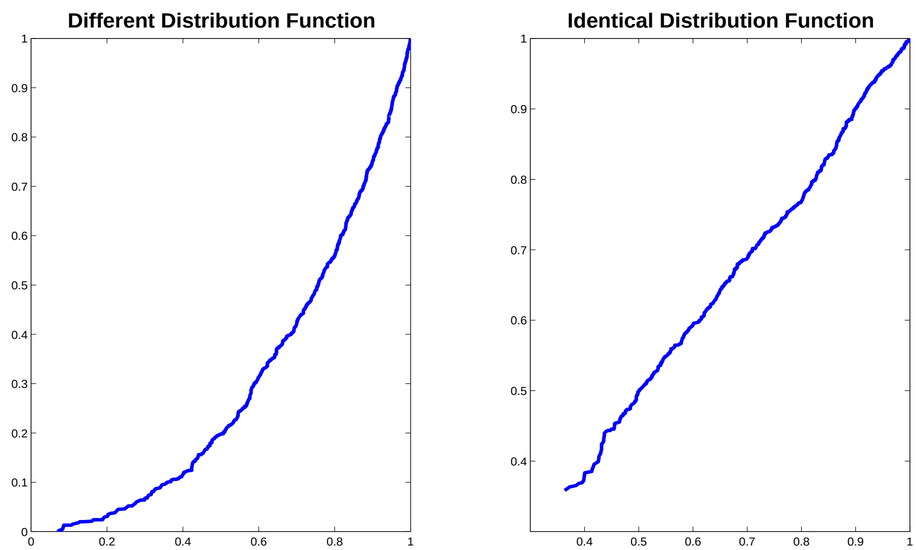


Figure 2.12: Extremality-Extremality-Plot

New Tolerance regions, detection of outliers and depth measure can be also examined with the concepts introduced in this paper. Work is currently underway on this extensions.

3

An extremality stochastic order

3.1 Introduction

One of the most relevant tools in risk evaluations of portfolios of hedge funds was introduced by Markowitz [43]. In his approach, risky investments comparisons are carried out through means and variances of the prospects: given a random vector of risky assets $\mathbf{X} = (X_1, \ldots, X_n)'$ and a real vector $\mathbf{w} = (\omega_1, \ldots, \omega_n)'$ describing the allocation of wealth, the risk averse decision maker assigns to the portfolio $Z_{\mathbf{X}} = \mathbf{w}'\mathbf{X}$ the utility $U(Z_{\mathbf{X}}) = E(Z_{\mathbf{X}}) - \alpha Var(Z_{\mathbf{X}})$, where $\alpha > 0$ is the degree of risk aversion, and choose among portfolios maximizing the utility $U(Z_{\mathbf{X}})$. Markowitz model has some drawbacks; for instance, it is not consistent with respect to the usual stochastic order (see Müller and Stoyan [46]), where the consistency is the monotonicity of a utility function or of a risk measure with respect to some stochastic order (see Bauerle and Müller [5] and references therein). In fact, starting from the assumption that utility functions are increasing and concave, which is common in economic theory, consistency means that stochastic comparisons between two different vectors $\mathbf{X}$ and $\mathbf{Y}$ of risky assets implies comparisons between the utilities $EU(Z_{\mathbf{X}})$ and $EU(Z_{\mathbf{Y}})$ for the same vector of allocations. The aim of this Chapter is to introduce a new multivariate stochastic order that may be useful in finding out new guidelines for allocation of risks in static portfolios.

Comparisons among random variables and vectors in different stochastic ways have been extensively considered during the last thirty years. Applications of these stochastic orderings have been provided in several disciplines, from economic theory to reliability and queueing theory (see, e.g., Barlow and Proschan [2], Stoyan [57], Shaked and Shanthikumar [55], Denuit et al. [14]). Among the stochastic orders defined and studied in the literature, most of them deal with comparisons between random vectors, like the multivariate usual stochastic order or the multivariate dispersion orders, with applications in decision making in multiple output scenarios.

In this Chapter we introduce a new multivariate stochastic order, called *extremality order*, that is a generalization of the upper and lower orthant order discussed in Shaked and Shanthikumar [55] and Marshall and Olkin [44], and that, unlike these two orders, allows comparisons of random vectors in different directions, determined by a unit vector. Essentially, it is based on rotation of the non-negative orthant in a direction given, obtaining a cone, which is isomorph to the non-negative orthant.

The extremality stochastic order introduced in this Chapter is based on the multivariate data ordering introduced in Chapter 2, where an extremality measure was defined to find multivariate extremes from a directional approach. Inspired on the extremality directional data ordering, we propose in this Chapter a probabilistic comparison between multivariate random vectors based on the probability assigned on some extremes sets on a given direction.

We present in portfolio comparisons some examples of application, in the determination of optimal allocations of wealth among risks in single period portfolio problems and we show that other directions can be more interesting than those used to define the upper and lower orthant orders.

The Chapter is organized as follows. Some preliminaries are described in Section 3.2 , while the properties of the extremality order, and its relationships with other multivariate stochastic orders (in particular, with the upper orthant and lower orthant orders), are described in Section 3.3. A list of its applications in portfolio theory are provided in Sections 3.4 and 3.5. Finally, in Section 3.6 we summarize the main conclusions.

3.2 Preliminaries

For ease of reference, first we briefly recall some notation that will be used throughout the Chapter. Random vectors taking values in $\mathbb{R}^n$ will be considered, unless otherwise stated. The space $\mathbb{R}^n$ is endowed with the usual componentwise partial order, which is defined as follows: given two vectors $\mathbf{x} = (x_1, \cdots, x_n)$ and $\mathbf{y} = (y_1, \cdots, y_n)$ in $\mathbb{R}^n$, then $\mathbf{x} \leq \mathbf{y}$ if $x_i \leq y_i$ for $i = 1, \cdots, n$. A function $\phi : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is said to be an increasing function when $\phi(\mathbf{x}) \leq \phi(\mathbf{y})$ for $\mathbf{x} \leq \mathbf{y}$. Throughout the paper the terms ‘increasing’ and ‘decreasing’ stand for ‘non-decreasing’ and ‘non-increasing’, respectively. Moreover, we shall adopt the following notations: for any random variable Z we shall denote its distribution function by $F_Z(x) = P(Z \leq x)$ and its survival function by $\bar{F}_Z(x) = P(Z > x)$; the notation $=_{st}$ stands for the equality in law; the notation $u \wedge v$ and $u \vee v$ stand for $\min\{u, v\}$ and $\max\{u, v\}$, respectively.

Concerning the stochastic comparisons, we first provide the definition of the orders usually considered in the univariate setting.

Definition 3.2.1 *Given two random variables X and Y we say that X is smaller than Y in the* usual stochastic order [convex order, increasing convex order, increasing concave order] *(denoted by $X \leq_{st} [\leq_{cx}, \leq_{icx}, \leq_{icv}]\ Y$) if and only if $E[\phi(X)] \leq E[\phi(Y)]$ for all increasing [convex, increasing convex, increasing concave] functions ϕ for which the expectations exist. Moreover, we say that X is smaller than Y in the* Laplace transform order *(denoted by $X \leq_{Lt} Y$) whenever $\phi(t) = -e^{-at}$ where a is any positive number.*

It should be recalled that in economics, where comparisons among expected utilities are commonly considered, the usual stochastic order and the increasing concave order are referred as *First order Stochastic Dominance* (FSD) and *Second order Stochastic Dominance* (SSD), respectively.

In the multivariate setting, the following stochastic orders have been defined as multivariate generalization of the usual stochastic order.

Definition 3.2.2 *Given two random vectors $\mathbf{X}$ and $\mathbf{Y}$, $\mathbf{X}$ is said to be smaller than $\mathbf{Y}$ in:*

(i) usual stochastic order *(denoted by $\mathbf{X} \leq_{st} \mathbf{Y}$) if $\mathrm{E}[\phi(\mathbf{X})] \leq \mathrm{E}[\phi(\mathbf{Y})]$ for any increasing function ϕ with finite expectations;*

(ii) upper orthant order *(denoted by $\mathbf{X} \leq_{uo} \mathbf{Y}$) if $\bar{F}_{\mathbf{X}}(x_1, \ldots, x_n) \leq \bar{F}_{\mathbf{Y}}(x_1, \ldots, x_n)$ for all $\mathbf{x}$, where $\bar{F}_{\mathbf{X}}, \bar{F}_{\mathbf{Y}}$ denote the survival function of $\mathbf{X}$ and $\mathbf{Y}$, respectively. (iii)* lower orthant order *(denoted by $\mathbf{X} \leq_{lo} \mathbf{Y}$) if $F_{\mathbf{X}}(x_1, \ldots, x_n) \geq F_{\mathbf{Y}}(x_1, \ldots, x_n)$ for all $\mathbf{x}$, where $F_{\mathbf{X}}, F_{\mathbf{Y}}$ denote the distribution function of $\mathbf{X}$ and $\mathbf{Y}$, respectively.*

It is easy to verify that both the upper orthant order and the lower orthant order are implied by the usual stochastic order. The following two multivariate stochastic orders have been defined to compare the strength of dependence between the components of vectors in the same Fréchet class (see Lehmann [38]). Recall that a function $\phi : \mathbb{R}^n \to \mathbb{R}$ is said to be supermodular if $\phi(\boldsymbol{x} \vee \boldsymbol{y}) + \phi(\boldsymbol{x} \wedge \boldsymbol{y}) \geq \phi(\boldsymbol{x}) + \phi(\boldsymbol{y})$ for all $\boldsymbol{x}$, $\boldsymbol{y} \in \mathbb{R}^n$.

Definition 3.2.3 *Given two random vectors $\mathbf{X}$ and $\mathbf{Y}$ having the same marginal distributions, i.e., such that $X_i =_{st} Y_i$ for all $i = 1, \ldots, n$, $\mathbf{X}$ is said to be smaller than $\mathbf{Y}$ in:*

(i) positive quadrant dependence order *(denoted by* $\mathbf{X} \leq_{PQD} \mathbf{Y}$*) if* $\mathbf{X} \leq_{uo} \mathbf{Y}$*, or, equivalently, if* $\mathbf{X} \geq_{lo} \mathbf{Y}$*;*

(ii) supermodular order *(denoted by* $\mathbf{X} \leq_{sm} \mathbf{Y}$*) if* $E[\phi(\mathbf{X})] \leq E[\phi(\mathbf{Y})]$ *for any supermodular function* ϕ *such that the expectations exist.*

It should be pointed out that the positive quadrant dependence order and the supermodular order are equivalent for dimension two, while this is not true in higher dimensions. Further details, properties and applications of all the stochastic orders defined above may be found, for example, in Müller and Stoyan [46], Shaked and Shanthikumar [55], or Denuit et al. [14].

3.3 Extremality order

The extremality order, defined here, is a generalization of the upper orthant and lower orthant orders, and allows for comparison of random vectors based on directions specified by a unit vector. Given $\mathbf{u} \in \mathbb{R}^n$ satisfying $\|\mathbf{u}\| = 1$, let $\mathcal{R}_{\mathbf{u}}$ be a rotation matrix such that

$$\mathcal{R}_{\mathbf{u}}\mathbf{u} = \frac{1}{\sqrt{n}}\,\mathbf{1}, \tag{3.3.1}$$

where $\mathbf{1} = [1, \ldots, 1]' \in \mathbb{R}^n$. We can now formulate the following definition to characterize the extremality order. From now on we assume that $\|\mathbf{u}\| = 1$.

Definition 3.3.1 *Given two random vectors* $\mathbf{X}$ *and* $\mathbf{Y}$ *in* $\mathbb{R}^n$*,* $\mathbf{X}$ *is said smaller than* $\mathbf{Y}$ *in* extremality order in the direction $\mathbf{u}$ *(denoted by* $\mathbf{X} \leq_{E_{\mathbf{u}}} \mathbf{Y}$*) if*

$$P\left[\mathcal{R}_{\mathbf{u}}\left(\mathbf{X} - \mathbf{t}\right) \geq 0\right] \leq P\left[\mathcal{R}_{\mathbf{u}}\left(\mathbf{Y} - \mathbf{t}\right) \geq 0\right], \quad \textit{for all } \mathbf{t} \textit{ in } \mathbb{R}^n. \tag{3.3.2}$$

In words, $\mathbf{X} \leq_{E_{\mathbf{u}}} \mathbf{Y}$ means that the probability that all components jointly assume “large values in the direction of $\mathbf{u}$” is lower for $\mathbf{X}$ than for $\mathbf{Y}$, where for “large values in the direction of a unit vector $\mathbf{u}$” we mean that $\mathbf{y}$ is larger than $\mathbf{x}$ if $\mathcal{R}_{\mathbf{u}}\left(\mathbf{y} - \mathbf{x}\right) \geq 0$. $\leq_{E_{\mathbf{u}}}$ is based on the multivariate data ordering introduced in the previous Chapter, where an extremality measure was defined to find multivariate extremes from a directional approach.

It is easy to observe that, for $\mathbf{u} = \frac{1}{\sqrt{n}}\,\mathbf{1}$, we have the natural componentwise order in $\mathbb{R}^n$. Therefore, if $\mathbf{u} = \frac{1}{\sqrt{n}}\,\mathbf{1}$, then

$$\mathbf{X} \leq_{E_{\mathbf{u}}} \mathbf{Y} \Longleftrightarrow \mathbf{X} \leq_{uo} \mathbf{Y} \quad \text{and} \quad \mathbf{X} \leq_{E_{-\mathbf{u}}} \mathbf{Y} \Longleftrightarrow \mathbf{X} \geq_{lo} \mathbf{Y}.$$

An equivalent definition of the order can be given by using the notion of oriented sub-orthants introduced in Chapter 2.

Definition 3.3.2 *Given a unit director vector* $\mathbf{u} \in \mathbb{R}^n$ *and a vertex* $\mathbf{t} \in \mathbb{R}^n$*, the* Oriented Sub-Orthant $\mathcal{C}_{\mathbf{t}}^{\mathbf{u}}$ *is the convex cone*

$$\mathcal{C}_{\mathbf{t}}^{\mathbf{u}} = \{\mathbf{x} \in \mathbb{R}^n \mid \mathcal{R}_{\mathbf{u}}(\mathbf{x} - \mathbf{t}) \geq \mathbf{0}\}. \tag{3.3.3}$$

Note that if $\mathbf{u} = \frac{1}{\sqrt{n}} \mathbf{1}$ then, $\mathcal{C}_{\mathbf{t}}^{\mathbf{u}} = \mathbf{t} + \mathbb{R}_+^n$, and therefore in this case the extremality order is equivalent to upper orthant order (and similarly for the lower orthant order).

According to Definition 2.2.3 another way of writing (3.3.2) is thus the following:

$$\mathbf{X} \leq_{E_{\mathbf{u}}} \mathbf{Y} \Longleftrightarrow P_{\mathbf{X}}\left(\mathcal{C}_{\mathbf{t}}^{\mathbf{u}}\right) \leq P_{\mathbf{Y}}\left(\mathcal{C}_{\mathbf{t}}^{\mathbf{u}}\right), \quad \text{for all } \mathbf{t} \in \mathbb{R}^n, \tag{3.3.4}$$

where $P_{\mathbf{X}}$ and $P_{\mathbf{Y}}$ are the probabilities induced by the joint distribution function of $\mathbf{X}$ and $\mathbf{Y}$, respectively. Note that (3.3.2) also is equivalent to

$$\mathbf{X} \leq_{E_{\mathbf{u}}} \mathbf{Y} \Longleftrightarrow E\left[I_{\mathcal{C}_{\mathbf{t}}^{\mathbf{u}}}(\mathbf{X})\right] \leq E\left[I_{\mathcal{C}_{\mathbf{t}}^{\mathbf{u}}}(\mathbf{Y})\right], \quad \text{for all } \mathbf{t} \in \mathbb{R}^n, \tag{3.3.5}$$

where $I_{\mathcal{C}_{\mathbf{t}}^{\mathbf{u}}}$ denotes the indicator function of $\mathcal{C}_{\mathbf{t}}^{\mathbf{u}}$. However, as shown in following Example 3.3.3, two random vectors can be ordered in extremality even if they are not comparable in the upper or in the lower orthant orders.

Example 3.3.3 *Consider two random vectors* $\mathbf{X} = (X_1, X_2)$ *and* $\mathbf{Y} = (Y_1, Y_2)$ *such that they are uniformly distributed and independent margins* $X_1 \sim U(0,c)$, $X_2 \sim U(a,b)$, $Y_1 \sim U(0,d)$, $Y_2 \sim U(0,b)$, *with* $b > a \geq 0$ $d > c \geq 0$. *Let* $F_{\mathbf{X}}$ *and* $F_{\mathbf{Y}}$ *be the joint distribution functions of* $\mathbf{X}$ *and* $\mathbf{Y}$*, respectively. We can see easily that,* $\bar{F}_{\mathbf{X}}(0,a) = 1$, $\bar{F}_{\mathbf{Y}}(0,a) = \frac{b-a}{b} < 1$, $\bar{F}_{\mathbf{X}}(c,a) = 0$ *and* $\bar{F}_{\mathbf{Y}}(c,a) = \frac{(d-c)(b-a)}{bd} > 0$. *Now* $F_{\mathbf{X}}(c,b) = 1$, $F_{\mathbf{Y}}(c,b) < 1$, $F_{\mathbf{X}}(c,a) = 0$ *and* $F_{\mathbf{Y}}(c,a) = \frac{ac}{bd} > 0$. *Therefore,* $\mathbf{X}$ *and* $\mathbf{Y}$ *are not ordered regarding the upper orthant order and are not ordered regarding the lower orthant order. However, taking* $\mathbf{u} = \frac{1}{\sqrt{2}}[1,-1]'$ *as unit vector, we have that*

$$P_{\mathbf{X}}\left(\mathcal{C}_{(x,y)}^{\mathbf{u}}\right) \begin{cases} 1 & \textit{if}, \quad x < 0, \ \ y > b, \\ \frac{y-b}{b-a} & \textit{if}, \quad x < 0, \ \ a \leq y \leq b, \\ \frac{c-x}{c} & \textit{if}, \quad 0 \leq x \leq c, \ \ y > b, \\ \left(\frac{c-x}{c}\right)\left(\frac{y-b}{b-a}\right) & \textit{if}, \quad 0 \leq x \leq c, \ \ a \leq y \leq b, \\ 0 & \textit{otherwise}. \end{cases}$$

$$P_{\mathbf{Y}}\left(\mathcal{C}_{(x,y)}^{\mathbf{u}}\right) \begin{cases} 1 & \textit{if}, \quad x < 0, \ \ y > b, \\ \frac{y}{b} & \textit{if}, \quad x < 0, \ \ 0 \leq y \leq b, \\ \frac{d-x}{d} & \textit{if}, \quad 0 \leq x \leq d, \ \ y > b, \\ \left(\frac{d-x}{d}\right)\left(\frac{y}{b}\right) & \textit{if}, \quad 0 \leq x \leq d, \ \ 0 \leq y \leq b, \\ 0 & \textit{otherwise}. \end{cases}$$

Assuming $(x, y) \notin (0, c)\times(a, b)$, is straightforward to see that $P_{\mathbf{X}}(\mathcal{C}^{\mathbf{u}}_{(x,y)}) \leq P_{\mathbf{Y}}(\mathcal{C}^{\mathbf{u}}_{(x,y)})$, while for $(x, y) \in (0, c) \times (a, b)$

$$P_{\mathbf{Y}}\left(\mathcal{C}^{\mathbf{u}}_{(x,y)}\right) - P_{\mathbf{X}}\left(\mathcal{C}^{\mathbf{u}}_{(x,y)}\right) = \left(\frac{d-x}{d}\right)\left(\frac{y}{b}\right) - \left(\frac{c-x}{c}\right)\left(\frac{y-b}{b-a}\right) \geq 0,$$

since that $\left(\frac{d-x}{d}\right) \geq \left(\frac{c-x}{c}\right)$ and $\left(\frac{y}{b}\right) \geq \left(\frac{y-b}{b-a}\right)$. Hence, $P_{\mathbf{X}}(\mathcal{C}^{\mathbf{u}}_{(x,y)}) \leq P_{\mathbf{Y}}(\mathcal{C}^{\mathbf{u}}_{(x,y)})$ for all $(x, y) \in \mathbb{R}^2$, and therefore, from (3.3.4), $\mathbf{X} \leq_{E_{\mathbf{u}}} \mathbf{Y}$.

Some properties of the extremality order are described next. The first one expresses a relation between the extremality order and the univariate stochastic order.

Property 3.3.4 *Let $\mathcal{R}^{(r.i)}_{\mathbf{u}}$ denotes the i-th row of the matrix $\mathcal{R}_{\mathbf{u}}$. If $\mathbf{X} \leq_{E_{\mathbf{u}}} \mathbf{Y}$, then $\mathcal{R}^{(r.i)}_{\mathbf{u}}\mathbf{X} \leq_{st} \mathcal{R}^{(r.i)}_{\mathbf{u}}\mathbf{Y}$ for every $i = 1, \ldots, n$.*

Proof. Since $\mathbf{X} \leq_{E_{\mathbf{u}}} \mathbf{Y} \Longleftrightarrow \mathcal{R}_{\mathbf{u}}\mathbf{X} \leq_{uo} \mathcal{R}_{\mathbf{u}}\mathbf{Y}$, the assertion immediately follows from Theorem 6.G.3-(c) in Shaked and Shanthikumar [55], that states that the margins of random vectors ordered in upper sense are ordered in the univariate usual stochastic order. □

An immediate consequence of Property 3.3.4 is that, whenever $\mathbf{X}$ and $\mathbf{Y}$ have finite means,

$$\mathbf{X} \leq_{E_{\mathbf{u}}} \mathbf{Y} \Longrightarrow \mathcal{R}_{\mathbf{u}}E[\mathbf{X}] \leq \mathcal{R}_{\mathbf{u}}E[\mathbf{Y}]. \tag{3.3.6}$$

Moreover, since for every vector $\mathbf{u}$ it is $\mathcal{R}'_{\mathbf{u}}\mathcal{R}_{\mathbf{u}} = I_n$, and, from (3.3.1), $[\sqrt{n}\mathcal{R}_{\mathbf{u}}\mathbf{u}]' = \mathbf{1}'$, it follows

$$\mathbf{1}'\mathcal{R}_{\mathbf{u}}\mathbf{X} = \left[\sqrt{n}\mathcal{R}_{\mathbf{u}}\mathbf{u}\right]'\mathcal{R}_{\mathbf{u}}\mathbf{X} = \sqrt{n}\mathbf{u}'\mathcal{R}'_{\mathbf{u}}\mathcal{R}_{\mathbf{u}}\mathbf{X} = \sqrt{n}\mathbf{u}'\mathbf{X}. \tag{3.3.7}$$

Thus, again from Property 3.3.4,

$$\mathbf{X} \leq_{E_{\mathbf{u}}} \mathbf{Y} \Longrightarrow \mathbf{u}'E[\mathbf{X}] \leq \mathbf{u}'E[\mathbf{Y}]. \tag{3.3.8}$$

One reason of interest in Property 3.3.4 is the fact that it provides a tool to compare linear combinations of random variables, a typical problem considered in portfolio theory. It is well known that, given two sets $\{X_1, \ldots, X_n\}$ and $\{Y_1, \ldots, Y_n\}$ of independent random variables, if $X_i \leq_{st} Y_i$ then $\sum_{i=1}^n a_iX_i \leq_{st} \sum_{i=1}^n a_iY_i$ whenever $a_i \geq 0$, $i = 1, \ldots, n$.

A generalization of this assertion was proved by Scarsini [53], who removed the assumption of independence, proving that if $(X_1, \ldots, X_n)'$ and $(Y_1, \ldots, Y_n)'$ have a common copula, then the stochastic order among the marginals implies the stochastic orders among the vectors and, as a corollary, the stochastic order among positive combinations of the marginals.

Since $\mathcal{R}_{\mathbf{u}}^{(r.i)}\mathbf{X}$ and $\mathcal{R}_{\mathbf{u}}^{(r.i)}\mathbf{Y}$ are linear combinations, not necessarily positive, of marginals of $\mathbf{X}$ and $\mathbf{Y}$, then Property 3.3.4 describes conditions to compare non-positive linear combinations of dependent random variables, as it is shown in the following example.

Example 3.3.5 *Let* $\mathbf{X} = (X_1, X_2)'$ *and* $\mathbf{Y} = (Y_1, Y_2)'$ *be two normally distributed vectors having the same covariance. Assume that* $Var(X_1) + Var(X_2) = Var(Y_1) + Var(Y_2)$*. Clearly,* $\mathbf{X}$ *and* $\mathbf{Y}$ *can have different copulas. Let* $\mathbf{u} = [1, 0]'$ *be the unit vector, so that the rotation matrix is given by*

$$\mathcal{R}_{\mathbf{u}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}.$$

Assume that $\mathcal{R}_{\mathbf{u}} \left(E[\mathbf{Y}] - E[\mathbf{X}]\right) \geq \mathbf{0}$*, i.e., that* $E[\mathbf{Y}]$ *belongs to oriented sub-orthant with vertex in* $E[\mathbf{X}]$ *and oriented by the vector* $\mathbf{u}$*. Under the assumptions above it is clear that* $Var\left(\mathcal{R}_{\mathbf{u}}^{(r.i)}\mathbf{X}\right) = Var\left(\mathcal{R}_{\mathbf{u}}^{(r.i)}\mathbf{Y}\right)$*,* $i = 1, 2$*. Besides,*

$$Cov\left(\mathcal{R}_{\mathbf{u}}^{(r.1)}\mathbf{X}, \mathcal{R}_{\mathbf{u}}^{(r.2)}\mathbf{X}\right) = \frac{1}{2}\left(Var(X_1) - Var(X_2)\right)$$

and

$$Cov\left(\mathcal{R}_{\mathbf{u}}^{(r.1)}\mathbf{Y}, \mathcal{R}_{\mathbf{u}}^{(r.2)}\mathbf{Y}\right) = \frac{1}{2}\left(Var(Y_1) - Var(Y_2)\right).$$

Therefore, from Theorem 3.3.21 in Müller and Stoyan [46] it follows that if

$$\left(Var(X_1) - Var(X_2)\right) \leq \left(Var(Y_1) - Var(Y_2)\right),$$

then $\mathbf{X} \leq_{E_{\mathbf{u}}} \mathbf{Y}$*. By using Property 3.3.4, we get both* $X_1 - X_2 \leq_{st} Y_1 - Y_2$ *and* $X_1 + X_2 \leq_{st} Y_1 + Y_2$*.*

The following property describes sufficient conditions to compare normal random vectors in the extremality order sense. Other sufficient conditions for the extremality comparison will be stated next.

Property 3.3.6 *Let* $\mathbf{X} \sim N\left(\mu_{\mathbf{X}},\ \Sigma_{\mathbf{Y}}\right)$ *and* $\mathbf{Y} \sim N\left(\mu_{\mathbf{Y}},\ \Sigma_{\mathbf{Y}}\right)$ *be two normally distributed random vectors, and let* $\mathbf{u}$ *be a unit vector such that* $\mathcal{R}_{\mathbf{u}}\left(\mu_{\mathbf{Y}} - \mu_{\mathbf{X}}\right) \geq 0$*.*

a) *If* $\Sigma_{\mathbf{X}} = \Sigma_{\mathbf{Y}} \Longrightarrow \mathbf{X} \leq_{E_{\mathbf{u}}} \mathbf{Y}$

b) *If* $\Sigma_{\mathbf{Y}} = \Sigma_{\mathbf{X}} + A$*, where* A *is a matrix such that* $\Sigma_{\mathbf{Y}}$ *is nonnegative definite and* $\mathcal{R}_{\mathbf{u}} A \mathcal{R}'_{\mathbf{u}}$ *has nonnegative components with zero diagonal elements, then* $\mathbf{X} \leq_{E_{\mathbf{u}}} \mathbf{Y}$*.*

Proof. Part a) follows from fact that $\mu_{\mathcal{R}_\mathbf{u}\mathbf{X}} \leq \mu_{\mathcal{R}_\mathbf{u}\mathbf{Y}}$ and $\Sigma_{\mathcal{R}_\mathbf{u}\mathbf{X}} = \Sigma_{\mathcal{R}_\mathbf{u}\mathbf{Y}}$. Therefore $\mathcal{R}_\mathbf{u}\mathbf{X} \leq_{st} \mathcal{R}_\mathbf{u}\mathbf{Y}$ by Theorem 3.3.13. in Müller and Stoyan [46], thus $\mathbf{X} \leq_{E_\mathbf{u}} \mathbf{Y}$. Part b) follows easily from Theorem 3.3.21.(a) in Müller and Stoyan [46] since $\mu_{\mathcal{R}_\mathbf{u}\mathbf{X}} \leq \mu_{\mathcal{R}_\mathbf{u}\mathbf{Y}}$ and $\Sigma_{\mathcal{R}_\mathbf{u}\mathbf{Y}} = \Sigma_{\mathcal{R}_\mathbf{u}\mathbf{X}} + \mathcal{R}_\mathbf{u} A \mathcal{R}'_\mathbf{u}$. □

For the next statement, let the *oriented upper set in the direction* $\mathbf{u}$ (denoted by $U_\mathbf{u}$) be a set such that $\mathbf{x} \in U_\mathbf{u}$ implies $\mathcal{C}_\mathbf{x}^\mathbf{u} \subset U_\mathbf{u}$.

Property 3.3.7 *Let* $\mathbf{X} = (X_1, \ldots, X_n)'$ *and* $\mathbf{Y} = (Y_1, \ldots, Y_n)'$ *be two random vectors. If* $E\left[I_{U_\mathbf{u}}(\mathbf{X})\right] \leq E\left[I_{U_\mathbf{u}}(\mathbf{Y})\right]$ *for all oriented upper set* $U_\mathbf{u}$ *in the direction* $\mathbf{u}$*, then* $\mathbf{X} \leq_{E_\mathbf{u}} \mathbf{Y}$*.*

Proof. Since the assumption is equivalent to $\mathcal{R}_\mathbf{u}\mathbf{X} \leq_{st} \mathcal{R}_\mathbf{u}\mathbf{Y}$, then $\mathcal{R}_\mathbf{u}\mathbf{X} \leq_{uo} \mathcal{R}_\mathbf{u}\mathbf{Y}$. Thus, $\mathbf{X} \leq_{E_\mathbf{u}} \mathbf{Y}$. □

Another sufficient condition for the extremality order, easily checked in practice, is stated in the next property. Recall that two random vectors $\mathbf{X}$ and $\mathbf{Y}$ with densities $f_\mathbf{X}$ and $f_\mathbf{Y}$, respectively, are such that $\mathbf{X}$ is smaller than $\mathbf{Y}$ in the *likelihood ratio order* (denoted $\mathbf{X} \leq_{lr} \mathbf{Y}$) if

$$f_\mathbf{X}(\mathbf{x}) f_\mathbf{Y}(\mathbf{y}) \leq f_\mathbf{X}(\mathbf{x} \vee \mathbf{y}) f_\mathbf{Y}(\mathbf{x} \wedge \mathbf{y})$$

for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Since the likelihood ratio order implies the usual stochastic order, the following property immediately follows from the chain of implications

$$\mathcal{R}_\mathbf{u}\mathbf{X} \leq_{lr} \mathcal{R}_\mathbf{u}\mathbf{Y} \Rightarrow \mathcal{R}_\mathbf{u}\mathbf{X} \leq_{st} \mathcal{R}_\mathbf{u}\mathbf{Y} \Rightarrow \mathcal{R}_\mathbf{u}\mathbf{X} \leq_{uo} \mathcal{R}_\mathbf{u}\mathbf{Y} \Rightarrow \mathbf{X} \leq_{E_\mathbf{u}} \mathbf{Y}.$$

Property 3.3.8 *Let* $\mathbf{X}$ *and* $\mathbf{Y}$ *be two random vectors with densities* $f_\mathbf{X}$ *and* $f_\mathbf{Y}$*, respectively. If*

$$f_\mathbf{X}(\mathcal{R}'_\mathbf{u}\mathbf{x}) f_\mathbf{Y}(\mathcal{R}'_\mathbf{u}\mathbf{y}) \leq f_\mathbf{X}(\mathcal{R}'_\mathbf{u}(\mathbf{x} \vee \mathbf{y})) f_\mathbf{Y}(\mathcal{R}'_\mathbf{u}(\mathbf{x} \wedge \mathbf{y})) \tag{3.3.9}$$

for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$*, then* $\mathbf{X} \leq_{E_\mathbf{u}} \mathbf{Y}$*.*

Proof. Since $\mathcal{R}_\mathbf{u}$ is an orthogonal matrix, it follows that the Jacobian of the transformations $\mathcal{R}_\mathbf{u}\mathbf{X}$ and $\mathcal{R}_\mathbf{u}\mathbf{Y}$ is equal to 1. Moreover, we also have that $\mathcal{R}_\mathbf{u}^{-1} = \mathcal{R}'_\mathbf{u}$. Therefore, $f_{\mathcal{R}_\mathbf{u}\mathbf{X}}(\mathbf{x}) = f_\mathbf{X}(\mathcal{R}'_\mathbf{u}\mathbf{x})$ and $f_{\mathcal{R}_\mathbf{u}\mathbf{Y}}(\mathbf{x}) = f_\mathbf{Y}(\mathcal{R}'_\mathbf{u}\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^n$. Hence, (3.3.9) iff $f_{\mathcal{R}_\mathbf{u}\mathbf{X}}(\mathbf{x}) f_{\mathcal{R}_\mathbf{u}\mathbf{Y}}(\mathbf{y}) \leq f_{\mathcal{R}_\mathbf{u}\mathbf{X}}(\mathbf{x} \vee \mathbf{y}) f_{\mathcal{R}_\mathbf{u}\mathbf{Y}}(\mathbf{x} \wedge \mathbf{y})$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. We thus get $\mathcal{R}_\mathbf{u}\mathbf{X} \leq_{lr} \mathcal{R}_\mathbf{u}\mathbf{Y}$ and consequently $\mathbf{X} \leq_{E_\mathbf{u}} \mathbf{Y}$. □

A random vector $\mathbf{X}$ has the MTP_2 property if $\mathbf{X} \leq_{lr} \mathbf{X}$ (see for instance Karlin and Rinott [30]). For $\mathbf{X}$ normally distributed it has a MTP_2 iff the off-diagonal elements of $\Sigma_{\mathbf{X}}^{-1}$are nonpositive. Particulary a bivariate normal density is MTP_2 if the correlation coefficient is nonnegative (Karlin and Rinott [31]). We show in Example 3.3.9 that for any normally distributed bivariate random vector $\mathbf{X}$, always there exists a rotation $\mathcal{R}_{\mathbf{u}}$ such that $\mathcal{R}_{\mathbf{u}}\mathbf{X}$ has the MTP_2 property. In fact, the same example also provides a sufficient condition for extremality order in terms of the likelihood order.

Example 3.3.9 *Consider* $\mathbf{X} = (X_1, X_2)'$ *a normally distributed random vector with covariance matrix* $\Sigma_{\mathbf{X}}$*. It is easily seen that* $\Sigma_{\mathbf{X}} = QDQ'$ *where* $Q = (q_{ij})$ *is an orthogonal matrix and* $D = (d_{ii})$ *is a diagonal matrix with nonnegative elements and* $d_{11} \geq d_{22}$*. Let* $\mathbf{u} = (q_{11}, q_{21})'$ *be the first column of the matrix* Q*. Then, according to (3.3.1),*

$$\mathcal{R}_{\mathbf{u}} = \frac{\sqrt{2}}{2} \begin{pmatrix} q_{11} + q_{21} & q_{21} - q_{11} \\ q_{11} - q_{21} & q_{11} + q_{21} \end{pmatrix}.$$

The vector $\mathcal{R}_{\mathbf{u}}\mathbf{X}$ *also is normaly distributed and* $\mathbf{\Sigma}_{\mathcal{R}_{\mathbf{u}}\mathbf{X}} = \mathcal{R}_{\mathbf{u}}\mathbf{\Sigma}_{\mathbf{X}}\mathcal{R}'_{\mathbf{u}}$ *(see, e.g., Valdez and Dhaene [62], Theorem 2).*

It is clear that

$$\mathbf{\Sigma}_{\mathcal{R}_{\mathbf{u}}\mathbf{X}} = \mathcal{R}_{\mathbf{u}}QDQ'\mathcal{R}'_{\mathbf{u}} = \frac{1}{2} \begin{pmatrix} d_{11} + d_{22} & d_{11} - d_{22} \\ d_{11} - d_{22} & d_{11} + d_{22} \end{pmatrix}.$$

As $d_{11} \geq d_{22}$ *we have that* $\mathcal{R}_{\mathbf{u}}\mathbf{X}$ *has the* MTP_2 *property. If we take* $\mathbf{Y} = \mathbf{X} + \mathcal{R}'_{\mathbf{u}}\mathbf{s}$ *where* $\mathbf{s} = (s_1, s_2)'$ *is such that*

$$\frac{d_{11} - d_{22}}{d_{11} + d_{22}} \leq \frac{s_2}{s_1} \leq \frac{d_{11} + d_{22}}{d_{11} - d_{22}},$$

then from Theorem 3.2-b in Rinott and Scarsini [50] we conclude that $\mathcal{R}_{\mathbf{u}}\mathbf{X} \leq_{lr} \mathcal{R}_{\mathbf{u}}(\mathbf{X} + \mathcal{R}'_{\mathbf{u}}\mathbf{s})$
and so, $\mathbf{X} \leq_{E_{\mathbf{u}}} \mathbf{Y}$.

The extremality order satisfies the following closure properties (closure with respect to convergence in distribution and closure with respect to mixture). Here, the notation $[Z|A]$ stands for the random object whose distribution is the conditional distribution of Z given the event A.

Property 3.3.10 a) *Let* $\{\mathbf{X}_j,\ j = 1, 2, \dots\}$ *and* $\{\mathbf{X}_j,\ j = 1, 2, \dots\}$ *be two sequences of random vectors such that* $\mathbf{X}_j \longrightarrow_d \mathbf{X}$ *and* $\mathbf{Y}_j \longrightarrow_d \mathbf{Y}$ *as* $j \longrightarrow \infty$*, where* $\longrightarrow_d$ *denotes convergence in distribution. If* $\mathbf{X}_j \leq_{E_{\mathbf{u}}} \mathbf{Y}_j$ *for all* $j = 1, 2, \dots,$ *then* $\mathbf{X} \leq_{E_{\mathbf{u}}} \mathbf{Y}$.

b) *Let* $\mathbf{X}$*,* $\mathbf{Y}$ *and* Θ *be random vectors such that* $[\mathbf{X} \mid \Theta = \theta] \leq_{E_{\mathbf{u}}} [\mathbf{Y} \mid \Theta = \theta]$ *for all* θ *in the support of* Θ*. Then* $\mathbf{X} \leq_{E_{\mathbf{u}}} \mathbf{Y}$.

Proof. a) Clearly, if $\mathbf{X}_j \longrightarrow_d \mathbf{X}$ and $\mathbf{Y}_j \longrightarrow_d \mathbf{Y}$ as $j \longrightarrow \infty$, then $\mathcal{R}_\mathbf{u}\mathbf{X}_j \longrightarrow_d \mathcal{R}_\mathbf{u}\mathbf{X}$ and $\mathcal{R}_\mathbf{u}\mathbf{Y}_j \longrightarrow_d \mathcal{R}_\mathbf{u}\mathbf{Y}$ as $j \longrightarrow \infty$. But $\mathcal{R}_\mathbf{u}\mathbf{X}_j \leq_{uo} \mathcal{R}_\mathbf{u}\mathbf{Y}_j$ since $\mathbf{X}_j \leq_{E_\mathbf{u}} \mathbf{Y}_j$, for $j = 1, 2, \ldots$. Applying Theorem 6.G.3-d in [55], it follows that $\mathcal{R}_\mathbf{u}\mathbf{X} \leq_{uo} \mathcal{R}_\mathbf{u}\mathbf{Y}$; thus also $\mathbf{X} \leq_{E_\mathbf{u}} \mathbf{Y}$. b) From $[\mathbf{X} \mid \Theta = \theta] \leq_{E_\mathbf{u}} [\mathbf{Y} \mid \Theta = \theta]$, it follows that $[\mathcal{R}_\mathbf{u}\mathbf{X} \mid \Theta = \theta] \leq_{uo} [\mathcal{R}_\mathbf{u}\mathbf{Y} \mid \Theta = \theta]$. The assertion follows from Theorem 6.G.3-e in Shaked and Shanthikumar [55]. □

The following statement describes a property that will be used in Section 3.4. Recall that the *copula* C of a random vector $\mathbf{X}$ is a cumulative distribution function with uniform marginals on [0,1] such that, for all $\mathbf{x} \in \mathbb{R}^n$,

$$F_\mathbf{X}(x_1, \ldots, x_n) = C(F_{X_1}(x_1), \ldots, F_{X_n}(x_n)).$$

The copula of the vector $\mathbf{X}$ describes dependence properties of its components, and it is unique if $F_{X_1}, \ldots, F_{X_n}$ are continuous. For more details about copulas, see Nelsen [47].

Theorem 3.3.11 *Let $\mathbf{X} = (X_1, \ldots, X_n)'$ and $\mathbf{Y} = (Y_1, \ldots, Y_n)'$ be two random vectors and $\mathbf{u} = (u_1, \ldots, u_n)$ a unit vector. If $\mathcal{R}_\mathbf{u}\mathbf{X}$ and $\mathcal{R}_\mathbf{u}\mathbf{Y}$ have the same copula, then*

$$\mathbf{X} \leq_{E_\mathbf{u}} \mathbf{Y} \Longrightarrow \phi(\mathcal{R}_\mathbf{u}^{(1)}\mathbf{X}, \ldots, \mathcal{R}_\mathbf{u}^{(n)}\mathbf{X}) \leq_{st} \phi(\mathcal{R}_\mathbf{u}^{(1)}\mathbf{Y}, \ldots, \mathcal{R}_\mathbf{u}^{(n)}\mathbf{Y})$$

for every increasing function ϕ. In particular, denoted with $\mathcal{R}_\mathbf{u}^{(c.i)}$ the i-th column of the matrix $\mathcal{R}_\mathbf{u}$, and the assumptions it holds

$$\mathbf{X} \leq_{E_\mathbf{u}} \mathbf{Y} \Longrightarrow \sum_{i=1}^{n} \mathbf{a}'\mathcal{R}_\mathbf{u}^{(c.i)} X_i \leq_{st} \sum_{i=1}^{n} \mathbf{a}'\mathcal{R}_\mathbf{u}^{(c.i)} Y_i$$

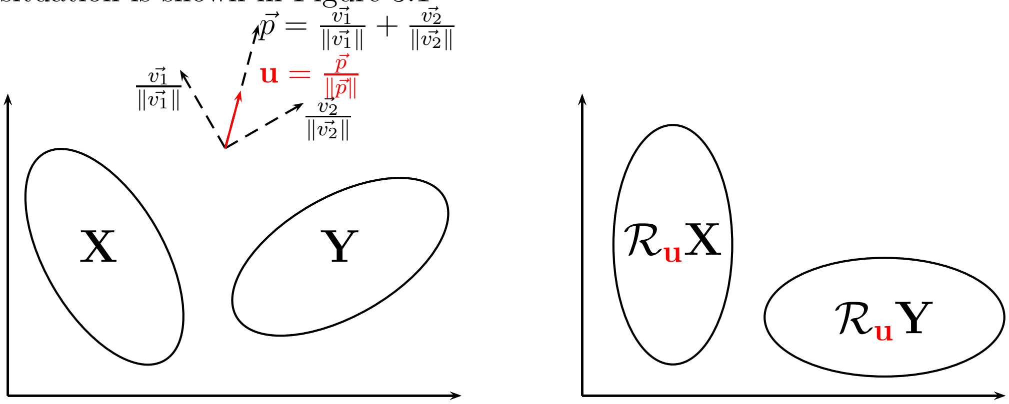
for every vector $\mathbf{a} = (a_1, \ldots, a_n)'$ with nonnegative components.

Proof. As $\mathbf{X} \leq_{E_\mathbf{u}} \mathbf{Y}$, then from Property 3.3.4, $\mathcal{R}_\mathbf{u}^{(r.i)}\mathbf{X} \leq_{st} \mathcal{R}_\mathbf{u}^{(r.i)}\mathbf{Y}$ for all $i = 1, \ldots, n$. Since $\mathcal{R}_\mathbf{u}\mathbf{X}$ and $\mathcal{R}_\mathbf{u}\mathbf{Y}$ have the same copula, by Theorem 6.B.14. in Shaked and Shanthikumar [55] it follows that $\mathcal{R}_\mathbf{u}\mathbf{X} \leq_{st} \mathcal{R}_\mathbf{u}\mathbf{Y}$. Now, by Theorem 3.3.11 in Müller and Stoyan [46], the first assertion follows. In particular, letting $\phi(x_1, \ldots, x_n) = (a_1, \ldots, a_n)(x_1, \ldots, x_n)'$, where $\mathbf{a} = (a_1, \ldots, a_n)'$ is a vector with non-negative components, we get

$$\mathbf{a}'\mathcal{R}_\mathbf{u}\mathbf{X} \leq_{st} \mathbf{a}'\mathcal{R}_\mathbf{u}\mathbf{Y},$$

i.e., the second assertion. □

Note that, in particular, Theorem 3.3.11 applies when the unit vector $\mathbf{u}$ is such that $\mathcal{R}_\mathbf{u}\mathbf{X}$ and $\mathcal{R}_\mathbf{u}\mathbf{Y}$ have the copula $C(u, v) = uv$, i.e., when they have independent components. To find a unit vector $\mathbf{u}$ such that $\mathcal{R}_{\mathbf{u}}\mathbf{X}$ and $\mathcal{R}_{\mathbf{u}}\mathbf{Y}$ have the same copula in some situations it is easy. For example, assume that $\mathbf{X}$ and $\mathbf{Y}$ are normally distributed with covariance matrix $\Sigma_{\mathbf{X}}$ and $\Sigma_{\mathbf{Y}}$, respectively. Assume that the eigenvectors $\vec{v_1}$, $\vec{v_2}$ of $\Sigma_{\mathbf{X}}$ and $\Sigma_{\mathbf{Y}}$ are the same. If we define $\vec{p} = \frac{\vec{v_1}}{\|\vec{v_1}\|} + \frac{\vec{v_2}}{\|\vec{v_2}\|}$ and $\mathbf{u} = \frac{\vec{p}}{\|\vec{p}\|}$, then $\mathcal{R}_{\mathbf{u}}\mathbf{X}$ and $\mathcal{R}_{\mathbf{u}}\mathbf{Y}$ will have the same copula. A graphical representation of this situation is shown in Figure 3.1



(a) different copula **(b) same copula**

Figure 3.1: Rotations of a Normal distribution

Figure 3.1 shows that in the original system, before the rotation, $\mathbf{X}$ has a negative dependency and $\mathbf{Y}$ has a positive dependence. Therefore, their copulas are different. After of the rotation in the directions $\mathbf{u}$, indicated above, $\mathcal{R}_{\mathbf{u}}\mathbf{X}$ and $\mathcal{R}_{\mathbf{u}}\mathbf{Y}$ have the same copula, i.e., $C(u, v) = uv$. It is interesting to observe that, if $\mathbf{X}$ and $\mathbf{Y}$ have a common copula, $\mathcal{R}_{\mathbf{u}}\mathbf{X}$ and $\mathcal{R}_{\mathbf{u}}\mathbf{Y}$ may not have common copula. In fact, consider $\mathbf{X}$ to have $N(\mu, I_2)$ distribution and $\mathbf{Y}$ have $N(\mu, D_2)$ distribution, where I_2 is the identical matrix and D_2 is a diagonal matrix with $d_{11} > d_{22}$. Clearly, $\mathbf{X}$ has a spherical distribution, and $\mathbf{X}$ and $\mathbf{Y}$ have a common copula. Let now $\mathbf{u} = [1, 0]'$ be the rotation vector. In the Figure 3.2 is shown that $\mathcal{R}_{\mathbf{u}}\mathbf{X}$ has the same copula as before the rotation, but $\mathcal{R}_{\mathbf{u}}\mathbf{Y}$ has a different copula copula since positive dependency can be observed.

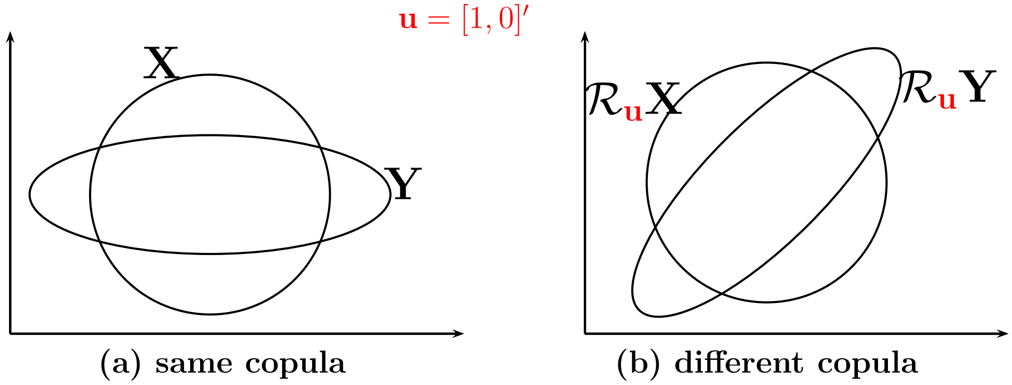


Figure 3.2: Rotations of a Normal distribution

Corollary 3.3.12 *Under the same assuptions of Theorem 3.3.11 it holds that*

$$\mathbf{X} \leq_{E_{\mathbf{u}}} \mathbf{Y} \Longrightarrow \mathbf{u}'\mathbf{X} \leq_{st} \mathbf{u}'\mathbf{Y}$$

Proof. We need only consider $\mathbf{a} = [1, \ldots, 1]'$ and the assertion follows immediately from Theorem 3.3.11 and formula (3.3.7) $\square$

The Corollary 3.3.12 follows also from the fact that if two random vectors are ordered in the multivariate stochastic order, then the sum of their components are ordered in the univariate stochastic order (see, e.g. Theorem 6.B.16-a in Shaked and Shanthikumar [55]). In Theorem 3.3.11 and Corollary 3.3.12 it is worth noting that the coefficients $\mathbf{a}'\mathcal{R}_{\mathbf{u}}^{(c.i)}$ and components of $\mathbf{u}'$ may also be negative; thus Theorem 3.3.11 gives conditions to compare, in usual stochastic order, non-positive linear combinations of dependent random variables.

If the vectors $\mathbf{X}$ and $\mathbf{Y}$ have the same means, then the following property holds.

Theorem 3.3.13 *Let $\mathbf{X} = (X_1, \ldots, X_n)'$ and $\mathbf{Y} = (Y_1, \ldots, Y_n)'$ be two random vectors such that $E[\mathbf{X}] = E[\mathbf{Y}]$. If there exists a vector $\mathbf{u}$ such that $\mathbf{X} \leq_{E_{\mathbf{u}}} \mathbf{Y}$, then*

$$\mathcal{R}_{\mathbf{u}}^{(r.i)}\mathbf{X} =_{st} \mathcal{R}_{\mathbf{u}}^{(r.i)}\mathbf{Y}, \quad \textit{for all } i = 1, \ldots, n.$$

Proof. Since $\mathbf{X} \leq_{E_{\mathbf{u}}} \mathbf{Y}$, from Property 3.3.4, we have $\mathcal{R}_{\mathbf{u}}^{(r.i)}\mathbf{X} \leq_{st} \mathcal{R}_{\mathbf{u}}^{(r.i)}\mathbf{Y}, \quad i = 1, \ldots, n$. Observe that

$$E[\mathbf{X}] = E[\mathbf{Y}] \Longrightarrow \mathcal{R}_{\mathbf{u}}E[\mathbf{X}] = \mathcal{R}_{\mathbf{u}}E[\mathbf{Y}] \Longrightarrow E[\mathcal{R}_{\mathbf{u}}\mathbf{X}] = E[\mathcal{R}_{\mathbf{u}}\mathbf{Y}]. \qquad (3.3.10)$$

Now the proof follows from the fact that variables ordered in usual stochastic order having the same mean should also have the same distribution (see Theorem 1.A.8. in Shaked and Shanthikumar [55]). $\square$

Relationships between extremality order and two positive dependence orders are obtained in the next result.

Theorem 3.3.14 *Let* $\mathbf{X} = (X_1, \ldots, X_n)'$ *and* $\mathbf{Y} = (Y_1, \ldots, Y_n)'$ *be two random vectors such that* $E[\mathbf{X}] = E[\mathbf{Y}]$. *If* $\mathbf{X} \leq_{E_\mathbf{u}} \mathbf{Y}$ *then* $\mathcal{R}_\mathbf{u}\mathbf{X} \leq_{PQD} \mathcal{R}_\mathbf{u}\mathbf{Y}$. *Moreover, if* $n = 2$, *it follows also that* $\mathcal{R}_\mathbf{u}\mathbf{X} \leq_{sm} \mathcal{R}_\mathbf{u}\mathbf{Y}$.

Proof. From Theorem 3.3.13 it is easily seen that $\mathcal{R}_\mathbf{u}\mathbf{X}$ and $\mathcal{R}_\mathbf{u}\mathbf{Y}$ have the same marginals. Since $\mathbf{X} \leq_{E_\mathbf{u}} \mathbf{Y} \Longleftrightarrow \mathcal{R}_\mathbf{u}\mathbf{X} \leq_{uo} \mathcal{R}_\mathbf{u}\mathbf{Y}$, the first assertion follows from definition of positive quadrant dependence order, while the second from the equivalency between $\leq_{PDQ}$ and $\leq_{sm}$ when $n = 2$. $\square$

Corollary 3.3.15 *Let* $\mathbf{X} = (X_1, X_2)'$ *and* $\mathbf{Y} = (Y_1, Y_2)'$ *be two random vectors such that* $E[\mathbf{X}] = E[\mathbf{Y}]$. *If* $\mathbf{X} \leq_{E_\mathbf{u}} \mathbf{Y}$ *then* $\mathbf{a}'\mathcal{R}_\mathbf{u}\mathbf{Y} \leq_{cv} \mathbf{a}'\mathcal{R}_\mathbf{u}\mathbf{X}$ *for all* $\mathbf{a} \in \mathbf{R}^2$ *such that* $a_1a_2 \geq 0$.

Proof. According to Theorem 3.3.14, it holds $\mathcal{R}_\mathbf{u}\mathbf{X} \leq_{sm} \mathcal{R}_\mathbf{u}\mathbf{Y}$. Hence, from Theorem 9.A.9 in Shaked and Shanthikumar [55],

$$\begin{pmatrix} a_1 & 0 \\ 0 & a_2 \end{pmatrix} \mathcal{R}_\mathbf{u}\mathbf{X} \leq_{sm} \begin{pmatrix} a_1 & 0 \\ 0 & a_2 \end{pmatrix} \mathcal{R}_\mathbf{u}\mathbf{Y}, \tag{3.3.11}$$

for all a_1, a_2 such that $a_1a_2 \geq 0$. Combining (3.3.11) and (9.A.19) in Shaked and Shanthikumar [55], we can assert that

$$\mathbf{a}'\mathcal{R}_\mathbf{u}\mathbf{X} \leq_{cx} \mathbf{a}'\mathcal{R}_\mathbf{u}\mathbf{Y} \text{ and } \mathbf{a}'\mathcal{R}_\mathbf{u}\mathbf{Y} \leq_{cv} \mathbf{a}'\mathcal{R}_\mathbf{u}\mathbf{X}.$$

$\square$ As a particular case of Corollary 3.3.15, with $\mathbf{a} = \frac{1}{\sqrt{2}}[1, 1]'$, we can also conclude that, under the same assumptions,

$$\mathbf{X} \leq_{E_\mathbf{u}} \mathbf{Y} \Longrightarrow \mathbf{u}'\mathbf{Y} \leq_{cv} \mathbf{u}'\mathbf{X}.$$

We finish this Section with a necessary condition for the extremality order. To this end, recall that given two non-negative random variables X and Y, X is said to be smaller than Y in Laplace transform order (briefly $X \leq_{Lt} Y$) if only if $E[\exp^{-sX}] \geq E[\exp^{-sY}]$, for all $s \in \mathbb{R}^+$.

Theorem 3.3.16 *Let* $\mathbf{X}$, $\mathbf{Y}$ *be two random vectors and* $\mathbf{u}$ *a unit vector such that* $\mathcal{R}_\mathbf{u}\mathbf{X}$ *and* $\mathcal{R}_\mathbf{u}\mathbf{Y}$ *are positive. If* $\mathbf{X} \leq_{E_{(-\mathbf{u})}} \mathbf{Y}$ *then* $\mathbf{u}'\mathbf{X} \geq_{Lt} \mathbf{u}'\mathbf{Y}$ *and, in particular,* $E[\mathbf{u}'\mathbf{X}] \geq E[\mathbf{u}'\mathbf{Y}]$.

Proof. Since $\mathbf{X} \leq_{E_{(-\mathbf{u})}} \mathbf{Y} \Longleftrightarrow \mathcal{R}_{\mathbf{u}}\mathbf{X} \geq_{lo} \mathcal{R}_{\mathbf{u}}\mathbf{Y}$, and since $\mathcal{R}_{\mathbf{u}}\mathbf{X}$ and $\mathcal{R}_{\mathbf{u}}\mathbf{Y}$ are positive, then from Theorem 6.G.14 in Shaked and Shanthikumar [55] it follows

$$\sum_{i=1}^{n} a_i \mathcal{R}_{\mathbf{u}}^{(r.i)}\mathbf{X} \geq_{Lt} \sum_{i=1}^{n} a_i \mathcal{R}_{\mathbf{u}}^{(r.i)}\mathbf{Y}, \quad \text{whenever } \ a_i \geq 0, \ i = 1, 2, \ldots, n.$$

Assuming $a_i = \frac{1}{\sqrt{n}}$, for all $i = 1, 2, \ldots, n$, the assertion follows from

$$\mathbf{1}'\mathcal{R}_{\mathbf{u}}\mathbf{X} = \sqrt{n}\mathbf{u}'\mathbf{X}.$$

The second inequality follows as a consequence of the comparison in Laplace transform order. □

3.4 Portfolio comparisons with the extremality order

In this Section we describe some examples of application of previous results to the the single period portfolio problem. We first describe the problem. Consider an economic agent with unitary initial capital. Suppose that the random variables $X_1, \cdots, X_n$ represent the outcome of n financial positions which can be chosen for investment. Thus we have risks $-X_1, \cdots, -X_n$. In this context, a portfolio is a random variable $Z_{\mathbf{a}} = \sum_{i=1}^{n} a_i X_i$, where the weights vector $\mathbf{a} = (a_1, \ldots, a_n)$ range in the subset $\mathbf{A}_n = \{\mathbf{a} = (a_1, \ldots, a_n) : \sum_{i=1}^{n} a_i = 1, a_i \geq 0, i = 1, \ldots, n\}$. When short selling are permitted, then the condition $a_i \geq 0$ can be removed. The goal of the single period portfolio problem consists in determining the allocation $\mathbf{a} = \{a_1, \ldots, a_n\}$ of the unitary wealth to the n risks that minimize the total risk, or that maximize the expected utility of the resulting final wealth $Z_{\mathbf{a}}$.

A first problem that the economic agent can consider is the minimization of the risk, which is commonly expressed as the minimization of the *value at risk* at a fixed quantile α (VaR_α) of the random loss $-Z_{\mathbf{a}}$ (see, e.g., Jorion [29]). It represents the α-quantile of the loss distribution of portfolio. This means that VaR_α is the better loss among the $(1-\alpha)100\%$ worst losses, and it is formally defined as follows: if F is the distribution of $-Z_{\mathbf{a}}$ and $\alpha \in (0, 1)$, then

$$VaR_\alpha(Z_a) = \inf\{z \in \mathbb{R}; | \ F(z) \geq \alpha\}. \tag{3.4.1}$$

Let us first consider the case that the risk manager wants to allocate the wealth to n risks, and he/she has to chose between two sets of dependent financial positions, say $\mathbf{X} = (X_1, \ldots, X_n)'$ or $\mathbf{Y} = (Y_1, \ldots, Y_n)'$. Assume that there exists a vector $\mathbf{u} = (u_1, \ldots, u_n)'$ such that the assumptions of Theorem 3.3.11 are satisfied (i.e., $\mathbf{X} \leq_{E_\mathbf{u}} \mathbf{Y}$ and the rotated vectors have the same copula). Let $\mathbf{a} = (a_1, \ldots, a_n)'$ be any vector such that $\frac{\mathbf{a}}{\mathbf{a}'\mathcal{R}_\mathbf{u}\mathbf{1}} \geq \mathbf{0}$.

Since the usual stochastic order implies the corresponding order between values at risk for every α (see, e.g., Bäuerle and Müller [5]), then, for every allocation

$$\omega = \left(\frac{\mathbf{a}'\mathcal{R}_\mathbf{u}^{(c.1)}}{\mathbf{a}'\mathcal{R}_\mathbf{u}\mathbf{1}}, \ldots, \frac{\mathbf{a}'\mathcal{R}_\mathbf{u}^{(c.n)}}{\mathbf{a}'\mathcal{R}_\mathbf{u}\mathbf{1}} \right)',$$

the economic agent will choose the portfolio $Z_\mathbf{Y} = \sum_{i=1}^n \omega_i Y_i$ rather than $Z_\mathbf{X} = \sum_{i=1}^n \omega_i X_i$. It is interesting to observe that some of the weights ω_i can be negative, so that even portfolios with short selling can be compared.

Similarly, if the economic agent is non-satiable, which means that he/she has an increasing utility function U, under the same conditions as above her/his expected utility will be maximized choosing the portfolio $Z_\mathbf{Y}$ instead of $Z_\mathbf{X}$, since

$$Z_\mathbf{X} \leq_{st} Z_\mathbf{Y} \Longrightarrow E\left[U\left(Z_\mathbf{X}\right)\right] \leq E\left[U\left(Z_\mathbf{Y}\right)\right],$$

for any increasing function U.

For illustrating this result, we introduce an example in the bivariate case.

Example 3.4.1 *Let $\mathbf{X} = (X_1, X_2)'$ and $\mathbf{Y} = (Y_1, Y_2)'$ be bivariate normally distributed random vectors with mean $\mu_\mathbf{X}$ and $\mu_\mathbf{Y}$, respectively and the same covariance matrix Σ. Choose any unit vector $\mathbf{u} = (u_1, u_2)'$ such that $\mu_\mathbf{Y} \in \mathcal{C}_{\mu_\mathbf{X}}^\mathbf{u}$, (for example the vector $\mathbf{u} = \frac{\mu_\mathbf{Y} - \mu_\mathbf{X}}{\|\mu_\mathbf{Y} - \mu_\mathbf{X}\|}$). According to (3.3.1), the rotation matrix is given by*

$$\mathcal{R}_\mathbf{u} = \frac{\sqrt{2}}{2} \begin{pmatrix} u_1 + u_2 & u_2 - u_1 \\ u_1 - u_2 & u_1 + u_2 \end{pmatrix}.$$

It is easy to check that $\mathcal{R}_\mathbf{u}\mathbf{X}$ and $\mathcal{R}_\mathbf{u}\mathbf{Y}$ have the same copula. Let now $Z_\mathbf{X}$ and $Z_\mathbf{Y}$ be two portfolios defined as

$$Z_\mathbf{X} = \frac{1}{2}\left(1 + \frac{a_1u_1 - a_2u_2}{a_1u_2 + a_2u_1}\right) X_1 + \frac{1}{2}\left(1 - \frac{a_1u_1 - a_2u_2}{a_1u_2 + a_2u_1}\right) X_2$$
$$Z_\mathbf{Y} = \frac{1}{2}\left(1 + \frac{a_1u_1 - a_2u_2}{a_1u_2 + a_2u_1}\right) Y_1 + \frac{1}{2}\left(1 - \frac{a_1u_1 - a_2u_2}{a_1u_2 + a_2u_1}\right) Y_2,$$

where $\mathbf{X}$ and $\mathbf{Y}$ are financial positions which can be chosen for investment. By Theorem 3.3.11, we have that

$$Z_\mathbf{X} \leq_{st} Z_\mathbf{Y}, \tag{3.4.2}$$

for every $\mathbf{a} = (a_1, a_2)'$ *such that* $\frac{\mathbf{a}}{\mathbf{a}'\mathcal{R}_{\mathbf{u}}\mathbf{1}} \geq \mathbf{0}$. *Therefore, if an investor measures the risk through value at risk, then he/she prefers the portfolio* $Z_{\mathbf{Y}}$ *instead* $Z_{\mathbf{Y}}$ *since from (3.4.2) it follows that,*

$$VaR_{\alpha}(-Z_{\mathbf{Y}}) \leq VaR_{\alpha}(-Z_{\mathbf{X}}),$$

for all $\alpha \in (0,1)$. *Indeed, if the investor is non-satiable the same conclusion can be drawn. It is also remarkable that the weights of the two portfolios can assume negative values.*

Now, let us consider the case where the economic agent has to chose between two sets of risks, say $\mathbf{X} = (X_1, X_2)$ or $\mathbf{Y} = (Y_1, Y_2)$, but assume here that $E[\mathbf{X}] = E[\mathbf{Y}]$. Thus the expectation is the same for any linear combination of the risks. In fact, in such a case, given any two portfolios

$$Z_{\mathbf{X}} = \sum_{i=1}^{2} b_i X_i \text{ and } Z_{\mathbf{Y}} = \sum_{i=1}^{2} b_i Y_i,$$

they cannot be ordered in usual stochastic order, since they have the same expectations. However, the economic agent can prefer one of the two portfolios if she/he, beside of being non-satiable, is risk averse. In fact, in this situation, the utility function U is increasing and concave (see, e.g., Yaari [63]), and a comparison in the concave order can be used as a criteria to choose between portfolios.

By a direct application of Corollary 3.3.15, it can be asserted that, if $\mathbf{X} \leq_{E_{\mathbf{u}}} \mathbf{Y}$, the the expected utility of the agent will be greater choosing the portfolio $Z_{\mathbf{X}} = \sum_{i=1}^{2} \mathbf{a}'\mathcal{R}_{\mathbf{u}}^{(c.i)} X_i$ rather than $Z_{\mathbf{Y}} = \sum_{i=1}^{2} \mathbf{a}'\mathcal{R}_{\mathbf{u}}^{(c.i)} Y_i$, for all $\mathbf{a} = [a_1, a_2]'$ such that $a_1 a_2 \geq 0$. As above, this fact becomes particularly interesting whenever some of the allocations to the risks are negative, thus allowing comparisons also in the case of short selling.

Example 3.4.2 *Let* $\mathbf{X} = (X_1, X_2)'$ *and* $\mathbf{Y} = (Y_1, Y_2)'$ *be two random vectors elliptically distributed such that* $E[\mathbf{X}] = E[\mathbf{Y}]$. *Assume that the covariance matrices* $\Sigma_{\mathbf{X}}$ *and* $\Sigma_{\mathbf{Y}}$ *have the same eigenvalues* d_1 *and* d_2. *Let*

$$Q = \begin{pmatrix} q_1 & q_2 \\ q_2 & -q_1 \end{pmatrix} \quad \textit{and} \quad T = \begin{pmatrix} t_1 & t_2 \\ t_2 & -t_1 \end{pmatrix}$$

be the eigenvectors matrices of $\Sigma_{\mathbf{X}}$ *and* $\Sigma_{\mathbf{Y}}$, *respectively. It is clear that* $\Sigma_{\mathbf{X}} = QDQ'$ *and* $\Sigma_{\mathbf{Y}} = TDT'$ *where* $D = diag(d_1, d_2)$. *Let* $\mathbf{u} = \frac{(Q+T)\mathbf{1}}{\|(Q+T)\mathbf{1}\|}$ *be the unit vector, this gives*

$$\mathbf{u} = \frac{1}{\| (Q+T)\mathbf{1} \|} [q_1 + t_1 + q_2 + t_2, \ q_2 + t_2 - q_1 - t_1]'.$$

From (3.3.1) we have that

$$\mathcal{R}_{\mathbf{u}} = \frac{\sqrt{2}}{\| (Q+T)\mathbf{1} \|} \begin{pmatrix} q_2+t_2 & -q_1-t_1 \\ q_1+t_1 & q_2+t_2 \end{pmatrix}.$$

By straightforward calculations we can see that

$$\Sigma_{\mathcal{R}_{\mathbf{u}}\mathbf{X}} = \mathcal{R}_{\mathbf{u}} Q D Q' \mathcal{R}'_{\mathbf{u}} \quad \textit{and} \ \Sigma_{\mathcal{R}_{\mathbf{u}}\mathbf{Y}} = \mathcal{R}_{\mathbf{u}} T D T' \mathcal{R}'_{\mathbf{u}},$$

have the same diagonal and their off-diagonal elements are given by $\rho = (d_1 - d_2)(q_1t_1 + q_2t_2 + 1)(q_1t_1 - q_2t_2)$ *and* $(-\rho)$ *respectively. Without lack of generality assume* $\rho \leq 0$*, then* $\mathcal{R}_{\mathbf{u}}\mathbf{X} \leq_{pqd} \mathcal{R}_{\mathbf{u}}\mathbf{Y}$ *(see Landsman and Tsanakas [35], Corollary 2.) and consequently* $\mathbf{X} \leq_{E_{\mathbf{u}}} \mathbf{Y}$*. Therefore, if we consider the random variables*

$$\begin{aligned} Z_{\mathbf{X}} &= \frac{1}{2}\left(1 + \frac{q_1+t_1}{q_2+t_2}\right) X_1 + \frac{1}{2}\left(1 - \frac{q_1+t_1}{q_2+t_2}\right) X_2, \\ Z_{\mathbf{Y}} &= \frac{1}{2}\left(1 + \frac{q_1+t_1}{q_2+t_2}\right) Y_1 + \frac{1}{2}\left(1 - \frac{q_1+t_1}{q_2+t_2}\right) Y_2, \end{aligned}$$

and the Corollary 3.3.15 with $\mathbf{a} = \frac{\|(Q+T)\mathbf{1}\|}{2\sqrt{2}(q_2+t_2)}[1,1]'$*, we have that* $Z_{\mathbf{Y}} \leq_{cv} Z_{\mathbf{X}}$ *and* $Z_{\mathbf{X}} \leq_{cx} Z_{\mathbf{Y}}$*. Then a risk averse rational decision maker would prefer the portfolio* $Z_{\mathbf{X}}$*.*

Moreover, the idea above allows also for criteria based on comparisons of either portfolio variances or risks measured through a convex measure in the sense of Föllmer and Schied [20]. *Conditional Value at Risk* (CVaR) is a convex risk measure; therefore under the same hypothesis as above, $Z_{\mathbf{X}}$ has less variance and smaller (CVaR) than the portfolio $Z_{\mathbf{Y}}$ since variance and CVaR are consistent with respect to the convex order (see Shaked and Shanthikumar [55] pages 110 to 112 and Bäuerle and Mülle [5]).

Consider now the case in which the agent has to allocate his capital in two different but not independent risky assets X_1 and X_2. A typical problem in portfolio theory is the determination of the allocation $\alpha \in [0,1]$ such that $Z_\alpha = (1-\alpha)X_1 + \alpha X_2$ maximizes the expected utility $h(\alpha) = E[U(Z_\alpha)]$, where U is the increasing and concave utility function of the agent. In 1971, Hadar and Russel proved that, if X_1 and X_2 are iid, then $h(\frac{1}{2}) \geq h(\alpha)$ for all $\alpha \in [0,1]$, thus proving that the maximal diversification gives the maximal expected utility under the assumptions above. This result was generalized in Ma [42] to the multivariate case, replacing the assumption of independence with the assumption of exchangeability.

Related results have been provided in Pellerey and Semeraro [48]. Specifically, they proved that if the vector (S, D) of the sum $S = X_1 + X_2$ and the difference

$D = X_2 - X_1$ of the risks is *positive quadrant dependent (PQD)*, i.e., if $(S, D) \geq_{PQD} (S^{\perp}, D^{\perp})$ where $(S^{\perp}, D^{\perp})$ is the independent version of (S, D), and if $E[X_2] \leq E[X_1]$, then $h(\alpha)$ is decreasing in $\alpha = [\frac{1}{2}, 1]$. Similarly, they proved that if the vector (S, D) is *negative quadrant dependent, NQD,* (i.e., if $(S, D) \leq_{PQD} (S^{\perp}, D^{\perp})$) and $E[X_2] \geq E[X_1]$, then $h(\alpha)$ is increasing in $\alpha = [0, \frac{1}{2}]$. As a consequence of these results we have that if the vector (X_1, X_2) is such that $E[X_2] = E[X_1]$, and S and D are uncorrelated, then $h(\frac{1}{2}) \geq h(\alpha)$ for all $\alpha \in [0, 1]$. A generalization of this result is given in the following Theorem.

Theorem 3.4.3 *Let* $\mathbf{X} = (X_1, X_2)$ *be random vector. Consider* $\mathbf{u} = [u_1, u_2]'$ *a unit vector and* $\mathbf{v} = \frac{\sqrt{2}}{2}[u_1 - u_2, u_1 + u_2]'$. *Let* $\mathbf{Z} = (Z_1, Z_2)$ *be any random vector of independent components with mean* $\mathcal{R}_{\mathbf{v}} E[\mathbf{X}]$ *and let* $\mathbf{Y} = \mathcal{R}_{\mathbf{v}}' \mathbf{Z}$. *If* $\mathbf{Y} \leq_{E_{\mathbf{v}}} [\geq_{E_{\mathbf{v}}}] \mathbf{X}$ *and* $u_1 E[X_2] \leq [\geq]\ u_2 E[X_1]$, *then for every increasing and concave utility function* U *it holds that*

$$
\begin{aligned}
&E\left[U(\frac{\sqrt{2}}{2}u_1X_1 + \frac{\sqrt{2}}{2}u_2X_2)\right] \\
&\geq E\left[U\left(\frac{\sqrt{2}}{2}\left(u_1 + u_2 - 2u_2\alpha\right)X_1 + \frac{\sqrt{2}}{2}\left(2u_1\alpha - u_1 + u_2\right)X_2\right)\right]
\end{aligned}
$$

for all $\alpha \in [\frac{1}{2}, 1]$ *[for all* $\alpha \in [0, \frac{1}{2}]$*].*

Proof. We see the proof the case $\leq$, the another case is similar. It is immediate that $E[\mathbf{Y}] = E[\mathbf{X}]$. From Theorem 3.3.14 we deduce that $\mathcal{R}_{\mathbf{v}}\mathbf{Y} \leq_{PQD} \mathcal{R}_{\mathbf{v}}\mathbf{X}$ since $\mathbf{Y} \leq_{E_{\mathbf{v}}} \mathbf{X}$. However, $\mathcal{R}_{\mathbf{v}}\mathbf{Y}$ has independent components, so $\mathcal{R}_{\mathbf{v}}\mathbf{X}$ is PQD. On account of (3.3.1) we obtain

$$
\begin{pmatrix} u_1 & u_2 \\ -u_2 & u_1 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \quad \textit{is PQD.} \tag{3.4.3}
$$

Using the formula (3.3.1) to construct $\mathcal{R}_{\mathbf{u}}$, with $\mathbf{u} = [u_1, u_2]'$, we have that

$$
\mathcal{R}_{\mathbf{u}}\mathbf{X} = \frac{\sqrt{2}}{2} \begin{pmatrix} u_1 + u_2 & u_2 - u_1 \\ u_1 - u_2 & u_1 + u_2 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}. \tag{3.4.4}
$$

Denote X_i^R the *i-th* component of the vector $\mathcal{R}_{\mathbf{u}}\mathbf{X}$, $i = 1, 2$, and let $(S^R, D^R)'$ be the vector of the sum and the difference of the components of $(X_1^R, X_2^R)'$. Therefore, from (3.4.4),

$$
\begin{aligned}
S^R &= X_1^R + X_2^R = \sqrt{2}u_1X_1 + \sqrt{2}u_2X_2, \\
D^R &= X_2^R - X_1^R = \sqrt{2}u_1X_2 - \sqrt{2}u_2X_1.
\end{aligned}
$$

From (3.4.3) we have that $(S^R, D^R)'$ is PQD. Since $u_1 E[X_2] \leq u_2 E[X_1]$, it follows that $E[X_2^R] \leq E[X_1^R]$. By using Theorem 2.1 in Pellerey and Semeraro [48], we get that

$$Z_\alpha = (1-\alpha)X_1^R + \alpha X_2^R = \frac{\sqrt{2}}{2}\left(u_1 + u_2 - 2u_2\alpha\right) X_1 + \frac{\sqrt{2}}{2}\left(2u_1\alpha - u_1 + u_2\right) X_2$$

is decreasing in the concave order in $\alpha \in [\frac{1}{2}, 1]$. Thus, in particular, for every increasing and concave utility function U it holds

$$\begin{aligned} & E\left[U(\frac{\sqrt{2}}{2}u_1X_1 + \frac{\sqrt{2}}{2}u_2X_2)\right] \\ & \geq E\left[U\left(\frac{\sqrt{2}}{2}\left(u_1 + u_2 - 2u_2\alpha\right) X_1 + \frac{\sqrt{2}}{2}\left(2u_1\alpha - u_1 + u_2\right) X_2\right)\right] \end{aligned}$$

for all $\alpha \in [\frac{1}{2}, 1]$. □

Remark 3.4.4 *Note that if there exists a vector $u = (u_1, u_2)$ such that*

$$\begin{pmatrix} u_1 & u_2 \\ -u_2 & u_1 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$$

has uncorrelated components (so that it is neither PQD nor NQD), then the maximal expected utility is reached in

$$Z_{1/2} = \frac{\sqrt{2}}{2}u_1X_1 + \frac{\sqrt{2}}{2}u_2X_2.$$

Since $\frac{\sqrt{2}}{2}u_1 + \frac{\sqrt{2}}{2}u_2 \leq 1$, this means that under such conditions it is not necessary to invest totally the available capital. It also is interesting to observe that some of the allocations can be negative, thus the case of short selling are allowed.

3.5 Optimal portfolio selection through rotations

We now consider the particular case that the risks have joint elliptical distributions that is a common assumption in portfolio theory since they allow for the presence of heavy tails and asymptotic tail dependence distributions. The importance and applications of elliptical distributions for risk management and insurance have been widely studied by Embrechts et al. [16], Landsman [34] and Landsman and Valdez [33].

Definition 3.5.1 *The random vector* $\mathbf{X} = (X_1, \ldots, X_n)'$ *is said to have an* elliptical distribution *with parameters* μ *and* $\mathbf{\Sigma}$ *if its characteristic function can be expressed as*

$$E[\exp(i\mathbf{t}'X)] = \exp(i\mathbf{t}'\mu)\phi\left(\mathbf{t}'\mathbf{\Sigma}\mathbf{t}\right), \quad \mathbf{t} = (t_1, \ldots, t_n)', \tag{3.5.1}$$

for some function ϕ, *and if* $\mathbf{\Sigma}$ *is such that* $\mathbf{\Sigma} = \mathbf{A}\mathbf{A}'$ *for some matrix* $\mathbf{A}(n \times m)$.

The moments of $\mathbf{X}$ do not necessarily exist. However, if the mean vector exits, then it is the parameter μ. Besides, if the covariance matrix exists, then it is given by $Cov(\mathbf{X}) = -2\phi'(0)\mathbf{\Sigma}$. A necessary condition for this covariance matrix to exist is $\| \phi'(0) \| < \infty$, where ϕ' denotes the first derivative of the characteristic generator ϕ of $\mathbf{X}$. Note that the class of multivariate elliptical distribution with $\phi(x) = \exp(-\frac{x}{2})$ corresponds to the class of multivariate normal distribution.

Recall that the components of a random vector $\mathbf{X} = (X_1, \ldots, X_n)'$ are said to be *exchangeable* if for any permutation matrix $\mathbf{P}$ the vector $\mathbf{PX}$ has the same distribution as $\mathbf{X}$. The following property will be used later.

Property 3.5.2 *Let* $\mathbf{X} = (X_1, X_2)$ *be a random vector elliptically distributed with parameters* $\mu = \mathbf{0}$ *and* $\mathbf{\Sigma}_{\mathbf{X}}$. *Then there exists a unit vector* $\mathbf{u}$ *such that* $\mathcal{R}_{\mathbf{u}}\mathbf{X}$ *is exchangeable.*

Proof. Since $\mathbf{\Sigma}_{\mathbf{X}}$ is a symmetric matrix, then it can be expressed as $\mathbf{\Sigma}_{\mathbf{X}} = QDQ'$, where $Q = (q_{ij})$ is an orthogonal matrix and $D = (d_{ii})$ is a diagonal matrix with non-negative elements. Let $\mathbf{u} = (q_{11}, q_{21})'$ be the first column of the matrix Q. Then, according to (3.3.1),

$$\mathcal{R}_{\mathbf{u}} = \frac{\sqrt{2}}{2}\begin{pmatrix} q_{11} + q_{21} & q_{21} - q_{11} \\ q_{11} - q_{21} & q_{11} + q_{21} \end{pmatrix}.$$

The vector $\mathcal{R}_{\mathbf{u}}\mathbf{X}$ also is elliptically distributed, with parameters $\mathcal{R}_{\mathbf{u}}\mu = \mathbf{0}$ and $\mathbf{\Sigma}_{\mathcal{R}_{\mathbf{u}}\mathbf{X}} = \mathcal{R}_{\mathbf{u}}\mathbf{\Sigma}_{\mathbf{X}}\mathcal{R}_{\mathbf{u}}'$ (see, e.g., Valdez and Dhaene [62], Theorem 2).

It is clear that

$$\mathbf{\Sigma}_{\mathcal{R}_{\mathbf{u}}\mathbf{X}} = \mathcal{R}_{\mathbf{u}}QDQ'\mathcal{R}_{\mathbf{u}}' = \frac{1}{2}\begin{pmatrix} d_{11} + d_{22} & d_{11} - d_{22} \\ d_{11} - d_{22} & d_{11} + d_{22} \end{pmatrix}.$$

Observe that $\mathbf{\Sigma}_{\mathcal{R}_{\mathbf{u}}\mathbf{X}}$ is symmetric, with the same diagonal elements, which implies that for any permutation matrix $\mathbf{P}$ we have that $\mathbf{P}'\mathbf{\Sigma}_{\mathcal{R}_{\mathbf{u}}\mathbf{X}}\mathbf{P} = \mathbf{\Sigma}_{\mathcal{R}_{\mathbf{u}}\mathbf{X}}$. It is easily seen that $\mathbf{P}'\mathcal{R}_{\mathbf{u}}\mathbf{X}$ is elliptically distributed. Hence, its characteristic function will be given by

$$E[\exp(i\mathbf{t}'X)] = \phi\left(\mathbf{t}'\mathbf{P}'\mathbf{\Sigma}_{\mathcal{R}_{\mathbf{u}}\mathbf{X}}\mathbf{P}\mathbf{t}\right) = \phi\left(\mathbf{t}'\mathbf{\Sigma}_{\mathcal{R}_{\mathbf{u}}\mathbf{X}}\mathbf{t}\right), \quad \mathbf{t} = (t_1, t_2)'. \tag{3.5.2}$$

Since the characteristic function determines the distribution, then by (3.5.2) we have that $\mathcal{R}_{\mathbf{u}}\mathbf{X} =_{st} \mathbf{P}\mathcal{R}_{\mathbf{u}}\mathbf{X}$, and as a consequence, the vector $\mathcal{R}_{\mathbf{u}}\mathbf{X}$ has exchangeable components. $\square$

The next statement describes conditions to compare in the concave order two different portfolios of elliptically distributed risky assets, with possibility of negative weights.

Theorem 3.5.3 *Let $\mathbf{X} = (X_1, X_2)$ be a elliptically distributed random vector with parameter $\mathbf{\Sigma}$ and vector of means $\mu = \mathbf{0}$. Let $\mathbf{u} = (u_1, u_2)'$ be an eigenvector of $\mathbf{\Sigma}$. If (α_1, α_2) and (β_1, β_2) are such that $\alpha_1 + \alpha_2 = \beta_1 + \beta_2$ and $\alpha_1 \leq \beta_1$, then*

$$\begin{aligned}
&\frac{\sqrt{2}}{2}\left[\alpha_1(u_1+u_2)+\alpha_2(u_2-u_1)\right]X_1+\frac{\sqrt{2}}{2}\left[(\alpha_1(u_1-u_2)+\alpha_2(u_1+u_2)\right]X_2\\
&\geq_{cv}\frac{\sqrt{2}}{2}\left[\beta_1(u_1+u_2)+\beta_2(u_2-u_1)\right]X_1+\frac{\sqrt{2}}{2}\left[(\beta_1(u_1-u_2)+\beta_2(u_1+u_2)\right]X_2.
\end{aligned}$$

Proof. As $\mathbf{u} = (u_1, u_2)'$ is a eigenvector of $\mathbf{\Sigma}$, we deduce from Property 3.5.2 that

$$\mathcal{R}_{\mathbf{u}}\mathbf{X} = \frac{\sqrt{2}}{2}\begin{pmatrix} u_1+u_2 & u_2-u_1 \\ u_1-u_2 & u_1+u_2 \end{pmatrix}\begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$$

is exchangeable. Applying Theorem 3.A.35 in Shaked and Shanthikumar [55] we conclude the proof. $\square$

The following Property is an extension of the Property 3.5.2 for high-dimension

Property 3.5.4 *Let $\mathbf{X} = (X_1, \ldots, X_n)'$ be a random vector elliptically distributed with parameters $\mu_{\mathbf{X}} = 0$ and $\mathbf{\Sigma}_{\mathbf{X}}$ is such that it has at least $n-1$ equal eigenvalues given by $\lambda_1 \geq \lambda_2 = \cdots = \lambda_n = \lambda > 0$. Then there exists a unit vector $\mathbf{u}$ such that $\mathcal{R}_{\mathbf{u}}\mathbf{X}$ has exchangeable components .*

Proof. Our proof starts with the observation that the singular values decomposition (SVD) of $\mathbf{\Sigma}_{\mathbf{X}}$ is given by $\mathbf{\Sigma}_{\mathbf{X}} = QDQ'$, where $D = diag\{\lambda_1, \lambda, \ldots, \lambda\}$ and $Q = (q_{ij})$ is an orthogonal matrix. Consider the unit vector $\mathbf{u} = Q^{(c.1)} = [q_{11}, q_{21}, \ldots, q_{n1}]'$. From (3.3.1) we have that $\mathcal{R}_{\mathbf{u}}Q^{(c.1)} = \frac{1}{\sqrt{n}}[1, \ldots, 1]'$. It is easy to check that

$$\mathbf{\Sigma}_{\mathcal{R}_{\mathbf{u}}\mathbf{X}} = \mathcal{R}_{\mathbf{u}}QDQ'\mathcal{R}_{\mathbf{u}}' = \left(\mathcal{R}_{\mathbf{u}}Q\sqrt{D}\right)\left(\mathcal{R}_{\mathbf{u}}Q\sqrt{D}\right)' \tag{3.5.3}$$

Taking $H = \mathcal{R}_{\mathbf{u}}Q$ we can rewrite (3.5.3) as $(H\sqrt{D})(H\sqrt{D})'$. Of course, H is an orthogonal matrix whose first column is $\frac{1}{\sqrt{n}}[1, \ldots, 1]'$.

Let σ^*_{ij} be the element (ij) of the matrix (3.5.3). Hence,

$$\begin{aligned}\sigma^*_{ij} &= \frac{\lambda_1}{n} + \lambda \sum_{j=2}^{n} h^2_{ij} = \frac{\lambda_1}{n} + \lambda \left(1 - \frac{1}{n}\right) = \frac{\lambda_1 + \lambda(n-1)}{n} \\ &= \frac{tr\,(\mathbf{\Sigma_X})}{n} = \frac{1}{n}\sum_{i=1}^{n} var(X_i), \quad if \quad i = j. \\ \sigma^*_{ij} &= \frac{\lambda_1}{n} + \lambda \sum_{k=2}^{n} h_{ik}h_{jk} = \frac{\lambda_1}{n} + \lambda \left(-\frac{1}{n}\right) = \frac{\lambda_1 - \lambda}{n}, \quad if \quad i \neq j.\end{aligned}$$

It is clear that the diagonal of $\mathbf{\Sigma}_{\mathcal{R}_\mathbf{u}\mathbf{X}}$ has the same elements and the off-diagonal also have the same elements. It follows that $P'\mathbf{\Sigma}_{\mathcal{R}_\mathbf{u}\mathbf{X}}P = \mathbf{\Sigma}_{\mathcal{R}_\mathbf{u}\mathbf{X}}$ for any $(n \times n)$ permutation matrix $\mathbf{P}$. It is easily seen that $\mathbf{P}'\mathcal{R}_\mathbf{u}\mathbf{X}$ is elliptically distributed (see, e.g., Valdez and Dhaene [62], Theorem 2). Hence, its characteristic function will be given by

$$E[\exp(i\mathbf{t}'X)] = \phi\left(\mathbf{t}'\mathbf{P}'\mathbf{\Sigma}_{\mathcal{R}_\mathbf{u}\mathbf{X}}\mathbf{P}\mathbf{t}\right) = \phi\left(\mathbf{t}'\mathbf{\Sigma}_{\mathcal{R}_\mathbf{u}\mathbf{X}}\mathbf{t}\right), \quad \mathbf{t} = (t_1, \ldots, t_n)'. \tag{3.5.4}$$

By the one-to-one correspondence between distribution functions and characteristic functions, and from (3.5.4), $\mathcal{R}_\mathbf{u}\mathbf{X} =_{st} \mathbf{P}\mathcal{R}_\mathbf{u}\mathbf{X}$, and as a consequence the vector $\mathcal{R}_\mathbf{u}\mathbf{X}$, has exchangeable components. □

Remark 3.5.5 *Property 3.5.4 is also valid when* $\mu_\mathbf{X} = kQ^{(c.1)}$ *for some* $k \in \mathbb{R}$ *since* $\mathcal{R}_\mathbf{u}\mu_\mathbf{X} = k\mathcal{R}_\mathbf{u}Q^{(c.1)} = \frac{k}{\sqrt{n}}[1, \ldots, 1]'$. *Therefore,* $\mathcal{R}_\mathbf{u}\mathbf{X}$ *has exchangeable components. In Theorem 3.5.4 obviously* $k = 0$.

We are thus led to the following strengthening of Theorem 3.5.3

Theorem 3.5.6 *Let* $\mathbf{X} = (X_1, \ldots, X_n)'$ *satisfy the hypotheses of the Property 3.5.4. Suppose that the (SVD) of* $\mathbf{\Sigma_X}$ *is given by* $\mathbf{\Sigma_X} = QDQ'$ *and let* $\mathbf{u} = Q^{(c.1)}$ *be the unit vector. If* $\mathbf{a} = (a_1, \ldots, a_n)'$ *is majorized by* $\mathbf{b} = (b_1, \ldots, b_n)'$*, then*

$$\sum_{i=1}^{n} \mathbf{a}'\mathcal{R}^{(c.i)}_\mathbf{u} X_i \geq_{cv} \sum_{i=1}^{n} \mathbf{b}'\mathcal{R}^{(c.i)}_\mathbf{u} X_i. \tag{3.5.5}$$

Proof. We conclude from Property 3.5.4 that $\mathcal{R}_\mathbf{u}\mathbf{X}$ has exchangeable components; hence the assertion follows by Theorem 3.A.35 in Shaked and Shanthikumar [55]. □

3.6 Conclusions

We have introduced in this paper a generalization of the upper and lower orthant orders. This new stochastic order allows for comparisons of random vectors in different directions. We also have given some properties and their relationships with other stochastic orders studied in the literature as necessary condition as sufficient conditions and we have given new conditions for comparing in the concave order linear combinations of random variables non necessary iid.

From applications point of view, we consider the single period portfolio problem of allocating the wealth to n risks. Some solutions to this problem are given when two random vectors are comparable in extremality order sense. In the special case of risks elliptically distributed, we have studied directions to rotate the distributions and finding easily the optimal allocations of the wealth in order to maximize the expected utility of a risk averse decision maker.

For the case of random variables elliptically distributed with mean zero, we showed that always is possible to find a rotation where the rotated distribution has exchangeable components in dimension two. We also can find the linear combinations of the random variables that can improve an utility function. For greater dimensions we studied the conditions under which the distribution can be rotated so that it has exchangeable component. The results of this Chapter are based on Laniado et al. [36]

Recall that a function $\phi : [0, \infty) \longrightarrow \mathbb{R}$ is said to be completely monotone if all its derivatives $\phi^{(n)}(x)$ exist and satisfy $(-1)^n \phi^{(n)}(x) \geq 0$, for all $x \geq 0$ and $n = 0, 1, 2 \ldots$. It is well known that X is said to be smaller than Y in Laplace transform order if, and only if,

$$E\left[\phi(X)\right] \geq E\left[\phi(Y)\right],$$

for all completely monotone function ϕ, provided expectation exist (see Shaked and Shanthikumar [55], pages 234 and 235). Therefore, there is a direction in which this research might be continued. For example, Theorem 3.3.16 in this Chapter establishes an interesting relationship between extremality stochastic order and the Laplace transform order. Hence, it would be interesting to study the consistency (preservation) of the extremality order in portfolios when the economic agent has a Constant Absolute Risk Aversion (CARA) or a Constant Relative Risk Aversion (CRRA) utility function since in those both cases, the utility function satisfies similar conditions to the completely monotone function.

4

Portfolio selection based on multivariate extremality order

4.1 Introduction

In this Chapter, we propose a strategy for selecting portfolios based on the extremality order introduced in Chapter 2. Instead of the usual optimization techniques, our approach compares portfolios by means of the ordering provided by the extremality measure when a direction is chosen by the economic agent. Before introducing our methodology, we provide some preliminaries on the classic problem of portfolio selection. Consider the general portfolio optimization problem

$$\min_{\mathbf{w}} \left[\hat{\rho}(R\mathbf{w}) - \frac{1}{\gamma} g(R\mathbf{w})\right], \quad s.t. \quad \sum_{i=1}^{n} w_i = 1, \tag{4.1.1}$$

where $\mathbf{w} = (w_1, \ldots, w_n)'$ is the vector of portfolio weights and R is a $m \times n$ data matrix being m the number of returns and n the number of assets. $\hat{\rho}$ is a risk measure based on the data; for instance, a risk statistic or a natural risk statistic (Heyde et al. [26] and Ahmed et al. [1]); $g : \mathbb{R}^m \longrightarrow \mathbb{R}$ quantifies the returns and γ is the risk-aversion parameter. When $\hat{\rho}$ is the variance and $g(\mathbf{x}) = \frac{\mathbf{x}'\mathbf{1}_m}{m}$, we have the classical mean-variance portfolio model discussed in Markowitz [43]. In this case, the problem (4.1.1) becomes

$$\min_{\mathbf{w}} \left[\mathbf{w}'\hat{\Sigma}\mathbf{w} - \frac{1}{\gamma}\hat{\mu}\mathbf{w}\right], \quad s.t. \quad \sum_{i=1}^{n} w_i = 1. \tag{4.1.2}$$

The model proposed in Markowitz [43] is relevant in modern portfolio theory where the main goal is to maximize return and minimize risk. Its philosophy is that an investor's decision regarding portfolio weights is based on the trade-off between expected return and risk. Markowitz [43] showed that an investor should hold a portfolio that sits at 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, and contains portfolios for which one cannot improve risk and return at the same time. As $\gamma \longrightarrow \infty$, the problem (4.1.2) corresponds to the minimum-variance portfolio and which has also been the subject of recent academic research. In this case, only the covariances need to be estimated and the model is thus, less vulnerable to estimation error than models with a finite risk-aversion parameter. The problem of estimating population moments by sample moments is widely explained in DeMiguel et al. [11]. Note that different values of γ may derive in different mean-variance portfolios on the efficient frontier.

In the hyperplane $\mathbf{w}'\mathbf{1} = 1$, each $\mathbf{w}$ generates a pair $(\hat{\rho}(R\mathbf{w}), g(R\mathbf{w}))$ of feasible portfolios. Figure 4.1 presents different possible linear combinations of the assets in the risk-return space. We see that the efficient frontier is given by 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. Also observe in Figure 4.1 that portfolios in the A square are more attractive than portfolios in the B square because they have higher returns with less risk.

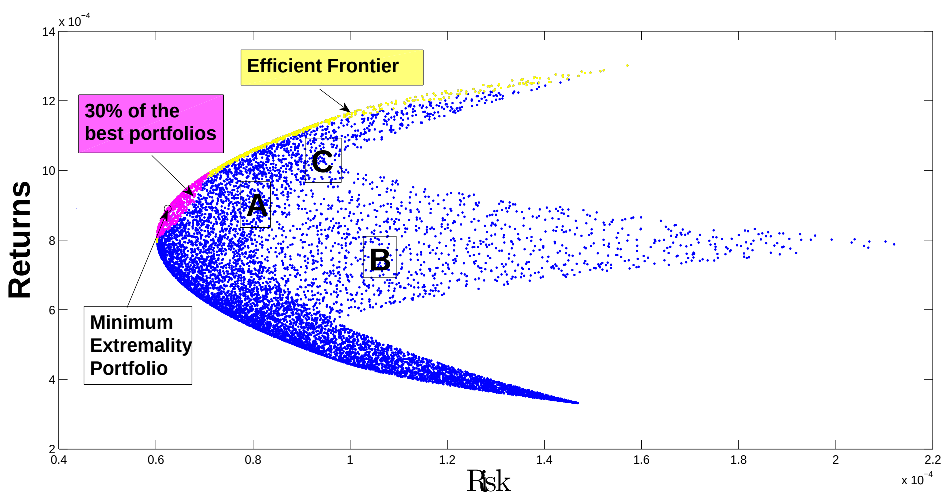


Figure 4.1: Feasible portfolios and efficient frontier

Similarly, portfolios in C are more attractive than portfolios in B, but those in A and C are not comparable in terms of return and risk, simultaneously. The proposal made in this chapter enables us to avoid this problem since it permits a total comparison between feasible portfolios. With our approach, portfolios in A can be compared with portfolios in C.

Many criteria for portfolio selection have been proposed in the literature (see DeMiguel et al. [12] and references therein). In general, the simpler strategy is to choose $\mathbf{w} = \frac{1}{n}\mathbf{1}_n$, (from now on called $\frac{1}{n}$-rule), which allocates the same proportion of wealth across each of the n assets. Almost all models can be expressed as (4.1.1) and in this case, the solution for the best portfolio will be that on the efficient frontier that minimizes its scalar projection on the vector $[1, -\frac{1}{\gamma}]'$, where γ is the risk-aversion parameter. The main difference between models is how the risk is measured and estimated. For example it is usual to consider the sample standard deviation as a risk measure although, other more stable estimators can be used for this purpose. For example, DeMiguel and Nogales [13] propose portfolio policies that are based on robust estimators. Another difference between models is comprised of the constraints on weights, as may be seen in Jagannathan and Ma [27] and DeMiguel et al. [11]) who discuss shortsales and constraining portfolio norms, respectively.

The main purpose of this Chapter is to present a novel methodology for comparing portfolios based on an extremality data order in some direction. Directions are chosen by the investor according to main criteria that he or she uses in selecting the portfolio. Because these criteria can be different from the traditional mean and variance, we also introduce new versions of the efficient frontiers which depend on criteria which are most relevant to the investor.

The structure of Chapter is thus as follows. In Section 4.2, we provide new definitions of alternative efficient frontiers and their graphical representation. Next, in Section 4.3, we present a methodology to select the portfolios through the extremality order defined in Chapter 2. A version of this order will also be provided for ordering portfolios and its main properties. In Section 4.4, we briefly present a methodology for evaluating the performance of the strategies developed in this Chapter and we introduce classic theories from the literature for purposes of comparison. Then, in Section 4.5, we present the performance of the strategies in real data and discuss the results. Finally in Section 4.6, we provide some conclusions and possible extensions of our approach.

To introduce our methodology, we start by introducing alternative definitions of efficient frontiers.

4.2 Alternative efficient frontiers

The Markowitz model of mean-variance is the most common formulation of portfolio selection problems. However, portfolio models that consider criteria other than returns and variances have been widely studied in the recent literature. For example, Rockafellar and Uryasev [52] and Quaranta and Zaffaroni [49] implemented the conditional Value-at-Risk instead of the variance as risk measure; Gaivoronski and Pflug [21] studied Value-at Risk in portfolio optimization and Usta and Kantar [61] proposed a model of mean-variance-skewness for portfolio selection. What is more, Usta and Kantar [61] also maximize entropy in the objective function to generate well-diversified asset portfolio within optimal asset allocation. Transaction costs are also considered in many models.

It is well known that in the portfolio optimization approach proposed by Markowitz, the problem is modeled as a mean-variance optimization problem where the expected return is maximized and the variance is minimized. The performance of the portfolio can be measured with the Sharpe-ratio introduced in Sharpe [56] and defined as the increasing of the portfolio return by unity of risk, commonly expressed as $\frac{\hat{\mu}}{\hat{\sigma}}$ if a risk-free asset is not considered. However, it is interesting to consider that the Markowitz model often leads to portfolios which are highly concentrated on a few assets which stands in contradiction to the notion of diversification. This phenomena also leads to high transaction costs every time that the portfolio is rebalanced. In order to improve diversification, maximization of Shannon's entropy has been accepted as a good criterion of diversity, (Usta and Kantar [61]). This Chapter includes different criteria for choosing the portfolio that depend on investor choice. Any consideration of criteria different from the, classical criteria of mean-variance must therefore contemplate new versions of efficient frontier. We formalize these ideas in the remainder of the Section.

Let Θ be a set of k criteria for evaluating the performance of the portfolio. In the classical Markowitz model $k = 2$ and corresponds to mean and variance of the portfolio. Consider any criterion $c_i \in \Theta$, $i = 1, \ldots, k$ and denote

$$\theta_{c_i} = \begin{cases} 1, & \text{if the investor wants a portfolio with a low value of the criterion } c_i \\ -1, & \text{if the investor wants a portfolio with a high value of the criterion } c_i \end{cases}$$

For example, if we assume $\Theta = \{\text{return}, \text{risk}, \text{Sharpe-ratio}, \text{entropy}\} = \{c_1, c_2, c_3, c_4\}$, clearly

$$\theta_{\text{return}} = \theta_{c_1} = -1, \quad \theta_{\text{risk}} = \theta_{c_2} = 1, \quad \theta_{\text{Sr}} = \theta_{c_3} = -1, \quad \theta_{\text{entropy}} = \theta_{c_4} = -1.$$

As we have mentioned in the Introduction, Markowitz defined the efficient frontier as the set of feasible portfolios which cannot be improved in terms of return and risk simultaneously.

Following the Markowitz idea, we introduce new definitions of efficient frontiers which are determined by the criteria $\{c_1, \ldots, c_k\}$ belong to Θ. We define an efficient frontier as the set of feasible portfolios which cannot be improved in terms of all the criteria from Θ, simultaneously.

Let Ω be the set of possible weights for a collection of n assets and let $\mathcal{P}$ be given by,

$$\begin{aligned} \mathcal{P} : \Omega &\longrightarrow \mathbb{R}^k \\ \mathbf{w} &\longrightarrow \mathcal{P}_{\mathbf{w}} = [c_1(\mathbf{w}), \ldots, c_k(\mathbf{w})]'. \end{aligned}$$

where $\mathcal{P}_{\mathbf{w}}$ denotes the vector of criteria considered in Θ evaluated in the weight vector $\mathbf{w} = [w_1, \ldots, w_n]'$. Therefore, we can define the $\Theta-$Efficient Frontier as follows

Definition 4.2.1 ($\Theta-$Efficient Frontier) *Consider the unit vector* $\mathbf{u} = \frac{1}{\sqrt{k}}[\theta_{c_1}, \ldots, \theta_{c_k}]'$, $c_i, \in \Theta$. *Let* $\mathcal{S} = \mathcal{P}(\Omega) \subset \mathbb{R}^k$ *be the set of possible values of* $\mathcal{P}$. *The* $\Theta-$*Efficient Frontier is the set given by*

$$\partial\Theta = \left\{ \mathcal{P}_{\mathbf{w}} \in \mathcal{S} : \text{ there is not another } \mathcal{P}_{\mathbf{w}'} \in \mathcal{S} \text{ such that } \mathcal{C}^{\mathbf{u}}_{\mathcal{P}_{\mathbf{w}}} \subset \mathcal{C}^{\mathbf{u}}_{\mathcal{P}_{\mathbf{w}'}} \right\}. \tag{4.2.1}$$

Recall that $\mathcal{C}^{\mathbf{u}}_{\mathcal{P}_{\mathbf{w}}}$ is an oriented orthant (see Definition 2.2.3) with vertex in $\mathcal{P}_{\mathbf{w}}$ and oriented by $\mathbf{u}$, i.e,

$$\mathcal{C}^{\mathbf{u}}_{\mathcal{P}_{\mathbf{w}}} = \{\mathcal{P}_{\mathbf{w}'} \in \mathcal{S} \mid \mathcal{R}_{\mathbf{u}}(\mathcal{P}_{\mathbf{w}'} - \mathcal{P}_{\mathbf{w}}) \geq 0\}, \tag{4.2.2}$$

where $\mathcal{R}_{\mathbf{u}}$ is the known orthogonal rotation matrix from (3.3.1) such that

$$\mathcal{R}_{\mathbf{u}}\mathbf{u} = \frac{1}{\sqrt{n}}\,\mathbf{1}. \tag{4.2.3}$$

Consider the set $\{\text{return, risk, Sharpe-ratio, entropy}\} = \{c_1, c_2, c_3, c_4\}$ of criteria for choosing some portfolio weights. If we fix $k = 2$ then six possible Θ'sets given by $\Theta_{ij} = \{c_i, c_j\}$, for $i \neq j$ can be considered. Hence, according to Definition 4.2.1, we have six different efficient frontiers which some points on these are displayed in Figure 4.2. In order to illustrate those efficient frontiers, we have used the 5Ind data set of monthly assets in the period 07/1963−12/2004 from Kenneth French web-site.

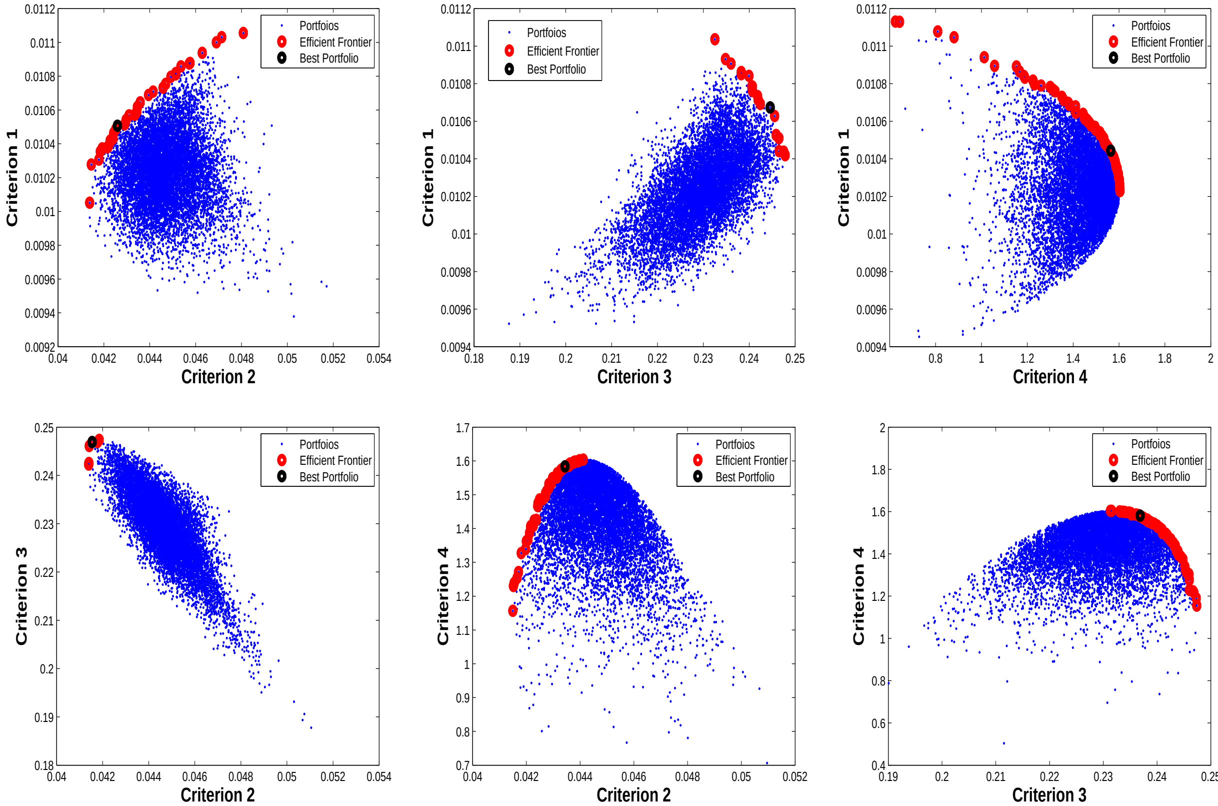


Figure 4.2: Points on the efficient frontier

In this case, we have considered Ω as the hyperplane $\{(w_1, \ldots, w_5) : \sum w_i = 1, \quad w_i \geq 0\}$ and we have simulated 10000 elements from previous set Ω via 10000 copies drawn from the random vector with a uniform distribution of independent marginals.

$$\mathbf{U} = \left(\sum_{i=1}^{5} U_i \right)^{-1} [U_1, \ldots, U_5]' \quad \text{where} \quad U_i \sim Uniform(0,1), \;\; \text{for all} \;\; i = 1, \ldots, 5. \tag{4.2.4}$$

Observe that the red points can not be improved for other points in terms of the respective criteria c_i, c_j, simultaneously. Hence, we say that the red points belong to the $\Theta_{ij}-$Efficient Frontier. In the previous cases the efficient frontiers can be determined by a curve, however our procedure allows us to build efficient frontiers when the investor wants to include more than two criteria to take decisions. In the cases when Θ' set is such that $k > 2$, the efficient frontiers will be determined by a surface in k dimension. For example, Figure 4.3 shows the efficient frontier when $k = 3$.

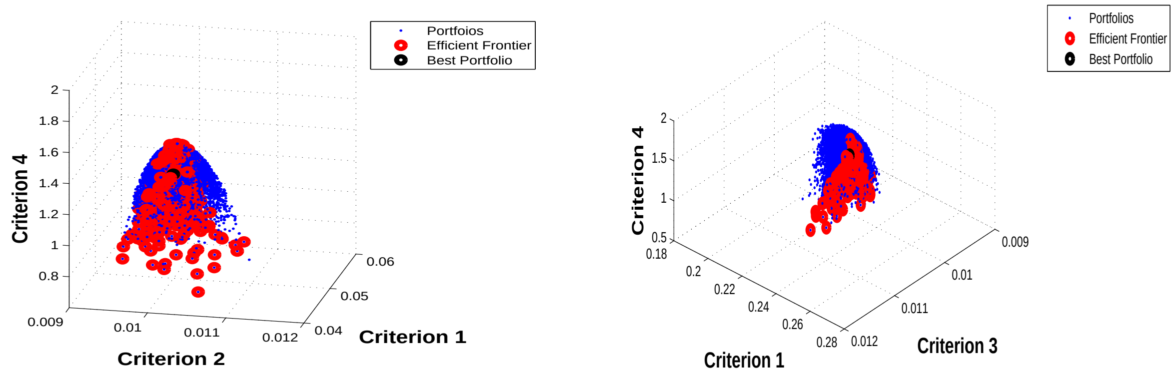


Figure 4.3: Points on the efficient frontier

Although Figure 4.2 is clearer than Figure 4.3, in both Figures it can be seen that the red points cannot be improved in terms of all the criteria from Θ, simultaneously.

Note that in all Figures displayed in this Section, we show the best portfolio (black point) in the same Figures which is obtained through the results given in the following Section.

4.3 Portfolio selection based on the extremality order

Chapter 2 proposes a methodology to sort multivariate data based on a new concept called extremality. This new concept induces an order which can be seen as a generalization of the usual componentwise order over a fixed direction $\mathbf{u}$. This multivariate order has also been developed in Laniado et al. [37]. In this Section we propose to sort feasible portfolios according to the order induced by a direction $\mathbf{u}$ which is determined by the criteria chosen for selecting the portfolio. Following the extremality order, our approach is to choose the smallest portfolio as the best portfolio instead of those traditional portfolios selected through optimization techniques. It should be noted that the advantage of our strategy depends on what the economic agent considers to be most interest when choosing a portfolio. What is more, we have not used optimization techniques since some criteria do not have the appropriate properties for optimization, for example, it is well known that the Value at Risk is not a convex risk measure and that its optimization process is complex. In addition, if the risk is measured with the variance, there is the problem of estimating a large number of parameters in the covariance matrix and as a consequence, the estimation errors will have a larger impact on portfolio weights when they come from some optimization process.

We rank a sample of feasible portfolio using the order introduced in Definition 2.3.2 where we say that the best portfolio is that with smallest extremality value. The feasible portfolios are built using Monte Carlo simulation. Our methodology consists in selecting a portfolio on each $\Theta-$Efficient Frontier and which is displayed as a black point in Figures 4.2. The portfolio is obtained by using the order induced by extremality studied in Chapter 2 and whose version to sort feasible portfolios is given as follows.

Definition 4.3.1 *Let* $\mathbf{w}$ *and* $\mathbf{w}' \in \Omega$ *and* $\mathbf{u} = \frac{1}{\sqrt{k}}[\theta_{c_1}, \ldots, \theta_{c_k}]'$. *We say that* $\mathbf{w}$ *dominates* $\mathbf{w}'$ *if, and only if,* $\mathcal{P}_{\mathbf{w}} \leq_{\mathcal{E}_{\mathbf{u}}} \mathcal{P}_{\mathbf{w}'}$.

Observe that the dominance of $\mathbf{w}$ on $\mathbf{w}'$ can also be written as

$$\mathcal{P}_{\mathbf{w}} \leq_{\mathcal{E}_{\mathbf{u}}} \mathcal{P}_{\mathbf{w}'} \equiv P\left(\mathcal{C}^{\mathbf{u}}_{\mathcal{P}_{\mathbf{w}}}\right) \geq P\left(\mathcal{C}^{\mathbf{u}}_{\mathcal{P}_{\mathbf{w}'}}\right). \tag{4.3.1}$$

Now, we give some details of what is meant by the term dominance in Definition 4.3.1. Let $\mathbf{w_1}, \mathbf{w_2} \in \Omega$ such that $\mathbf{w_1}$ dominates to $\mathbf{w_2}$ according to Definition 4.3.1, that is, from (4.3.1), we can affirm

$$P\left(\mathcal{C}^{\mathbf{u}}_{\mathcal{P}_{\mathbf{w_1}}}\right) \geq P\left(\mathcal{C}^{\mathbf{u}}_{\mathcal{P}_{\mathbf{w_2}}}\right) \quad \text{implying by (4.2.2) that,}$$

for any $\mathbf{w} \in \Omega$ randomly chosen

$$P\left[\mathcal{R}_{\mathbf{u}}(\mathcal{P}_{\mathbf{w}} - \mathcal{P}_{\mathbf{w_1}}) \geq 0\right] \geq P\left[\mathcal{R}_{\mathbf{u}}(\mathcal{P}_{\mathbf{w}} - \mathcal{P}_{\mathbf{w_2}}) \geq 0\right]. \tag{4.3.2}$$

Observe that if we consider the criteria c_1, c_2 as in the classical risk-return model, $\mathcal{P}_{\mathbf{w}}$, $\mathbf{w} \in \Omega$ represents a couple $(risk_{\mathbf{w}}, return_{\mathbf{w}}) \in \mathcal{P}(\Omega)$. Obviously, in this case $\mathbf{u} = \frac{1}{\sqrt{2}}[1, -1]'$ and the respective orthogonal rotation matrix (stated in (4.2.3)) will be given by

$$\mathcal{R}_{\mathbf{u}} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \text{ since } \mathcal{R}_{\mathbf{u}}\mathbf{u} = \frac{1}{\sqrt{2}}[1, 1]'.$$

Hence, the inequality (4.3.2) can be written as

$$P\left[\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} risk_{\mathbf{w}} - risk_{\mathbf{w_1}} \\ return_{\mathbf{w}} - return_{\mathbf{w_2}} \end{pmatrix} \geq \begin{pmatrix} 0 \\ 0 \end{pmatrix}\right] \geq$$
$$P\left[\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} risk_{\mathbf{w}} - risk_{\mathbf{w_2}} \\ return_{\mathbf{w}} - return_{\mathbf{w_2}} \end{pmatrix} \geq \begin{pmatrix} 0 \\ 0 \end{pmatrix}\right],$$

equivalent to

$$P\left[\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} risk_{\mathbf{w}} \\ return_{\mathbf{w}} \end{pmatrix} \geq \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} risk_{\mathbf{w_1}} \\ return_{\mathbf{w_1}} \end{pmatrix}\right] \geq$$
$$P\left[\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} risk_{\mathbf{w}} \\ return_{\mathbf{w}} \end{pmatrix} \geq \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} risk_{\mathbf{w_2}} \\ return_{\mathbf{w_2}} \end{pmatrix}\right].$$

Therefore, it is easily seen that

$$P\left[risk_{\mathbf{w}} \geq risk_{\mathbf{w_1}} \text{ and } return_{\mathbf{w}} \leq return_{\mathbf{w_1}}\right]$$
$$\geq P\left[risk_{\mathbf{w}} \geq risk_{\mathbf{w_2}} \text{ and } return_{\mathbf{w}} \leq return_{\mathbf{w_2}}\right],$$

which means that the probability of finding a worse portfolio, in terms of risk-return, is greater for $\mathcal{P}_{\mathbf{w_1}}$ than $\mathcal{P}_{\mathbf{w_2}}$ and this is what is meant by the dominance of $\mathbf{w_1}$ on $\mathbf{w_2}$. Hence, our strategy for selecting portfolios consists in choosing the more dominant portfolio, i.e, the smallest in terms of extremality. In Figure 4.2 it can be observed that the best portfolio is drawn as the black point on each Θ-efficient frontier. Figure 4.1 also shows the best portfolio obtained through extremality order, and in fact the same Figure also shows 30% of the best portfolios. We emphasize again in Figure 4.1 that with our methodology it is possible to compare portfolios in the A-square with those from the C-square.

Consider $\mathbf{u} = \frac{1}{\sqrt{k}}[\theta_{c_1}, \ldots, \theta_{c_k}]'$, the following property indicates that a portfolio which is a vertex of some oriented orthant $\mathcal{C}^{\mathbf{u}}_{\mathcal{P}_{\mathbf{w}}}$ has a better performance when compared to another portfolio that belongs to that orthant.

Property 4.3.2 *Let* $\mathbf{w}$ *and* $\mathbf{w}' \in \Omega$ *and* $\mathbf{u} = \frac{1}{\sqrt{k}}[\theta_{c_1},, \ldots, \theta_{c_k}]'$. *Then*

$$\mathcal{P}_{\mathbf{w}} \leq_{\mathcal{E}_{\mathbf{u}}} \mathcal{P}_{\mathbf{w}'}, \quad \textit{for all } \mathcal{P}_{\mathbf{w}'} \in \mathcal{C}^{\mathbf{u}}_{\mathcal{P}_{\mathbf{w}}}.$$

Proof. Suppose that $\mathcal{P}_{\mathbf{w}'} \in \mathcal{C}_{\mathcal{P}^{\mathbf{u}}_{\mathbf{w}}}$ which implies that, $\mathcal{C}^{\mathbf{u}}_{\mathcal{P}_{\mathbf{w}'}} \subset \mathcal{C}^{\mathbf{u}}_{\mathcal{P}_{\mathbf{w}}}$. Therefore,

$$P_F\left(\mathcal{C}^{\mathbf{u}}_{\mathcal{P}_{\mathbf{w}'}}\right) \leq P_F\left(\mathcal{C}^{\mathbf{u}}_{\mathcal{P}_{\mathbf{w}}}\right) \Longrightarrow \mathcal{P}_{\mathbf{w}} \leq_{\mathcal{E}_{\mathbf{u}}} \mathcal{P}_{\mathbf{w}'}.$$

□

The order of the criteria can affect the position of the Θ-Efficient frontier, however the extremality portfolio order given in the Definition 4.3.1 is invariant to the order selected for the criteria. This is demonstrated in the following property.

Property 4.3.3 *Let* Π *be a permutation matrix. If* $\mathcal{P}_{\mathbf{w}} \leq_{\mathcal{E}_{\mathbf{u}}} \mathcal{P}_{\mathbf{w}'}$, *then* $\Pi\mathcal{P}_{\mathbf{w}} \leq_{\mathcal{E}_{\Pi\mathbf{u}}} \Pi\mathcal{P}_{\mathbf{w}'}$.

Proof. The assertion follows on immediately from the invariance of the extremality under orthogonal transformations. (see Property 2.4.3). □

The following Proposition shows that our strategy can be linked with the Markowitz solution since the portfolio selected under criteria Θ belongs to the Θ-efficient frontier.

Proposition 4.3.4 *If $\mathcal{P}_{\mathbf{w}} \leq_{\mathcal{E}_{\mathbf{u}}} \mathcal{P}_{\mathbf{w}'}$, for all $\mathcal{P}_{\mathbf{w}'} \in \mathcal{S}$, then $\mathcal{P}_{\mathbf{w}}$ belongs to the Θ-Efficient Frontier.*

Proof. Let $\Theta = \{c_1, \ldots, c_k\}$ be the criteria used by the investor for selecting his\her portfolio. Therefore, we already know that the unit vector that determines the direction for ordering the feasible portfolios is given by $\mathbf{u} = \frac{1}{\sqrt{k}}[\theta_{c_1}, \ldots, \theta_{c_k}]'$. Now, suppose that $\mathcal{P}_{\mathbf{w}}$ does not belong to the $\Theta-$Efficient Frontier. From (4.2.1) there is a different portfolio $\mathcal{P}_{\mathbf{v}} = [c_1(\mathbf{v}), \ldots, c_k(\mathbf{v})]' \in \mathcal{S}$, such that $\mathcal{C}^{\mathbf{u}}_{\mathcal{P}_{\mathbf{w}}} \subset \mathcal{C}^{\mathbf{u}}_{\mathcal{P}_{\mathbf{v}}}$ implying that $\mathcal{P}_{\mathbf{w}} \in \mathcal{C}^{\mathbf{u}}_{\mathcal{P}_{\mathbf{v}}}$ which means from Property 4.3.2 that $\mathcal{P}_{\mathbf{v}}$ has better performance than $\mathcal{P}_{\mathbf{w}}$, i.e, $\mathcal{P}_{\mathbf{v}} <_{\mathcal{E}_{\mathbf{u}}} \mathcal{P}_{\mathbf{w}}$ which contradicts the hypothesis. □

DeMiguel et al. [12] show the advantages of ignoring the data on asset returns and using the $\frac{1}{n}$ rule to allocate an equal proportion of wealth for each one of n assets whose performance in terms of Sharpe ratio and transaction cost is, in general, very well accepted among investors. We also propose another strategy that considers the standard deviation vector of data as a director vector. When the risk is measured as the variance, we prove in the 2-dimensional case that the portfolio selected with this strategy has less risk than $\frac{1}{n}$ rule. However, the computational performance of this strategy for higher dimensions reveals that this result could hold for $n > 2$. This strategy will be denoted as P_R and is introduced as follows.

Consider $\mathbf{X} = (X_1, \ldots, X_n)'$ a random vector of risky assets such that $Var(X_i) < \infty$, for $i = 1, \ldots, n$. Let $\sigma = (\sigma_1, \ldots, \sigma_n)'$ be the standard deviation vector of the n random variables. Let $\mathbf{u} = \frac{\sigma}{\|\sigma\|}$ be a unit vector in the direction of the standard deviations. Let $\mathcal{R}_{\mathbf{u}}$ be as in (4.2.3), so $\mathcal{R}_{\mathbf{u}}\frac{\sigma}{\|\sigma\|} = \frac{1}{\sqrt{n}}\mathbf{1_n}$, i.e., in this case $\mathcal{R}_{\mathbf{u}}$ is the matrix becoming the standard deviation vector in a homogeneous vector with norm unitary. Here we propose a portfolio whose weights are given by

$$\mathbf{w} = \frac{\mathcal{R}_{\mathbf{u}}\mathbf{1_n}}{\mathbf{1_n'}\mathcal{R}_{\mathbf{u}}\mathbf{1_n}}, \quad \text{where} \quad \mathbf{u} = \frac{\sigma}{\| \sigma \|}. \tag{4.3.3}$$

The main idea of the strategy (4.3.3) is to ensure that the riskiest asset receives the smallest proportion of the budget available for investment. The goal with this strategy is basically to improve the $\frac{1}{n}$-rule. We will show show that P_R works better in terms of risk, and almost as well in terms of turnover, as the the $\frac{1}{n}$-rule. The following Proposition shows that if the components of $\mathbf{X}$ have the same variances, the strategy P_R is equivalent to $\frac{1}{n}$-rule.

Proposition 4.3.5 *Let $\mathbf{X} = (X_1, \ldots, X_n)'$ be a random vector of risky assets such that $Var(X_i) = c$, for $i = 1, \ldots, n$. Then the portfolio weights given in (4.3.3) are equivalents to $\frac{1}{n}$.*

Proof. Since the variances are the same, it is easily seen that $\mathbf{u} = \frac{\sigma}{\|\sigma\|} = \frac{1}{\sqrt{n}}\mathbf{1_n}$. Clearly from (4.2.3) the matrix $\mathcal{R}_\mathbf{u}$ has to be the identical matrix. Hence, from (4.3.3)

$$\mathbf{w} = \frac{\mathcal{R}_\mathbf{u}\mathbf{1_n}}{\mathbf{1_n'}\mathcal{R}_\mathbf{u}\mathbf{1_n}} = \frac{\mathbf{1_n}}{\mathbf{1_n'}\mathbf{1_n}} = \frac{1}{n}\mathbf{1_n}.$$

$\square$

Some special cases where $\mathbf{X}$ satisfies the conditions as in Proposition 4.3.5 are when $X_1, \ldots, X_n$ i.i.d. or $\mathbf{X}$ a random vector with exchangeable components.

Assuming that the risk of the portfolio is measured through its variance, we show below that for the 2-dimensional case the $\frac{1}{n}$-rule is risker than the portfolio P_R. This is stated in the following Property.

Property 4.3.6 *Let $\mathbf{X} = (X_1, X_2)'$ be random vector with finite margin variance σ_1^2 and σ_2^2, respectively and σ_{12} the covariance. Consider $\mathbf{w_1}$ as in (4.3.3) and $\mathbf{w_2} = (\frac{1}{2}, \frac{1}{2})'$. Then*

$$Var\left(\mathbf{w_1'}\mathbf{X}\right) \leq Var\left(\mathbf{w_2'}\mathbf{X}\right).$$

Proof. Since the unit vector $\mathbf{u} = \frac{1}{\|\sigma\|}(\sigma_1, \sigma_2)'$, the orthogonal rotation matrix stated in (4.2.3) is given by

$$\mathcal{R}_\mathbf{u} = \frac{\sqrt{2}}{2\parallel\sigma\parallel}\begin{pmatrix} \sigma_1+\sigma_2 & \sigma_2-\sigma_1 \\ \sigma_1-\sigma_2 & \sigma_1+\sigma_2 \end{pmatrix}, \text{ so the condition } \mathcal{R}_\mathbf{u}\mathbf{u} = \frac{1}{\sqrt{n}}\mathbf{1_n} \text{ holds.} \tag{4.3.4}$$

Replacing the matrix (4.3.4) in (4.3.3), we obtain that $\mathbf{w_1} = (\frac{\sigma_2}{\sigma_1+\sigma_2}, \frac{\sigma_1}{\sigma_1+\sigma_2})'$. Hence,

$$Var\left(\mathbf{w_1'}\mathbf{X}\right) = Var\left(\frac{\sigma_2}{\sigma_1+\sigma_2}X_1 + \frac{\sigma_1}{\sigma_1+\sigma_2}X_2\right) = \frac{\sigma_1\sigma_2}{(\sigma_1+\sigma_2)^2}\left(2\sigma_1\sigma_2 + 2\sigma_{12}\right).$$

The variance of the $\frac{1}{n}$-rule is given by

$$Var\left(\mathbf{w_2'}\mathbf{X}\right) = Var\left(\frac{1}{2}X_1 + \frac{1}{2}X_2\right) = \frac{1}{4}\left(\sigma_1^2 + \sigma_2^2 + 2\sigma_{12}\right).$$

The proof is completed by the simple fact that

$$\frac{\sigma_1\sigma_2}{(\sigma_1+\sigma_2)^2} \leq \frac{1}{4} \text{ and } (2\sigma_1\sigma_2 + 2\sigma_{12}) \leq \left(\sigma_1^2 + \sigma_2^2 + 2\sigma_{12}\right), \text{ for all } \sigma_1, \sigma_2, \sigma_{12}.$$

□

As we have commented previously, we have not the proof for the case $n > 2$ being this result a conjecture. However, we have computational evidence that it holds.

4.4 Portfolio performance measures

In this Section, we illustrate the approach developed in the previous Section and compare it with other alternatives in the literature of portfolio optimization. To evaluate the performance of the portfolios obtained in this Chapter, we will compare five classic strategies: Mean-variance portfolio with shortsales constrained and risk aversion parameter $\gamma = 1$ (MEAN), Mean-Variance portfolio with shortsales unconstrained (MEANU), Minimum-Variance portfolio with shortsales constrained (MIN), Minimum-Variance portfolio with shortsales unconstrained (MINU) and Equally-weighted Portfolio ($\frac{1}{n}$). The comparison is carried out using out-of-sample portfolio returns, out-of-sample portfolio risk, out-of-sample portfolio Sharpe ratio and portfolio turnovers. We use the technique "rolling-horizon" implemented in DeMiguel et al. [11], which depends on a window τ to perform the estimation. Thus, using the monthly data over the estimation window, we estimate the feasible portfolios set S through the simulation of ten thousand points of the hyperplane $\sum_{i=1}^{n} w_i = 1$ with $w_i \geq 0$ of the same way as in (4.2.4) and we sort the feasible portfolio following the Definition 4.3.1 depending on the criteria selected. In this Chapter we are working with strategies based on a set of two criteria, i.e, $k = 2$. The notation for our strategies is given in Table 4.1. For each estimation window, we compute six previous

Tables 4.1: Portfolios notation

Criteria	returns and variance	returns and Sharpe ratio
Portfolio notation	P_{12}	P_{13}
Criteria	returns and entropy	variance and Sharpe ratio
Portfolio notation	P_{14}	P_{23}
Criteria	variance and entropy	Sharpe ratio and entropy
Portfolio notation	P_{24}	P_{34}

strategies from Table 4.1 and the strategy P_R introduced in (4.3.3). In order to establish comparisons, we also compute the (MEAN), (MEANU), (MIN), (MINU) and $(\frac{1}{n})$ strategies. 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 et al. [11], DeMiguel and Nogales [13] for additional information):

- Out-of-sample portfolio returns

$$\hat{\mu}^i = \frac{1}{m-\tau} \sum_{t=\tau}^{m-1} (w_t^i)' r_{t+1}$$

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 standard deviation

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

- Out-of-sample portfolio Sharpe ratio

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

In order to statistically evaluate the difference in Sharpe ratios, we use the test devised by Memmel [45] , which is a corrected version of the original proposal of Jobson and Korkie [28]. It is illustrated in Usta and Kantar [61] as follows:

Let a and b be two portfolio selection models generating two Sharpe ratios SR_a and SR_b , respectively. Under the null hypothesis that the Sharpe ratios are the same, the test statistic for $SR_a - SR_b$ is asymptotically normally distributed with mean zero and variance ϑ:

$$\vartheta = \frac{1}{m-\tau} \left[2 - 2\rho_{ab} + \frac{1}{2} \left(SR_a^2 + SR_b^2 - 2SR_a SR_b \rho_{ab}^2 \right) \right],$$

where m, τ is the sample size and windows size in the rolling horizon, respectively. ρ_{ab} is the correlation coefficient between portfolio returns obtained from a and b models. Thus, the test statistic for difference in Sharpe ratios is calculated as follows:

$$Z = \frac{SR_a - SR_b}{\sqrt{\vartheta}}.$$

In this study, the p-value corresponding to the previous test statistic will be calculated for each model with respect to the $\frac{1}{n}$-rule, which is taken as a benchmark due to its easy implementation and widespread use. Additionally, in the literature, DeMiguel et al. [12] show that the $\frac{1}{n}$-rule has good behavior in the out-of-sample case.

- Turnover

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

where $w_{j,t}$ is the portfolio weight of the asset j chosen at time t and 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. The Turnover for $\frac{1}{n}$ strategy may be different to zero due to changes in asset prices between t and $t+1$. The Turnover can be interpreted as the average percentage of wealth traded period to period and it is related to transaction costs.

4.5 Empirical study

In this Section we give the descriptions of the empirical data sets used in this work and also present the results of the empirical study.

4.5.1 Data description

Regarding the data sets, we have used portfolios of different sizes. They have been built using a set of 50 companies from Spain and taking monthly returns from 10/2001 to 01/2008. The 50 companies are given in Table 4.2.

With companies from Table 4.2, we build some portfolio sets to evaluate the strategies discussed in this Chapter. These portfolios are given in Table 4.3. We also use other two data sets more. One of them has been taken from Kenneth French web-site:
$http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data\_library.html$,
which is a monthly returns data set of 48 industry portfolios representing the U.S. stock market and which is denoted by **48Ind**. The other dataset consists of monthly eight international indexes, which are taken from yahoo finance web site, NASDAC,

S&P500 from US, FTSE100 DAX, NIKKEI225UK from Europe and HSI, STRAITS-TIMES from Asia for the period from January 2000 to April 2012 and which are denoted by **8Indexes**.

Tables 4.2: Companies from Spain for constructing the portfolios

Code	**1**	**2**	**3**
Company	ACERINOX	ADOLFO DOMINGUEZ	ALTADIS
Code	4	5	6
Company	ACCIONA	AGUAS BARCELONA -A-	AMPER
Code	**7**	**8**	**9**
Company	AZKOYEN	TAVEX ALGODONERA	BARON DE LEY
Code	**10**	**11**	**12**
Company	BAYER	BEFESA	VISCOFAN
Code	**13**	**14**	**15**
Company	BANCO CASTILLA	CEPSA	CAMPOFRIO
Code	**16**	**17**	**18**
Company	DINAMIA	ENDESA	EUROPISTAS
Code	**19**	**20**	**21**
Company	FCC	FUNESPANA	GAS NATURAL
Code	**22**	**23**	**24**
Company	IBERDROLA	IBERPAPEL	IBERIA
Code	**25**	**26**	**27**
Company	INBESOS	INDITEX	LINGOTES ESPECIALES
Code	**28**	**29**	**30**
Company	METROVACESA	NH HOTELES	PRISA
Code	**31**	**32**	**33**
Company	REPSOL YPF	BODEGAS RIOJANAS	BANCO SANTANDER
Code	**34**	**35**	**36**
Company	SOGECABLE	SOL MELIA	TELEFONICA
Code	**37**	**38**	**39**
Company	TUBACEX	ACUMULADOR TUDOR	URBAS
Code	**40**	**41**	**42**
Company	UNION FENOSA	UNIPAPEL	URALITA
Code	**43**	**44**	**45**
Company	VIDRALA	AMERICA MOVIL-L	BBVA
Code	**46**	**47**	**48**
Company	TELMEX	VALE-R-12	VOLCAN-B
Code	**49**	**50**	
Company	ZELTIA	ZARDOYA	

Tables 4.3: Portfolio data sets from Spain

Set	Company code
5Spain	10 – 14 – 17 – 21 – 22
6Spain	9 – 23 – 32 – 39 – 47 – 48
10Spain	9 – 14 – 23 – 24 – 25 – 32 – 34 – 39 – 47 – 48
25Spain	3 – 4 – 6 – 9 – 10 – 13 – 14 – 15 – 19 – 21 – 22 – 23 – 24 25 – 31 – 32 – 34 – 35 – 38 – 39 – 42 – 45 – 47 – 48 – 49
40Spain	2 – 3 – 4 – 5 – 6 – 8 – 9 – 10 – 11 – 12 – 13 – 14 – 15 – 16 18 – 19 – 21 – 22 – 23 – 24 – 25 – 26 – 27 – 28 – 30 – 31 – 32 33 – 34 – 35 – 37 – 38 – 39 – 41 – 42 – 44 – 45 – 47 – 48 – 49

The summary of data considered in this study are given in the following Table

Tables 4.4: Data sets of monthly asset returns.

Data Sets	**Abrev.**	**n**	**Period**
5 companies from Spain see Table 4.3	5Spain	5	10/2001-01/2008
6 companies from Spain see Table 4.3	6Spain	6	10/2001-01/2008
10 companies from Spain see Table 4.3	10Spain	10	10/2001-01/2008
25 companies from Spain see Table 4.3	25Spain	25	10/2001-01/2008
40 companies from Spain see Table 4.3	40Spain	40	10/2001-01/2008
48 industry portfolio representing the U.S. stock market	48Ind	48	07/1963-12/2004
8 international indexes from US, Europe, Asia	8ndexes	8	01/2000-04/2012

These data are monthly asset returns and are presented in the Table 4.4 , with the abbreviation used to refer to each data set, the number of assets in each data set and the time period.

4.5.2 Results of the empirical study

Tables 4.5, 4.6, 4.7 and 4.8 report the out-of-sample criteria for the portfolios considered and the strategies analyzed. The windows size τ in the technique "rolling-horizon" is $\tau = 50$ for data sets from Spain and also for 8Index data set, and $\tau = 300$ for 48Ind data set.

Tables 4.5: Portfolio returns

Strategy	5Spain	6Spain	10Spain	25Spain	40Spain	48Ind	8Indexes
in this paper							
P_{12}	0.0282 (0.9472)	0.0347 (0.8341)	0.0361 (0.5075)	0.0194 (0.7978)	0.0169 (0.7605)	0.0111 (0.9341)	0.0047 (0.9240)
P_{13}	0.0312 (0.7430)	0.0446 (0.5458)	0.0367 (0.4705)	0.0234 (0.5854)	0.0177 (0.7078)	0.0107 (0.9834)	0.0049 (0.9039)
P_{14}	0.0280 (0.9598)	0.0397 (0.7016)	0.0293 (0.7715)	0.0205 (0.7453)	0.0164 (0.7908)	0.0111 (0.9485)	0.0042 (0.9833)
P_{23}	0.0320 (0.7060)	0.0243 (0.7487)	0.0256 (0.9180)	0.0192 (0.8088)	0.0133 (0.9833)	0.0106 (0.9738)	0.0042 (0.9846)
P_{24}	0.0278 (0.9759)	0.0235 (0.7487)	0.0208 (0.8531)	0.0182 (0.8787)	0.0133 (0.9833)	0.0111 (0.9345)	0.0037 (0.9591)
P_{34}	0.0285 (0.9223)	0.0310 (0.9764)	0.0292 (0.7622)	0.0199 (0.7826)	0.0162 (0.9833)	0.0107 (0.9981)	0.0042 (0.9788)
P_R	0.0283 (0.9378)	0.0135 (0.3829)	0.0168 (0.6492)	0.0115 (0.6727)	0.0102 (0.8148)	0.0107 (0.9864)	0.0039 (0.9790)
for comparison							
1/n	0.0274	0.0304	0.0240	0.0164	0.0130	0.0108	0.0041
MEAN	0.0246 (0.8421)	0.0023 (0.1140)	0.0030 (0.1368)	0.0088 (0.5199)	−0.0083 (0.1558)	0.0102 (0.9143)	0.0026 (0.8323)
MEANU	0.0327 (0.6770)	0.0220 (0.6557)	0.0310 (0.6814)	0.0427 (0.2437)	0.0398 (0.5245)	0.0072 (0.6197)	−0.0156 (0.7417)
MIN	0.0297 (0.8544)	0.0045 (0.1485)	0.0082 (0.2640)	0.0085 (0.4517)	0.0136 (0.9611)	0.0082 (0.5604)	0.0033 (0.9003)
MINU	0.0306 (0.8002)	0.0029 (0.1260)	0.0080 (0.2628)	−0.0042 (0.0644)	0.0001 (0.4270)	0.0041 (0.1315)	−0.0016 (0.3476)

Tables 4.6: Portfolio standard deviations

Strategy	5Spain	6Spain	10Spain	25Spain	40Spain	48Ind	8Indexes
in this paper							
P_{12}	0.0390	0.0651	0.0656	0.0388	0.0457	0.0380	0.0436
P_{13}	0.0417	0.0844	0.0613	0.0463	0.0437	0.0382	0.0486
P_{14}	0.0389	0.0904	0.0660	0.0451	0.0461	0.0396	0.0468
P_{23}	0.0452	0.0489	0.0476	0.0355	0.0419	0.0379	0.0424
P_{24}	0.0400	0.0683	0.0569	0.0385	0.0418	0.0393	0.0439
P_{34}	0.0401	0.0719	0.0599	0.0440	0.0431	0.0393	0.0465
P_R	0.0426	0.0508	0.0472	0.0386	0.0412	0.0390	0.0442
for comparison							
1/n	0.0392	0.0809	0.0628	0.0434	0.0440	0.0396	0.0459
MEAN	0.0595	0.0292	0.0279	0.0398	0.0591	0.0446	0.0474
MEANU	0.0493	0.0462	0.0579	0.1016	0.2032	0.0780	0.5845
MIN	0.0456	0.0328	0.0298	0.0295	0.0388	0.0358	0.0346
MINU	0.0494	0.0331	0.0311	0.0328	0.0675	0.0363	0.0367

Tables 4.7: Portfolio Sharpe ratios

Strategy	5Spain	6Spain	10Spain	25Spain	40Spain	48Ind	8Indexes
in this paper							
P_{12}	0.7218	0.5333	0.5498	0.5006	0.3700	0.2929	0.1070
	(0.6948)	(0.1315)	(0.0418)	(0.0314)	(0.0956)	(0.0965)	(0.3158)
P_{13}	0.7478	0.5279	0.5989	0.5056	0.4044	0.2789	0.1003
	(0.6084)	(0.1399)	(0.0378)	(0.0854)	(0.0179)	(0.5170)	(0.4829)
P_{14}	0.7196	0.4391	0.4438	0.4558	0.3564	0.2793	0.0896
	(0.6466)	(0.0519)	(0.2303)	(0.0978)	(0.0819)	(0.3309)	(0.8759)
P_{23}	0.7080	0.4962	0.5375	0.5406	0.3166	0.2801	0.0985
	(0.9093)	(0.2988)	(0.1723)	(0.0178)	(0.5215)	(0.4466)	(0.5582)
P_{24}	0.6941	0.3446	0.3656	0.4735	0.3182	0.2836	0.0848
	(0.8454)	(0.3012)	(0.7308)	(0.0610)	(0.5137)	(0.1533)	(0.6856)
P_{34}	0.7114	0.4308	0.4881	0.4514	0.3766	0.2731	0.0910
	(0.6893)	(0.1397)	(0.0025)	(0.0198)	(0.0204)	(0.8809)	(0.7383)
P_R	0.6657	0.2659	0.3550	0.2986	0.2471	0.2737	0.0879
	(0.3425)	(0.1807)	(0.7324)	(0.3307)	(0.4175)	(0.8022)	(0.9461)
for comparison							
1/n	0.6997	0.3753	0.3815	0.3791	0.2955	0.2719	0.0883
MEAN	0.4132	0.0804	0.1075	0.2213	−0.1400	0.2296	0.0555
	(0.0750)	(0.1902)	(0.1999)	(0.4145)	(0.0024)	(0.4806)	(0.7131)
MEANU	0.6632	0.4750	0.5354	0.4201	0.1960	0.0921	−0.0267
	(0.7598)	(0.3314)	(0.1060)	(0.8452)	(0.6209)	(0.0519)	(0.4246)
MIN	0.6502	0.1373	0.2745	0.2881	0.3500	0.2293	0.0961
	(0.5314)	(0.2605)	(0.5303)	(0.5073)	(0.5276)	(0.4326)	(0.8968)
MINU	0.6199	0.0871	0.2577	−0.1271	0.0012	0.1123	−0.0426
	(0.4932)	(0.1989)	(0.4981)	(0.0276)	(0.0948)	(0.0393)	(0.0640)

Tables 4.8: Portfolio turnovers

Strategy	5Spain	6Spain	10Spain	25Spain	40Spain	48ind	8Indexes
in this paper							
P_{12}	0.4573	0.3272	0.4951	0.6106	0.6028	0.6468	0.5152
P_{13}	0.3310	0.2931	0.5066	0.6377	0.6691	0.6510	0.5937
P_{14}	0.1358	0.1839	0.2367	0.4187	0.4728	0.5358	0.1946
P_{23}	0.2387	0.3343	0.4985	0.5773	0.6297	0.6622	0.5324
P_{24}	0.0988	0.1164	0.1969	0.3775	0.4804	0.5012	0.2282
P_{34}	0.1401	0.1520	0.2019	0.4212	0.4934	0.4963	0.3084
P_R	0.0448	0.0414	0.0468	0.0439	0.0449	0.0345	0.0365
for comparison							
1/n	0.0348	0.0558	0.0476	0.0378	0.0368	0.0324	0.0342
MEAN	0.1572	0.2112	0.2163	0.3503	0.4930	0.1112	0.2996
MEANU	0.1976	0.2056	0.2419	1.1022	3.9736	0.9607	39.0375
MIN	0.0848	0.0597	0.0698	0.1313	0.2292	0.0422	0.0744
MINU	0.1442	0.0682	0.0986	0.3243	1.4437	0.2312	0.2830

4.5.3 Discussion of results

We also want to emphasize that obtaining the best portfolio through an extremality order for any strategy is fast to compute and may be applied to portfolios of high-dimensional data. Considering k criteria chosen by the investor and a size m of feasible portfolios to be simulated, the computational cost of this calculation has complexity $\mathcal{O}(k.m)$, where $k << m$

- **Portfolio return. Table 4.5**

  Table 4.5 reports the out-of-sample returns of the benchmark portfolios and the portfolios obtained through the strategies proposed in this Chapter. The corresponding p-value of the difference between the return of each strategy and that $\frac{1}{n}$-rule and we affirm that the difference is significant if the p- value is smaller than 0.05. The p-values are computed by using the well-known t-student test for difference of means.

  We observe that the strategies that consider the return as a selected criterion of Θ, for example, P_{12}, P_{13} and P_{14}, in general, have a better out-of sample return than those used as benchmark. Note also that the best behavior, in terms of the returns, for all data sets is given for one of our strategies, except for the 5Spain and 40Spain sets. We can also see that all of our strategies have higher returns than (MEAN) strategy for all data set.

  Although we cannot see a significant statistical difference, the out-of-sample portfolio returns based on the strategies discussed in this Chapter are higher than those of other classical strategies used for comparison. This is attractive for the investors, but it is not the main criterion for choosing a good strategy, since investors would also want to take risk into account.

- **Portfolio standard deviation. Table 4.6**

  We can see from Table 4.6 that Property 4.3.6 seems be true for portfolios with a major numbers of assets, since our P_R strategy has a lower out-of-sample standard deviation than $\frac{1}{n}$-rule for all data sets except for 5Spain set. Observe also that the best among our strategies are those where the standard deviation has been considered as a criterion for selecting the portfolio, i.e., P_{12}, P_{23} and P_{24}. Another observation to highlight is that P_{23} and P_{24} also have a out-of-sample standard deviation than $\frac{1}{n}$-rule for all data sets except for 5Spain set. In general, (MIN) and (MINU) are the strategies which behave well with respect to standard deviation, however their performance in terms of returns and the Sharpe ratio is lower than any of the strategies studied in this Chapter.

- **Portfolio Sharpe ratio. Table 4.7**

  Table 4.7 shows the out-of-sample Sharp ratio for the different portfolios and their corresponding p-values that the Sharpe ratio for each one of these strategies is different from that for the $\frac{1}{n}$ rule. The p-values are computed by using the test for difference in Sharpe ratios as previously defined. We see that the P_{12} strategy always has higher Sharpe ratios than the policies for comparison in all data sets and that the differences are wide in almost all cases and statistically significant for 10Spain and 25Spain. Observe also that the P_{13} strategy always has higher Sharpe ratios than those policies for comparison in all data sets and the differences are wide in almost all cases but are only statistically significant for 10Spain and 40Spain. P_{14} attains a better Sharpe ratio than the benchmark strategies except for 6Span and 10Spain. The same applies for P_{23} which, in terms of Sharpe ratio, is only surpassed by (MIN) in 40Spain data. We see that P_{14} has a better behavior than (MEAN) and (MINU) for all data sets. This strategy also has higher Sharpe ratio than (MEAN) except for 6Spain, 10Spain data sets and it has higher Sharpe ratio than (MIN) except 40Spain data set. On the other hand, P_{34} attains higher Sharpe ratios than all policies of comparison except for (MEANU). Note also that our P_R strategy reaches higher Sharpe ratios than both the (Mean) and the (MINU) strategies for any of the data sets.

  Summarizing, we can affirm that the strategies introduced in this Chapter, generally have higher Sharpe ratios than the policies that we have considered as benchmark. Observe also that among our strategies, those that have the Sharpe ratio as criterion attain a very good performance in this aspect, i.e., P_{13}, P_{23}, P_{34}

- **Portfolio Turnovers. Table 4.8**

  In Table 4.8 it may be observed that those of our strategies that employ entropy as criterion of selection ,i.e, P_{14}, P_{24}, P_{34} attain a lower turnover than the strategies with other criteria. Obviously the best strategy in terms of turnover is the $\frac{1}{n}$ rule, since this diversification always reaches the maximum entropy. However, the P_R strategy not only performs well with respect to the turnover but, in fact behaves better for all data sets than any strategy except for the $\frac{1}{n}$ rule. Observe that P_R even attains the best turnover for all strategies in

6Spain and 10Spain data sets.

4.5.4 Sensitivity

In order to analyze the sensitivity of the results obtained with our strategies, we have used 6FF which is a monthly data set taken from the Kenneth French web-site, for the time period $07/1963 - 12/2004$. Figure 4.4 depicts the box-plots for different numbers of portfolios generated through

$$\mathbf{U} = \left(\sum_{i=1}^{6} U_i\right)^{-1} [U_1, \ldots, U_6]' \quad \text{where} \quad U_i \sim Uniform(0,1), \ \text{ for all } \ i = 1, \ldots, 6. \tag{4.5.1}$$

for each number of portfolios feasible, we have calculated 5000 times the performance measure discussed in this Chapter, i.e, returns, risk, Sharpe ratio and turnovers and the box plots of each single measure are reported in the following Figure 4.4.

We can see that when there is an increase in the number of generated feasible portfolios, there is an improvement in the performance of the criteria considered. In general, we can affirm that the results given for these strategies are stable since the box plots display little variability and the distributions of the performance measures do not present outliers.

Although, we have made the box plots for each of the strategies discussed in this Chapter, we only show the box plots associated with strategy P_{12}. The same conclusions, however, can be observed with the other strategies described in the Chapter.

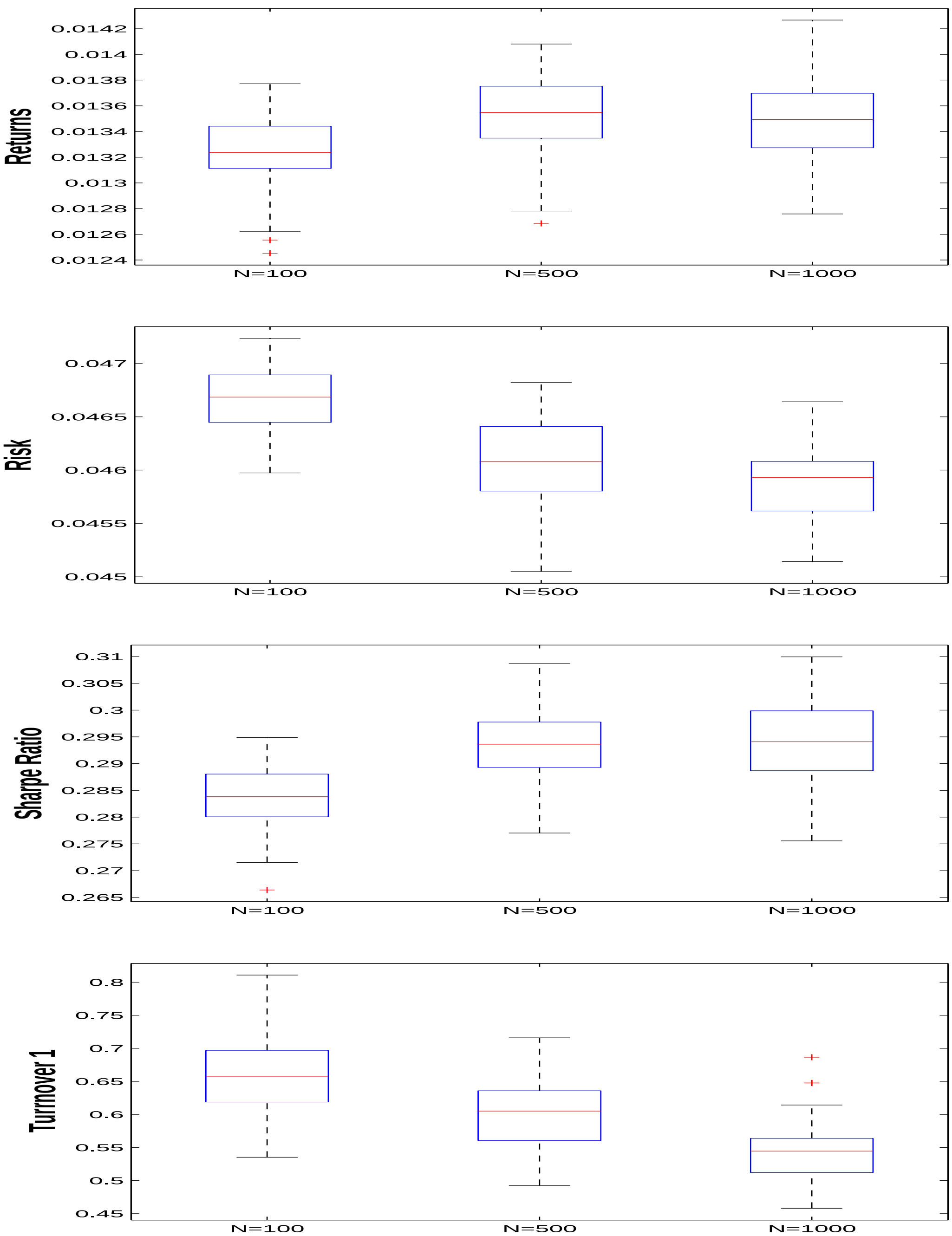


Figure 4.4: Box Plot for different numbers of feasible portfolio

4.6 Conclusions

We present portfolio selection strategies based on the data order induced by extremality and propose a procedure to sort feasible portfolios. We provide a methodology for selecting portfolio weights considering some criteria which are different from the classical criteria of mean and variance. Through the order discussed in this Chapter, we rank feasible portfolios and we select the best portfolio as the smallest in terms of extremality. Therefore, this framework is not based on optimization. We also define alternative efficient frontiers which depend on the criteria chosen by the investor, which could be surfaces rather than curves, and we show that the portfolio selected with the strategies studied belongs to the efficient frontier. The performance of the strategies introduced is compared with the performance of classic policies from the literature and we find that our strategies are better in terms of the Sharpe ratio than those used as a benchmark.

We also have introduced the strategy P_R that depends on the standard deviation vector of assets. We find that this policy often has a lower variance than $\frac{1}{n}$ rule. While we have proved that the previous statement holds in the case of two assets, We do not have the results for the general case. It would therefore be interesting to investigate the conditions that extend the result for a data set of n assets.

Our methodology can be extended by considering the criteria set Θ with $k > 2$ or by considering other criteria different from those studied in this Chapter. For example, it is well known that the popular Value at Risk (VaR) is not a convex measure risk, therefore its optimization process is complex, and so we also propose to develop our strategies taking (VaR) as a criterion of selection.

5 Conclusions and main contributions

This chapter summarizes the main contributions of the thesis. In general terms, this work has provided a methodology based on directions for comparing both multivariate data and random vectors. Each chapter offers a theoretical basis for constructing either multivariate data orders or multivariate stochastic orders. Also included are the respective applications of the orders introduced, especially applications that are addressed toward the financial framework.

Although each chapter of this thesis states its main conclusions and contributions, the principal aspects developed are briefly described as follows:

- We have developed an alternative approach for analyzing multivariate extremes by considering directions. In order to do this, an oriented convex cone has been constructed to calculate a new multivariate measure referred to throughout the work as extremality.
  - Through the extremality measure, a new way to sort multivariate data based on directions is introduced. Each orientation of the convex cone allows us to define different forms to rank the sample.
  - The inclusion of directions is a versatile way of introducing new definitions of multivariate quantiles and generalizes some quantiles already studied in the literature.
  - As an application of extremality, an oriented multivariate Value at Risk is defined as its level set. A direction chosen by the investor allows to identify the riskiest events. The VaR introduced in this work can be seen, by inclusion of directions, as a generalization of some VaR's already studied in the literature.

  - A study is made of the relationship between the oriented multivariate Value at Risk and the classical one-dimensional Value at Risk when an investor wants to diversify his/her wealth in a portfolio.

- We have provided a multivariate stochastic order which in this thesis also is referred as extremality order. Unlike previous comparisons of multivariate data, the second contribution is an order to compare random vectors. The new multivariate stochastic order is defined through a oriented convex cone. Both necessary and sufficient conditions are investigated and a study is also made of the relationships between other multivariate stochastic orders introduced in the literature.

  - In this work, the well known upper orthant order and lower orthant orders have been generalized by including directions. We have shown some examples where other directions can be more useful than those used to define the classical upper and lower orthant orders
  - We present some examples of application in the determination of optimal allocations of wealth among risks in single period portfolio problems. We present some examples of application in the determination of optimal allocations of wealth among risks in single period portfolio problems
  - We have studied the case in which an agent has to allocate his capital in different but not independent risky assets and we find an optimal solution based on rotations of the risky assets (random variables) such that the maximal diversification in the rotated vector gives the maximal expected utility.

- Markowitz defined the efficient frontier as the set of feasible portfolios which cannot be improved in terms of risk and return simultaneously. Inspired by the Markowitz idea, we have introduced new concepts of efficient frontier which depend on some indexes that the investor can choose and which may be different from the classical variance- return in Markowitz's model.

  - We have provided a version of extremality data order in order to sort feasible portfolio. This portfolio order allows us to estimate which portfolio is best. The procedure used does not use optimization techniques.
  - The methodology proposed is also versatile since the investor can choose the portfolio depending on the criteria which are most relevant and which may be different those of mean and variance.

- The different policies introduced in this work frequently attain a better performance in terms of the Sharpe ratio than those classical strategies used as a benchmark.
- A strategy is developed which is based on rotations in the direction of the standard deviation vector of data. We have proved that at least for the bivariate case, the policy mentioned has less risk than the $\frac{1}{n}$ rule which assigns the same weight to every risky asset.

Bibliography

[1] S. Ahmed, D. Filipovic, and G. Svindland. A note on natural risk statistics. *Oper. Res. Lett.*, 36(6):662–664, 2008.

[2] R. Barlow. *Statistical theory of reliability and life testing. Probability models.* Holt, Rinehart and Winston, New York, 1975.

[3] V. Barnett. The ordering of multivariate data. *Journal of the Royal Statistical Society. Series A (General)*, 139(3):318–355, 1976.

[4] F. Belzunce, A. Castaño, A. Olvera-Cervantes, and A. Suárez-Llorens. Quantile curves and dependence structure for bivariate distributions. *Computational Statistics & Data Analysis*, 51(10):5112–5129, 2007.

[5] N. Bäuerle and A. Müller. Stochastic orders and risk measures: consistency and bounds. *Insurance: Mathematics and Economics*, 38(1):132–148, 2006.

[6] C. Burgert and L. Ruschendorf. Consistent risk measures for portfolio vectors. *Insurance: Mathematics and Economics*, 38(2):289–297, 2006.

[7] I. Cascos. Data depth: multivariate statistics and geometry. In *New Perspectives in Stochastic Geometry, Oxford University (W.S. Kendall and I. Molchanov Eds.)*, pages 398–423, Oxford University Press, 2010.

[8] I. Cascos, A. López, and J. Romo. Data depth in multivariate statistics. *Boletín de Estadística e Investigación Operativa*, 27(3):151–174, 2011.

[9] I. Cascos and I. Molchanov. Multivariate risks and depth-trimmed regions. *Finance and Stochastics*, 11(3):373–397, 2007.

[10] P. Chaudhuri. On a geometric notion of quantiles for multivariate data. *Journal of the American Statistical Association*, 91(434):862–872, 1996.

[11] V. DeMiguel, L Garlappi, F.J. Nogales, and R. Uppal. A generalized approach to portfolio optimization: improving performance by constraining portfolio norms. *Management Science*, 55(5):798–812, 2009.

[12] V. DeMiguel, L. Garlappi, and R. Uppal. Optimal versus naive diversification: how innefficient is the $1/n$ portfolio strategy? *The Review of Financial Studies*, 22(5):1915–1953, 2009.

[13] V. DeMiguel and F.J. Nogales. Portfolio selection with robust estimation. *Operations Research*, 57(3):560–577, 2009.

[14] M. Denuit, J. Dhaene, M. Goovaerts, and R. Kaas. *Actuarial theory for dependent risks: measures, orders and models.* John Wiley & Sons, 2006.

[15] J.H.J. Einmahl and D.M. Mason. Generalized quantile processes. *The Annals of Statistics*, 20(2):1062–1078, 1992.

[16] P. Embrechts, A. McNeil, and D. Straumann. Correlation and dependence in risk management: properties and pitfalls. In M. A. H. Dempster, editor, *Risk Management: Value at Risk and Beyond*, pages 176–223. Cambridge University Press, Cambridge, 2002.

[17] P. Embrechts and G. Puccetti. Bounds for functions of multivariate risks. *Journal of Multivariate Analysis*, 97(2):526–547, 2006.

[18] P. Embrechts and G. Puccetti. Bounds for the sum of dependent risks having overlapping marginals. *Journal of Multivariate Analysis*, 101(1):177–190, 2010.

[19] J.M. Fernández-Ponce and A. Suárez-Lloréns. Central regions for bivariate distributions. *Austrian Journal of Statistic*, 31(2-3):141–156, 2002.

[20] H. Föllmer and A. Schied. Convex measures of risk and trading constraints. *Finance and Stochastics*, 6(4):429–447, 2002.

[21] A.A. Gaivoronski and G. Pflug. Value at risk in portfolio optimization: properties and computational approach. *Journal of Risk*, 7(2):1–28, 2000.

[22] P. Gänssler. *Empirical processes.* Lecture notes-monograph series. Institute of Mathematical Statistics, 1983.

[23] J.E. Gentle. *Numerical linear algebra for applications in statistics*. Statistics and computing. Springer, 1998.

[24] J. Hadar and W. R. Russell. Stochastic dominance and diversification. *Journal of Economic Theory*, 3(3):288–305, 1971.

[25] M. Hallin, D. Paindaveine, and M. Šiman. Multivariate quantiles and multiple-output regression quantiles: from L_1 optimization to halfspace depth. *The Annals of Statistics*, 38(2):635–669, 2010.

[26] C.C. Heyde, S.G. Kou, and X.H. Peng. What is a good external risk measure: bridging the gaps between robustness, subadditivity, and insurance risk measure. *Workinpaper*, 2009.

[27] R. Jagannathan and T. Ma. Risk reduction in large portfolios: why imposing the wrong constraints helps. Working Paper 8922, National Bureau of Economic Research, 2002.

[28] J. D. Jobson, Korkie, and Bob M. Performance hypothesis testing with the sharpe and treynor measures. *The Journal of Finance*, 36(4):889–908, 1981.

[29] P. Jorion. *Value at Risk: the New Benchmark for Managing Financial Risk*. McGraw-Hill, 2007.

[30] S. Karlin and Y. Rinott. Classes of orderings of measures and related correlation inequalities. i. multivariate totally positive distributions. *Journal of Multivariate Analysis*, 10(4):467–498, 1980.

[31] S. Karlin and Y. Rinott. Total positivity properties of absolute value multinormal variables with applications to confidence interval estimates and related probabilistic inequalities. *The Annals of Statistics*, 9(4):1035–1049, 1981.

[32] L. Kong and I. Mizera. Quantile tomography: using quantiles with multivariate data. *ArXiv e-prints*, 2008.

[33] Z. Landsman. Tail conditional expectations for elliptical distributions. *American Actuarial Journal*, 7(4):55–71, 2003.

[34] Z. Landsman. On the generalization of esscher and variance premiums modified for the elliptical family of distributions. *Insurance: Mathematics and Economics*, 35(3):563–579, 2004.

[35] Z. Landsman and A. Tsanakas. Stochastic ordering of bivariate elliptical distributions. *Statistics & Probability Letters*, 76(5):488–494, 2006.

[36] H. Laniado, R.E. Lillo, F. Pellerey, and J. Romo. Portfolio selection through an extremality stochastic order. *Insurance: Mathematics and Economics*, 51(1):1–9, 2012.

[37] H. Laniado, R.E. Lillo, and J. Romo. Multivariate extremality measure. Preprint 10-19 08, Universidad Carlos III de Madrid, Spain, 2010.

[38] E.L. Lehmann. Some concepts of dependence. *The Annals of Mathematical Statistics*, 37(5):1137–1153, 1966.

[39] J. Li and R.Y. Liu. Multivariate spacings based on data depth: I. construction of nonparametric multivariate tolerance regions. *The Annals of Statistics*, 36(3):1299–1323, 2008.

[40] R.Y. Liu. On a notion of data depth based on random simplices. *Annals of Statistics*, 18:405–414, 1990.

[41] R.Y. Liu, J. M. Parelius, and K. Singh. Multivariate analysis by data depth: descriptive statistics, graphics and inference. *The Annals of Statistics*, 27(3):783–840, 1999.

[42] C. Ma. Convex orders for linear combinations of random variables. *Journal of Statistical Planning and Inference*, 84(1-2):11–25, 2000.

[43] H. Markowitz. Mean-variance analysis in portfolio choice and capital markets. *Journal of Finance*, 7(1):77–91, 1952.

[44] A.W. Marshall and I. Olkin. *Inequalities: Theory of Majorization and Its Application.* Academic Press, 1979.

[45] C. Memmel. Performance hypothesis testing with the sharpe ratio. *Finance Letters*, 1(1):21–23, 2003.

[46] A. Müller and D. Stoyan. *Comparison Methods for Stochastic Models and Risks.* Wiley series in probability and statistics. Wiley, 2002.

[47] R.B. Nelsen. *An Introduction to Copulas.* Springer Series in Statistics. Springer, 2006.

[48] F. Pellerey and P. Semeraro. A note on the portfolio selection problem. *Theory and Decision*, 59(4):295–306, 2005.

[49] A.G. Quaranta and A. Zaffaroni. Robust optimization of conditional value at risk and portfolio selection. *Journal of Banking and Finance*, 32(10):2046 – 2056, 2008.

[50] Y. Rinott and M. Scarsini. Total positivity order and the normal distribution. *Journal of Multivariate Analysis*, 97(5):1251–1261, May 2006.

[51] R.T. Rockafellar. *Convex analysis*. Princeton Mathematical Series. Princeton University Press, 1997.

[52] R.T. Rockafellar and S. Uryasev. Optimization of conditional value-at-risk. *Journal of Risk*, 2:21–41, 2000.

[53] M. Scarsini. Multivariate stochastic dominance with fixed dependence structure. *Operations Research Letters*, 7(5):237 – 240, 1988.

[54] R. Serfling. Generalized quantile processes based on multivariate depth function, with applications in nonparametric multivariate analysis. *J. Multivar. Anal.*, 83.

[55] M. Shaked and J.G. Shanthikumar. *Stochastic orders*. Springer series in statistics. Springer, 2007.

[56] W. Sharpe. Adjusting for risk in portfolio performance measurement. *The Journal of Portfolio Management*, 1(2):7–19, 1975.

[57] D. Stoyan and D.J. Daley. *Comparison methods for queues and other stochastic models*. Wiley series in probability and mathematical statistics: Applied probability and statistics. Wiley, 1983.

[58] L. Tibiletti. On a new notion of multidimensional quantile. *Metron. International Journal of Statistics*, 51(3-4):77–83, 1993.

[59] L. Tibiletti. Incremental value at risk: traps and misinterpretations. In *Trends in Mathematics*, pages 355–364, Basel, Switzerland, October 5-7 2000. Ed. M. Kohlmann and Shanjian Tang, Birkhauser Verlag.

[60] John W. Tukey. Mathematics and the picturing of data. In *Proceedings of the International Congress of Mathematicians (Vancouver, B. C., 1974), Vol. 2*, pages 523–531. Canad. Math. Congress, Montreal, Que., 1975.

[61] I. Usta and Y. M. Kantar. Mean-variance-skewness-entropy measures: a multi-objective approach for portfolio selection. *Entropy*, 13(1):117–133, 2011.

[62] E.A. Valdez and J. Dhaene. Convex order bounds for sum of dependent log-elliptical random variables. In *7th International Congress on Insurance: Mathematics & Economics*.

[63] M.E Yaari. The dual theory of choice under risk. *Econometrica*, 55(1):95–115, January 1987.

[64] Y. Zuo and R. Serfling. General notions of statistical depth function. *The Annals of Statistics*, 28(2):pp. 461–482, 2000.