EXTREMAL DEPENDENCE OF MULTIVARIATE DISTRIBUTIONS AND ITS APPLICATIONS

By

YANNAN SUN

A dissertation submitted in partial fulfillment of the
requirements for the degree of

DOCTOR OF PHILOSOPHY

WASHINGTON STATE UNIVERSITY
Department of Mathematics

MAY 2010

I certify that I have read this dissertation and certify that in my opinion it is fully adequate, in scope and in quality, as a dissertation for the degree of Doctor of Philosophy.

Haijun Li, Ph.D., Chair

Ari Ariyawansa, Ph.D.

Alan Genz, Ph.D.

EXTREMAL DEPENDENCE OF MULTIVARIATE DISTRIBUTIONS

AND ITS APPLICATIONS

Abstract

by Yannan Sun, Ph.D.
Washington State University
MAY 2010

Chair: Haijun Li

Stochastic dependence arises in many fields including electrical grid reliability, network/internet traffic analysis, environmental and agricultural impact assessment, and financial risk management. Models in these fields exhibit the stochastic behavior of strong dependence—often among their extreme values. Extremal dependence is directly related to the dynamics of a system. It is critically important to understand this relationship. If not, the extremal dependence can cause long term, contagious damage.

The tail dependence of multivariate distributions is frequently studied via the tool of copulas, but copula-based methods become inefficient in higher dimensions. The theories of multivariate regular variation and extreme values bring in many results for multivariate distributions based on various measures. Applying these theories, however, is difficult because the measures are not explicit. In this dissertation, we establish the link between tail dependence parameters and the theory of multivariate regular variation. For a multivariate regularly varying distribution, we show that the upper tail dependence function and the intensity measure are equivalent in describing its extremal dependence structure. With this new characterization we can efficiently evaluate tail dependence for multivariate distributions when the copulas are not explicitly accessible.

As applications, we derive tractable formulas for tail dependence parameters of heavy-tailed scale mixtures of multivariate distributions. We discuss the structural traits of multivariate tail dependence, such as the monotonicity properties of tail dependence for bivariate elliptical distributions. We also give a general method to approximate the Value-at-Risk for aggregated data in terms of the tail dependence function of the data and the Value-at-Risk for one of the variables. Explicit approximations are obtained for Archimedean copulas, Archimedean survival copulas and Pareto distributions of Marshall-Olkin type. Additionally, to visualize the monotonicity properties of the tail dependence parameters, we give some simulation results for Pareto distributions in three and four dimensions.

Acknowledgments

There are many people I would like to acknowledge for helping me finish this dissertation as well as successfully completing my career at Washington State University. Because of them being a student and instructor in the Mathematics Department was a joy. I thank my advisor Haijun Li for his guidance, patience and support. He not only brought me to the forefront of our research area, but is also an inspiration as a passionate and enthusiastic mathematician. I am also very grateful to professors Ari Ariyawansa, Alan Genz and Marc Evans who always had time to give me advice.

I have always felt lucky and blessed that I have such a loving and supportive family. I thank my parents for everything they have done for me. Without their unending encouragement and all kinds of support, my time studying abroad and becoming a mathematician would have been all the more arduous. I can feel their love no matter where I am or what I've been up to. The appearance of my husband Andrew Stevens is the best surprise I have ever had. I thank him for being an understanding friend, a patient listener, my personal editor, the household mechanic and an occasional trouble maker in my life. I also thank my extended family; Aunt Pearl Mingzhu Sun, Aunt Yan Li and Uncle Yibing Zhang for all the care and love I have received. I am very proud to have all of them.

I would also like thank my friends and colleagues in Pullman for their help, support, company and all the great memories in the past six years. Life would have been very boring without them. Specifically, I thank Huan Zhao, Liying Zheng, Miao Miao, Pengyu Liu, Limin Yang, Yin Chan, Elisabeth Ellison, and Sherod Eubanks for the things I have forgotten or do not have space to include.

List of Figures

Contents

Chapter 1

Preliminaries

1.1 Introduction

Stochastic dependence arises in a wide range of applications, including financial risk management, network/internet traffic analysis, and environmental and agricultural impact assessment. In such applications, different physical processes exhibit the stochastic behavior of strong dependence – often among their extreme values. Extremal dependence is directly related to the system dynamics. It is critically important to understand this relationship. If not, the extremal dependence can cause long term, contagious, cascading damage. Take for example the global financial crisis during 2008; it is easy to see the effects of extremal dependence.

To model extremal events and their dependence the most commonly understood distribution, the normal distribution, does not work. This is because its tails are too thin. To allow for events in the tails to occur more often we need to use heavy-tailed (or regularly varying) distributions, such as t-distributions and various Pareto distributions. Tail dependence exists when the dependence among random variables concentrates in the extremely high (or low) values. In the bivariate case, tail dependence has been studied using the tool of copulas. A copula is the joint distribution of the percentiles of the random variables. By transforming all the marginal distributions to uniform distributions on the unit square (cube), the tail dependence structure is separated from the margins. Copula-based methods are widely used in financial risk management and actuarial science for statistical modeling. Copula functions can help determine the multivariate distribution that a sample of data follows. Yet, they become increasingly cumbersome and ineffective in higher dimensions. The theories of regular variation and extreme values bring in many results for multivariate distributions based on various measures. Applying these theories, however, is difficult because the measures are not explicit. A remedy, which my research addresses, is to express the measures in terms of tail dependence functions.

This dissertation focuses on characterizing the extremal dependence of high dimensional distributions and determining an efficient algorithm for extremal dependence measures. In this Chapter, we introduce some of the notation needed in later chapters. We review the definitions and basic properties of weak convergence and the generalized inverse of monotone functions to prepare for introducing copulas and regular variation. More detail can be found in [32].

In Chapter 2, we review the definition and properties of copulas. Some fundamental properties are given including Sklar's Theorem and Fréchet bounds. The family of Archimedean copulas is introduced because of its important role in modeling dependence structures. Next, we review three types of dependence measures: linear correlation, rank correlation, and the tail dependence parameters which are derived from copulas. We will see that the tail dependence parameters have better properties than linear and rank correlation. Tail dependence functions are defined in terms of copulas. Using tail dependence functions we can obtain various types of tail dependence parameters. These parameters describe the amount of dependence in the upper (lower) tail of a multivariate distribution. They measure the interactivity (e.g. possibility of contagion) among the multivariate margins of the random vector and play an important role in analyzing extremal dependence. Finally, elliptical distributions are introduced as the most widely used radially symmetric multivariate distributions. Elliptical distributions do not have explicit copulas in general. They play an important role, however, in analyzing tail dependence.

In Chapter 3, we first review the theory of regular variation and its relationship to extreme value theory. We survey the properties for regular variation of functions and of measures in both single-variable and multi-variable cases. The equivalent condition for elliptical distributions being multivariate regularly varying is given. In Section 3.3, we establish the link between tail dependence and the theory of multivariate regular variation. For a multivariate regularly varying distribution, we show that the upper tail dependence function and the intensity measure are equivalent in describing its extremal dependence structure. The proof employs the theory of regular variation even though the tail dependence is defined in terms of copulas. With this new characterization we can efficiently evaluate tail dependence for multivariate distributions when the copulas are not explicitly accessible.

In Chapter 4, we give several applications of Theorem 3.3.1, the main result of Chapter 3. First, we derive tractable formulas for the tail dependence parameters of heavy-tailed scale mixtures of multivariate distributions whose copulas are not explicitly accessible in general. These formulas depend on the joint moments and the heavy-tail index of the mixing random variable. Unlike the copula method, they avoid taking the complicated marginal transforms on the entire distribution. We also discuss the structural traits of multivariate tail dependence, such as monotonicity properties of tail dependence for bivariate elliptical distributions. Bivariate elliptical distributions possess the upper (lower) tail dependence property if their generating random variable is regularly varying [17, 33]. The formula for the extremal dependence γ of bivariate elliptical distributions is also derived in [14]. Our method yields similar results for multivariate elliptical distributions. Second, we give a general method to approximate the Value-at-Risk for aggregated data in terms of the tail dependence function of the data and the Value-at-Risk for one of the variables. The explicit approximations are obtained for Archimedean copulas, Archimedean survival copulas and Pareto distributions of Marshall-Olkin type. Last, to visualize the monotonicity properties of the tail dependence parameters, we give some simulation results for Pareto distributions with three and four dimensions.

1.2 Notation

In this section, we introduce some notation that will be used throughout the dissertation. We try to maintain notation consistent with the literature.

Numbers and Sets

$\mathbb{N}$ denotes the set of natural numbers along with zero $\{0, 1, 2, \ldots\}$. $\mathbb{R}$ denotes the real line $(-\infty, \infty)$. $\mathbb{R}_+$ denotes the nonnegative reals $(0, \infty)$. $\mathbb{R}^d$ is the d-dimensional real space $\mathbb{R} \times \cdots \times \mathbb{R}$. $\mathbb{R}_+^d$ is the first orthant of $\mathbb{R}^d$, i.e., the space of all the d-dimensional vectors with every component positive. The d-dimensional unit box is denoted $[0, 1]^d$. $\overline{\mathbb{R}}$ denotes the extended real line $[-\infty, \infty]$. $\overline{\mathbb{R}}^d$ is the d-dimensional extended real space $\overline{\mathbb{R}} \times \cdots \times \overline{\mathbb{R}}$. The notations $a \vee b$ and $a \wedge b$ mean the maximum and the minimum of a and b respectively. For a set S, $|S|$ is the number of elements in S. The set $I = \{1, 2, \ldots, d\}$ denotes the index set.

Vectors and Matrices

We use vector notation for points in $\overline{\mathbb{R}}^d$, e.g., $\boldsymbol{x} = (x_1, \ldots, x_d)$. For two vectors $\boldsymbol{a}, \boldsymbol{b} \in \overline{\mathbb{R}}^d$ we write $\boldsymbol{a} \leq \boldsymbol{b}$ if $a_i \leq b_i$ for each i; and $\boldsymbol{a} < \boldsymbol{b}$ if $a_i < b_i$ for each i. We denote by $[\boldsymbol{a}, \boldsymbol{b}]$ the Cartesian product $[a_1, b_1] \times \cdots \times [a_d, b_d]$ for $\boldsymbol{a} \leq \boldsymbol{b}$. Also denote $(x_1^p, \ldots, x_d^p)$ by $\boldsymbol{x}^p$ where $p > 0$, $x_i > 0$ for each i.

For a matrix A, A^T is the transpose of A. The matrix I_d is the identity matrix in $\mathbb{R}^{d \times d}$. $D[a_1, \ldots, a_d]$ is a $d \times d$ diagonal matrix with main diagonal entries $a_1, \ldots, a_d$.

Functions

For a real function $H : \overline{\mathbb{R}}^d \to \mathbb{R}$, $\mathrm{Dom}(H)$ and $\mathrm{Ran}(H)$ denote the domain and the range of H respectively. The terms "increasing" and "decreasing" are in the weak sense and mean non-decreasing and non-increasing respectively. We will use "strictly" for the strong sense of monotonicity. The function H is said to be monotone if it is monotone in every component. For two functions $f(x)$ and $g(x)$, we use the notation $f(x) \sim g(x)$ as $x \to a$ if $\lim_{x \to a} \frac{f(x)}{g(x)} = 1$.

Probability

For two random variables X and Y, $X \overset{d}{=} Y$ means that X and Y follow the same distribution.

1.3 Uniform Convergence

Let $\{h_n, n \in \mathbb{N}\}$ be a sequence of real functions. On a set $A \subset \mathbb{R}$, we say that h_n converges *uniformly* to h_0 if

$$\lim_{n \to \infty} \sup_{t \in A} |h_n(t) - h_0(t)| = 0. \tag{1.3.1}$$

If (1.3.1) holds for any close interval $A \subset \mathbb{R}$, then we say that h_n converges *locally uniformly* to h_0. Furthermore, if h_n are increasing and h_0 is continuous, then h_n converges locally uniformly to h_0 if h_n converges pointwise to h_0, i.e., $\lim_{n\to\infty} h_n(t) = h_0(t)$, for all $t \in \mathbb{R}$.

Let $\{H_n, n \in \mathbb{N}\}$ be a sequence of increasing real functions. Define

$$\mathcal{C}(H_n) := \{t \in \mathbb{R} : H_n \text{ is finite and continuous at } t\}.$$

Then H_n converges weakly to H_0 if H_n converges pointwise to H_0 on $\mathcal{C}(H_0)$, i.e.,

$$\lim_{n\to\infty} H_n(t) = H_0(t), \text{ for all } t \in \mathcal{C}(H_0). \tag{1.3.2}$$

In probability and statistics, if H_n are distribution functions of random variables X_n and (1.3.2) holds, then we say that X_n converges *in distribution* to X_0.

1.4 Generalized Inverses of Monotone Functions

Let $H : \mathbb{R} \to \overline{\mathbb{R}}$ be an increasing function. We define the (left-continuous) *generalized inverse* $H^{\leftarrow} : \overline{\mathbb{R}} \to \mathbb{R}$ of H as

$$H^{\leftarrow}(t) := \inf\{s : H(s) \geq t\}. \tag{1.4.1}$$

For a sequence of increasing real functions $\{H_n, n \in \mathbb{N}\}$ with range (a, b), let H_n converge weakly to H_0. Then we have

$$\lim_{n\to\infty} H_n^{\leftarrow}(t) = H_0^{\leftarrow}(t), \quad \text{for all } t \in (a, b) \cap \mathcal{C}(H_0^{\leftarrow}).$$

Chapter 2

Copulas and Dependence

In this chapter, we introduce the basic properties of copulas and several dependence measures. The copula is widely used tool for modeling the dependence structure of a random vector. The dependence measures derived from copulas have better properties than linear correlations. Using copulas, we define tail dependence functions which can be used to obtain various tail dependence parameters. Much of this chapter is based on [28, 29]. Proofs, examples, more details and further references can be found there.

2.1 Copulas

In recent years, copulas have become a popular tool in analysis of dependence. Copula functions are often used in financial risk management and actuarial science for modeling dependence structures in a sample of data. The basic idea is to transform all the marginal variables so that each of them follows the uniform distribution on $[0, 1]$. For a random variable X, if F is its distribution function (d.f.), then $F(X)$ has a uniform distribution. Similarly, for a d-dimensional random vector $\boldsymbol{X} = (X_1, \ldots, X_d)$ with marginal distributions $F_1, \ldots, F_d$, the transformed vector $(F_1(X_1), \ldots, F_d(X_d))$ has uniformly distributed margins and the copula of $\boldsymbol{X}$ is the joint distribution of the transformed vector. Copulas separate the dependence of $\boldsymbol{X}$ from the margins, which is particularly helpful in the understanding of the extremal dependence among the components of a random vector.

In this section, we give the definitions for copulas and survival copulas, and some fundamental properties including Sklar's Theorem and Fréchet bounds. Additionally, we introduce an important family of copulas, Archimedean copulas. When proving or applying the results given in this section, if the inverse of a monotone function does not exist, then we use the generalized inverse (1.4.1) instead.

2.1.1 Definitions and Basic Properties

In order to define a copula, we need to introduce the following "volume" function first.

Definition 2.1.1 Let H be a d-dimensional real function on $\overline{\mathbb{R}}^d$ such that $\mathrm{Dom}(H) = S_1 \times S_2 \times \cdots \times S_d$, where S_i are nonempty subsets of $\overline{\mathbb{R}}$. Let E be the set of all the vertices of a d-dimensional box $B = [\boldsymbol{a}, \boldsymbol{b}]$ such that $E \subseteq \mathrm{Dom}(H)$. Then the H*-volume* of B is given by

$$V_H(B) := \sum_{\boldsymbol{c} \in E} \mathrm{sgn}(\boldsymbol{c}) H(\boldsymbol{c}), \tag{2.1.1}$$

where $\mathrm{sgn}(\boldsymbol{c}) = \begin{cases} 1, & \text{if } |\{j : c_j = a_j\}| \text{ is even;} \\ -1, & \text{otherwise.} \end{cases}$

Take the bivariate case for example. Let H be a 2-dimensional real function such that $\mathrm{Dom}(H) = S_1 \times S_2$ where S_1, S_2 are nonempty subsets of $\overline{\mathbb{R}}$. Let $B = [a_1, b_1] \times [a_2, b_2]$ be a rectangle and $E = \{(a_1, b_1), (a_2, b_1), (a_1, b_2), (a_2, b_2)\} \subseteq \mathrm{Dom}(H)$. Then the H-volume of B is given by

$$V_H(B) = H(a_1, b_1) + H(a_2, b_2) - H(a_2, b_1) - H(a_1, b_2). \tag{2.1.2}$$

Definition 2.1.2 A d-dimensional *copula* $C(\boldsymbol{u}) = C(u_1, \ldots, u_d)$ is a distribution function, defined on the unit cube $[0,1]^d$, with uniform one-dimensional margins. In addition, the following properties need to be satisfied:

(1) For any $\boldsymbol{u} \in [0,1]^d$, $C(\boldsymbol{u}) = 0$ if $\exists i \in I$ such that $u_i = 0$;

(2) $C(1, \ldots, 1, u_i, 1, \ldots, 1) = u_i$, for all $i \in I$;

(3) For any $\boldsymbol{a}, \boldsymbol{b} \in [0,1]^d$ such that $\boldsymbol{a} \le \boldsymbol{b}$, $V_C([\boldsymbol{a}, \boldsymbol{b}]) \ge 0$.

The following theorem is the foundation of the theory of copulas. It reveals the relationship between multivariate distribution functions and their univariate margins.

Theorem 2.1.3 (Sklar's Theorem, [35]) Let F be a d-dimensional joint distribution function with margins $F_1, \ldots, F_d$. Then there exists a copula C such that for all $\boldsymbol{t} \in \overline{\mathbb{R}}^d$,

$$F(t_1, \ldots, t_d) = C(F_1(t_1), \ldots, F_d(t_d)). \tag{2.1.3}$$

If $F_1, \ldots, F_d$ are all continuous, then the corresponding copula C is unique and can be written as $C(u_1, \ldots, u_d) = F(F_1^{-1}(u_1), \ldots, F_d^{-1}(u_d))$, $(u_1, \ldots, u_d) \in [0,1]^d$. Thus, for multivariate distributions with continuous margins, the univariate margins and the multivariate dependence structure (as described by their copulas) can be separated.

For a random vector $\boldsymbol{X}$ with distribution function F and continuous margins $F_1, \ldots,$ F_d, the copula of F (or $\boldsymbol{X}$) is the distribution function C of $(F_1(X_1), \ldots, F_d(X_d))$. The following proposition shows that a copula is invariant under strictly increasing transformations of the margins.

Proposition 2.1.4 Let $(X_1, \ldots, X_d)$ be a random vector with copula C and continuous margins. If $T_1, \ldots, T_d$ are strictly increasing functions, then the copula of the random vector $(T_1(X_1), \ldots, T_d(X_d))$ is also C.

Fréchet bounds are another important property of copulas. We give the bounds and the copulas that achieve the bounds in the following theorem and examples.

Theorem 2.1.5 (Fréchet Bounds) For every copula $C(u_1, \ldots, u_d)$,

$$\max\{\sum_{i=1}^{d} u_i + 1 - d,\, 0\} \leq C(\boldsymbol{u}) \leq \min\{u_1, \ldots, u_d\}. \tag{2.1.4}$$

We denote the lower and upper bounds by $W(u_1, \ldots, u_d)$ and $M(u_1, \ldots, u_d)$ respectively.

Example 2.1.6 Let $(X_1, \ldots, X_d)$ be a random vector with copula C and continuous margins. Here we give several fundamental copulas.

1. If $X_1, \ldots, X_d$ are independent, then $C(\boldsymbol{u}) = \prod_{i=1}^{d} u_i$, which is called the *independence copula*. We denote it by $\Pi(\boldsymbol{u})$ hereafter.

2. If $X_1, \ldots, X_d$ are perfectly positively dependent, i.e., $X_i = T_i(X_1)$, $i = 2, \ldots, d$ where T_i are almost surely strictly increasing functions, then $C(\boldsymbol{u}) = M(u_1, \ldots, u_d)$ $= \min\{u_1, \ldots, u_d\}$, which is called the *comonotonicity copula*.

3. If X_1, X_2 are perfectly negatively dependent, i.e., $X_2 = T(X_1)$ where T is an almost surely strictly decreasing function, then $C(\boldsymbol{u}) = W(u_1, u_2) = \max\{u_1 + u_2 - 1,\, 0\}$, which is called the *countermonotonicity copula*. It is easy to see that we cannot extend this case to higher dimensions.

The survival copula is defined similarly. Consider a random vector $(X_1, \ldots, X_d)$ with continuous margins $F_1, \ldots, F_d$ and copula C. Observe that $\bar{F}_i(X_i) = 1 - F_i(X_i)$, $1 \leq i \leq d$, is also uniformly distributed over $[0, 1]$, thus

$$\hat{C}(u_1, \ldots, u_d) := \Pr\{\bar{F}_1(X_1) \leq u_1, \ldots, \bar{F}_d(X_d) \leq u_d\}$$

is a copula. $\hat{C}$ is called the *survival copula* of $(X_1, \ldots, X_d)$. The joint survival function of random vector $(X_1, \ldots, X_d)$ is expressed as

$$\bar{F}(t_1, \ldots, t_d) := \Pr\{X_1 > t_1, \ldots, X_d > t_d\} = \hat{C}(\bar{F}_1(t_1), \ldots, \bar{F}_d(t_d)), \quad (t_1, \ldots, t_d) \in \mathbb{R}^d.$$

It also follows that for any $(u_1, \ldots, u_d) \in [0, 1]^d$,

$$\bar{C}(u_1, \ldots, u_d) := \Pr\{F_1(X_1) > u_1, \ldots, F_d(X_d) > u_d\} = \hat{C}(1 - u_1, \ldots, 1 - u_d), \quad (2.1.5)$$

where $\bar{C}$ is the joint survival function of copula C.

2.1.2 Archimedean Copulas

Archimedean copulas are an important family of copulas and often used when modeling dependence in loss data. They have simple forms by definition, yet we can construct a variety of dependence structures using them. Many commonly used copulas belong to this family such as Clayton copulas and Gumbel copulas. In later chapters, we will see that many theorems can be easily applied on Archimedean copulas to derive simple formulas. Here we start with bivariate Archimedean copulas and then extend to higher dimensions.

Definition 2.1.7 (Pseudo-inverse) Let $\phi : [0, 1] \to [0, \infty]$ be a continuous strictly decreasing function. Define $\phi^{[-1]} : [0, \infty] \to [0, 1]$ as follows:

$$\phi^{[-1]}(t) = \begin{cases} \phi^{-1}(t), & 0 \le t \le \phi(0); \\ 0, & \phi(0) \le t \le \infty. \end{cases} \quad (2.1.6)$$

We say $\phi^{[-1]}$ is the *pseudo-inverse* of ϕ.

It is easy to see that $\phi^{[-1]}$ is continuous and decreasing on $[0, \infty]$. Particularly, $\phi^{[-1]}$ is strictly decreasing on $[0, \phi(0)]$. Moreover, if $\phi(0) = \infty$, then $\phi^{[-1]} = \phi$.

Theorem 2.1.8 (Bivariate Archimedean Copula) Let $\phi : [0,1] \to [0,\infty]$ be a continuous strictly decreasing function such that $\phi(1) = 0$ and $\phi^{[-1]}$ be the pseudo-inverse of ϕ from (2.1.6). Then

$$C(u_1, u_2) = \phi^{[-1]}(\phi(u_1) + \phi(u_2)) \tag{2.1.7}$$

is a copula if and only if ϕ is convex.

Copulas of form (2.1.7) are called bivariate *Archimedean copulas*. The function ϕ is called a *generator* of the Archimedean copula. In addition, if $\phi(0) = \infty$, then ϕ is a *strict* generator, in which case $\phi^{[-1]} = \phi$ and $C(u_1, u_2) = \phi^{[-1]}(\phi(u_1) + \phi(u_2))$ is called a *strict* Archimedean copula.

Proposition 2.1.9 Let C be a bivariate Archimedean copula with generator ϕ. Then C has the following properties:

1. C is symmetric: $C(u_1, u_2) = C(u_2, u_1)$ for any $u_1, u_2 \in [0,1]$;

2. C is associative: $C(C(u_1, u_2), u_3) = C(u_1, C(u_2, u_3))$ for any $u_1, u_2, u_3 \in [0,1]$;

3. For any positive constant c, $c\phi$ is also a generator of C.

Example 2.1.10 The following two copula families are special cases of the Archimedean copulas.

1. Let $\phi(t) = \frac{1}{\theta}(t^{-\theta} - 1)$, $\theta \geq -1$, $\theta \neq 0$. Then the *Clayton copula* is given by

$$C(u_1, u_2) = (\max\{u_1^\theta + u_2^\theta - 1,\ 0\})^{-1/\theta}. \tag{2.1.8}$$

   When $\theta > 0$, we have $C(u_1, u_2) = (u_1^\theta + u_2^\theta - 1)^{-1/\theta}$. In this case, the Clayton copula is a strict Archimedean copula. This type of copula is known as a *comprehensive copula*, because its lower and upper limits are the Fréchet bounds.

2. Let $\phi(t) = (-\ln t)^\theta$, $\theta \geq 1$. Then the *Gumbel copula* is given by

$$C(u_1, u_2) = \exp(-((-\ln u_1)^\theta + (-\ln u_2)^\theta)^{1/\theta}). \tag{2.1.9}$$

This copula is a strict Archimedean copula for any $\theta \geq 1$.

To construct a d-dimensional Archimedean copula, we use the same pattern as in (2.1.7). We also need the following definition and theorem to define a proper distribution function.

Definition 2.1.11 ([38]) A continuous function $h(t)$ is called *completely monotonic* on an interval $[a, b]$ if it satisfies

$$(-1)^k \frac{d^k}{dt^k} h(t) \geq 0 \tag{2.1.10}$$

for $k = 0, 1, 2, \ldots$ and $t \in (a, b)$.

Theorem 2.1.12 ([20]) Let

$$C(u_1, \ldots, u_d) := \phi^{-1}(\phi(u_1) + \cdots + \phi(u_d)) \tag{2.1.11}$$

where $\phi : [0, 1] \to [0, \infty]$ is a strict Archimedean copula generator. Then C is a d-dimensional (Archimedean) copula if and only if $\phi^{-1} : [0, \infty] \to [0, 1]$ is completely monotonic.

Now we generalize Example 2.1.10 to higher dimensional cases.

Example 2.1.13 1. Let $\phi(t) = \frac{1}{\theta}(t^{-\theta} - 1)$, $\theta > 0$. Then the d-dimensional Clayton copula is given by

$$C(\boldsymbol{u}) = \left(\sum_{i=1}^{d} u_i^\theta - d + 1 \right)^{-1/\theta}. \tag{2.1.12}$$

2. Let $\phi(t) = (-\ln t)^\theta$, $\theta \geq 1$. Then the d-dimensional Gumbel copula is given by

$$C(\boldsymbol{u}) = \exp\left(-\left(\sum_{i=1}^{d}(-\ln u_i)^\theta\right)^{1/\theta}\right). \tag{2.1.13}$$

Note that the lower bound for both of these copulas is the independence copula $\Pi(\boldsymbol{u}) = \prod_{i=1}^{d} u_i$. This occurs when ϕ^{-1} is completely monotonic.

2.2 Dependence Measures

In this section we introduce three kinds of dependence measures: linear correlation, rank correlation and tail dependence parameters. Linear correlation and rank correlation coefficients are only defined for two random variables. In higher dimensions, we need to use a correlation matrix which consists of the correlation coefficients of the respective random variables. Tail dependence parameters are defined as scalars for any d-dimensional space in terms of copulas. Lastly, we introduce tail dependence functions which allow us to analyze tail dependence structure in a more general manner. Tail dependence functions are the primary tool in this dissertation.

2.2.1 Linear Correlation

Let X_1, X_2 be two random variables. The *linear correlation coefficient* ρ of X_1 and X_2 is defined as:

$$\rho(X_1, X_2) = \frac{\mathrm{Cov}(X_1, X_2)}{\sqrt{\mathrm{Var}(X_1)\mathrm{Var}(X_2)}}. \tag{2.2.1}$$

Linear correlation is a symmetric dependence measure with range $[-1, 1]$. Linear correlation is invariant under strictly increasing linear transformations, but is not invariant under nonlinear transformations. In addition, The linear correlation coefficient is only defined when the variances of X_1 and X_2 are finite and nonzero.These are drawbacks for analyzing dependence structures of heavy tailed distributions. In many cases, the data follow a joint heavy tailed distribution and have an increasing trend in their means so that the variances are not finite. The marginal distributions of the data might not be explicitly accessible. Moreover, taking any increasing transformation on the variables should not change the dependence structure between them. Because of these restrictions, linear correlation is of limited use in extremal dependence analysis.

2.2.2 Rank Correlation

Rank correlation coefficients of a pair of random variables measure the extent to which one variable follows the other in an increasing/decreasing fashion. The coefficients have negative values if one variable increases while the other variable decreases. Unlike linear correlation, rank correlation depends only on the copula of a bivariate distribution and not on the margins. Rank correlation coefficients are invariant under strictly increasing transformations. We introduce the two most widely known rank correlation coefficients, Kendall's tau and Spearman's rho [23]. The dependence that they measure is known as concordance. Given two points $(x_1, x_2),\ (\tilde{x}_1, \tilde{x}_2) \in \mathbb{R}^2$, we say that they are *concordant* if $(x_1 - \tilde{x}_1)(x_2 - \tilde{x}_2) > 0$ and *discordant* if $(x_1 - \tilde{x}_1)(x_2 - \tilde{x}_2) < 0$.

Definition 2.2.1 1. Let random vectors $(X_1, X_2),\ (\tilde{X}_1, \tilde{X}_2)$ be independent and identically distributed (i.i.d.). Then *Kendall's tau* is given by

$$\rho_\tau(X_1, X_2) = \Pr\{(X_1 - \tilde{X}_1)(X_2 - \tilde{X}_2) > 0\} - \Pr\{(X_1 - \tilde{X}_1)(X_2 - \tilde{X}_2) < 0\}. \quad (2.2.2)$$

2. Let $(X_1, X_2),\ (\tilde{X}_1, \tilde{X}_2),\ (\hat{X}_1, \hat{X}_2)$ be i.i.d. random vectors. Then *Spearman's rho*

is given by

$$\rho_S(X_1, X_2) = 3(\Pr\{(X_1 - \tilde{X}_1)(X_2 - \hat{X}_2) > 0\} - \Pr\{(X_1 - \tilde{X}_1)(X_2 - \hat{X}_2) < 0\}). \tag{2.2.3}$$

Spearman's rho can also be given in terms of Kendall's tau and the linear correlation coefficient:

$$\rho_S(X_1, X_2) = \rho(F_1(X_1), F_2(X_2)), \tag{2.2.4}$$

where F_1, F_2 are the marginal distribution functions of (X_1, X_2).

Kendall's tau and Spearman's rho share many properties. They are both symmetric dependence measures with range $[-1, 1]$. It is necessary that they both have value zero for independent random variables, but not sufficient. The following proposition shows that they depend only on the copula of two variables, which implies that they are invariant under strictly increasing transformations.

Proposition 2.2.2 Let (X_1, X_2) be a random vector with copula C and continuous margins. Then we have

$$\rho_\tau(X_1, X_2) = 4\int_0^1 \int_0^1 C(u_1, u_2)\, dC(u_1, u_2) - 1, \tag{2.2.5}$$

$$\rho_S(X_1, X_2) = 12\int_0^1 \int_0^1 C(u_1, u_2) - u_1 u_2 \; du_1 du_2. \tag{2.2.6}$$

Before introducing tail dependence parameters we reiterate that both linear correlation and rank correlation are ill-suited to the task of extremal dependence analysis. This is primarily because they do not scale well to higher dimensions.

2.2.3 Tail Dependence Parameters

Tail dependence parameters describe the amount of dependence in the upper (or lower) tail of a multivariate distribution. They measure the interactivity (e.g. possibility of contagion) among the multivariate margins of the random vector and play an important role in analyzing extremal dependence. Like the rank correlation coefficients, tail dependence parameters depend only on the copula of the random vector and thus are invariant under strictly increasing transformations. We will look at the bivariate case first. Then we give the definition for high dimensional cases.

Definition 2.2.3 Let (X_1, X_2) be a random vector with continuous margins F_1, F_2. Then the *upper tail dependence parameter* of X_1 and X_2 is given by

$$\lambda_U := \lim_{u \to 1^-} \Pr\{F_2(X_2) > u \mid F_1(X_1) > u\}, \tag{2.2.7}$$

and the *lower tail dependence parameter* of X_1 and X_2 is given by

$$\lambda_L := \lim_{u \to 0^+} \Pr\{F_2(X_2) \le u \mid F_1(X_1) \le u\}, \tag{2.2.8}$$

provided the limits exist.

The range of both λ_U and λ_L is $[0, 1]$. If λ_U (or λ_L) is zero, then we say that X_1 and X_2 are *asymptotically independent* in the upper (lower) tail. To illustrate that the tail dependence parameters depend only on the bivariate copula, we have the following theorem.

Theorem 2.2.4 Let (X_1, X_2) be a random vector with continuous margins F_1, F_2 and copula C. Then the tail dependence parameters of X_1 and X_2 can be expressed in terms of their bivariate copula:

$$\lambda_U = 2 - \lim_{u \to 1^-} \frac{1 - C(u,u)}{1-u} = \lim_{u \to 0^+} \frac{\hat{C}(u,u)}{u}, \tag{2.2.9}$$

$$\lambda_L = \lim_{u \to 0^+} \frac{C(u,u)}{u}, \tag{2.2.10}$$

where $\hat{C}$ is the survival copula of C.

Example 2.2.5 (Bivariate Gaussian Copula) Let (X_1, X_2) have the standard bivariate normal distribution and (linear) correlation ρ. Then

$$\Pr(X \leq x, Y \leq y) = C(\Phi(x), \Phi(y)),$$

where C is the bivariate Gaussian copula. Since C is exchangeable, i.e., $C(u_1, u_2) = C(u_2, u_1)$, applying L'Hopital's rule we have

$$\lambda = \lambda_U = \lambda_L = 2 \lim_{u \to 0^+} \Pr(X_2 \leq \Phi^{-1}(u) \mid X_1 = \Phi^{-1}(u)).$$

It is known that $X_2 \mid X_1 = x \sim \mathcal{N}(\rho x, 1 - \rho^2)$. Hence when $\rho < 1$ we have

$$\lambda = 2 \lim_{x \to \infty} \left(1 - \Phi\left(x \frac{\sqrt{1-\rho}}{\sqrt{1+\rho}}\right)\right) = 0.$$

Therefore, the Gaussian copula is asymptotically independent in both the upper and lower tails. In statistical modeling, whenever normal distributions are fitted to the data, the Gaussian copula is inherent and extremal dependence can not be considered. In most cases, however, extremal dependence exists and should not be neglected. It has been said that the use of the Gaussian copula in credit derivatives was the primary cause of the financial crisis in 2008. Whitehouse stated in [37] that "The Gauss-copula is the worst invention ever for credit risk management".

Example 2.2.6 (Bivariate t Copula) Let (X_1, X_2) have the bivariate t-distribution. Then we have the following stochastic representation

$$(X_1, X_2) =_d (\mu_1, \mu_2) + \frac{\sqrt{\nu}}{\sqrt{S}}(Z_1, Z_2),$$

where $S \sim \chi^2_\nu$ and $(Z_1, Z_2) \sim \mathcal{N}_2(\mathbf{0}, \Sigma)$ are independent. Since the t copula is exchangeable, using an argument similar to Example 2.2.5 we have

$$\lambda = -2t_{\nu+1}\sqrt{\frac{(\nu+1)(1-\rho)}{1+\rho}},$$

when $\rho > -1$. Hence the t copula is asymptotically dependent in both the upper and lower tails. It is more reasonable to use the t copula than the Gaussian copula for data that show extremal dependence.

In the following definition we define tail dependence parameters and the extremal dependence parameter for any d-dimensional random vectors. This is a generalization of the bivariate case. Detailed discussion can be found in [25, 33].

Definition 2.2.7 Let $\boldsymbol{X} = (X_1, \ldots, X_d)$ be a random vector with continuous margins $F_1, \ldots, F_d$ and copula C.

1. $\boldsymbol{X}$ is said to be *upper-orthant tail dependent* if for some subset $\emptyset \neq J \subset I$, the following limit exists and is positive.

$$\tau_J = \lim_{u \to 1^-} \Pr\{F_j(X_j) > u, \forall j \in I \backslash J \mid F_i(X_i) > u, \forall i \in J\} > 0. \tag{2.2.11}$$

If for all $\emptyset \neq J \subset I$, $\tau_J = 0$, then we say $\boldsymbol{X}$ is upper-orthant tail independent.

2. $\boldsymbol{X}$ is said to be *upper extremal dependent* if the following limit exists and is positive.

$$\gamma = \lim_{u \to 1^-} \Pr\{F_j(X_j) > u, \forall j \in I \mid F_i(X_i) > u, \exists i \in I\} > 0. \tag{2.2.12}$$

If $\gamma = 0$, then we say $\boldsymbol{X}$ is upper extremal independent.

The limits τ_J are called the *upper tail dependence parameters*. γ is called the *upper extremal dependence parameter*. The lower tail dependence of a copula is defined as the upper tail dependence of its survival copula in (2.1.5). Obviously, tail dependence is a copula property; these parameters do not depend on the marginal distributions. If $X_1, \dots, X_n$ are independent, then the corresponding upper tail (extremal) dependence parameters are all zeros. Clearly, $\tau_J \geq \gamma$, for all nonempty $J \subset I$. Thus, the extremal dependence parameter provides a lower bound for orthant tail dependence parameters. The multivariate tail dependence parameters τ_J's have been used in [16] to analyze the contagion risk in banking systems, and the extremal dependence parameter γ in the bivariate case has been used in [7] to analyze the extremal dependence in financial return data.

2.2.4 Tail Dependence Functions

Tail dependence functions are a more general tool for analyzing extremal dependence. They can measure the dependence when the components approach to the tails at various speeds.

Let F be the distribution function of a d-dimensional random vector $\boldsymbol{X} = (X_1, \dots, X_d)$ with continuous margins $F_1, \dots, F_d$ and copula C. The *lower and upper tail dependence functions*, denoted by $b(\cdot; C)$ and $b^*(\cdot; C)$ respectively, are introduced in [21, 30, 19] as follows,

$$b(\boldsymbol{w}; C) := \lim_{u \to 0^+} \frac{C(uw_j, \forall j \in I)}{u}, \ \forall \boldsymbol{w} = (w_1, \dots, w_d) \in \mathbb{R}_+^d;$$

$$b^*(\boldsymbol{w}; C) := \lim_{u \to 0^+} \frac{\overline{C}(1 - uw_j, \forall j \in I)}{u}, \ \forall \boldsymbol{w} = (w_1, \dots, w_d) \in \mathbb{R}_+^d, \quad (2.2.13)$$

provided that the limits exist. Since $b(\boldsymbol{w};\hat{C}) = b^*(\boldsymbol{w};C)$ where $\hat{C}(u_1,\dots,u_d) = \overline{C}(1-u_1,\dots,1-u_d)$ is the survival copula in (2.1.5), this dissertation focuses only on upper tail dependence. The explicit expression of b^* for elliptical distributions was obtained in [21]. A theory of tail dependence functions was developed in [19, 30] based on Euler's homogeneous representation:

$$b^*(\boldsymbol{w};C) = \sum_{j=1}^{d} w_j t_j(w_i, i \neq j \mid w_j),\ \forall \boldsymbol{w} = (w_1,\dots,w_d) \in \mathbb{R}_+^d, \tag{2.2.14}$$

where

$$t_j(w_i, i \neq j \mid w_j) := \lim_{u \to 0^+} \Pr\{F_i(X_i) > 1 - w_i u,\ \forall i \neq j \mid F_j(X_j) = 1 - w_j u\},\quad j \in I.$$

The t_j's are called *upper conditional tail dependence functions*. For copulas with explicit expressions, the tail dependence functions are obtained directly from the copulas with relative ease. For copulas without explicit expressions, the tail dependence functions can be obtained from (2.2.14) by exploring closure properties of the related conditional distributions. In [30], for example, the tail dependence function of the multivariate t distribution is obtained by (2.2.14).

If follows from (2.2.11) and (2.2.13) that the upper tail dependence parameters can be expressed as

$$\tau_J = \frac{b^*(1,\dots,1;C)}{b^*(1,\dots,1;C_J)},\ \text{for all } \emptyset \neq J \subset I, \tag{2.2.15}$$

where C_J is the multivariate margin of C with component indices in J. It is shown in [19] that $b^*(\boldsymbol{w};C) > 0$ for all $\boldsymbol{w} \in \mathbb{R}_+^d$ if and only if $b^*(1,\dots,1;C) > 0$. Unlike the τ_J, however, the tail dependence function provides all the extremal dependence information of the copula C as specified by its *extreme value copula* (EV copula). The upper EV copula of C, denoted by C^{UEV}, is defined as $C^{UEV}(u_1,\dots,u_d) := \lim_{n\to\infty} C^n(u_1^{1/n},\dots,u_d^{1/n})$ for any $(u_1,\dots,u_d) \in [0,1]^d$ if the limit exists [18].

Define the upper *exponent function* $a^*(\cdot\,; C)$ of copula C as

$$a^*(\boldsymbol{w}; C) := \sum_{S\subseteq I, S\neq\emptyset} (-1)^{|S|-1} b_S^*(w_i, i\in S; C_S), \tag{2.2.16}$$

where $b_S^*(w_i, i\in S; C_S)$ denotes the upper tail dependence function of the margin C_S of C with component indices in S. It follows from (2.2.12) and (2.2.16) that the extremal dependence parameter can be expressed as

$$\gamma = \frac{b^*(1,\dots,1; C)}{a^*(1,\dots,1; C)}. \tag{2.2.17}$$

Similar to tail dependence functions, the exponent function has the following homogeneous representation:

$$a^*(\boldsymbol{w}; C) = \sum_{j=1}^{d} w_j \bar{t}_j(w_i, i\neq j \mid w_j),\ \forall \boldsymbol{w} = (w_1,\dots,w_d)\in \mathbb{R}_+^d, \tag{2.2.18}$$

where

$$\bar{t}_j(w_i, i\neq j \mid w_j) = \lim_{u\to 0^+} \Pr\{F_i(X_i)\le 1-w_iu, \forall i\neq j \mid F_j(X_j) = 1-w_ju\},\quad j\in I.$$

It is shown in [19] that tail dependence functions $\{b_S^*(w_i, i\in S; C_S)\}$ and the exponent function $a^*(\boldsymbol{w}; C)$ are uniquely determined by each other.

Theorem 2.2.8 ([19], [30]) Let C be a d-dimensional copula. Then the upper EV copula

$$C^{UEV}(u_1,\dots,u_d) = \exp\{-a^*(-\log u_1,\dots,-\log u_d)\},\ \forall (u_1,\dots,u_d)\in[0,1]^d$$

where a^* is given by either (2.2.16) or (2.2.18).

The following examples given in [19] show that the tail dependence functions of Archimedean copulas have very simple forms. More detail can be found in [15, 19].

Example 2.2.9 1. Consider a d-dimensional Clayton copula given in (2.1.12). The margin of the last $d-k$ variables has a similar copula form:

$$C(u_{k+1}, \ldots, u_d) = (\sum_{i=k+1}^{d} u_i^{-\theta} - (d-k-1))^{-1/\theta}. \tag{2.2.19}$$

It is shown that the lower tail dependence function is

$$b(w_1, \ldots, w_d) = (\sum_{i=1}^{d} w_i^{-\theta})^{-1/\theta}. \tag{2.2.20}$$

2. If a d-dimensional Archimedean copula given in (2.1.11) has generator ϕ and there exists $\alpha > 0$ such that

$$\frac{\phi^{-1}(st)}{\phi^{-1}(t)} \to s^{-\alpha} \text{ as } t \to \infty \text{ for any } s > 0, \tag{2.2.21}$$

then we have

$$b(\boldsymbol{w}) = (\sum_{i=1}^{d} w_i^{-1/\alpha})^{-\alpha} \quad \text{and} \quad a^*(\boldsymbol{w}) = (\sum_{i=1}^{d} w_i^{\alpha})^{1/\alpha}. \tag{2.2.22}$$

In the next chapter we will see that (2.2.21) is equivalent to saying that ϕ^{-1} is regularly varying at ∞ or that ϕ is regularly varying at 1.

2.2.5 Elliptical Distributions

In this section we introduce a large family of radially symmetric multivariate distributions – elliptical distributions. Some well known distributions such as the normal, t-, Laplace, Cauchy and logistic distributions belong to this family. In general, elliptical distributions do not have explicit copulas. However, they play an important role in analyzing tail dependence. Elliptical distributions are defined in terms of spherical distributions. More detail of the properties of spherical and elliptical distributions can be found in [11, 28].

Definition 2.2.10 A random vector $\boldsymbol{X} = (X_1, \ldots, X_d)$ is said to have a *spherical distribution* if $U\boldsymbol{X} \stackrel{d}{=} \boldsymbol{X}$ for any orthogonal matrix $U \in \mathbb{R}^{d\times d}$ satisfying $UU^T = U^TU = I_d$.

If $\boldsymbol{X}$ has a spherical distribution, then there exists a function $\psi : \mathbb{R} \to \mathbb{R}$ such that for all $\boldsymbol{t} \in \mathbb{R}^d$ we have

$$\phi_{\boldsymbol{X}}(\boldsymbol{t}) = \mathrm{E}(e^{i\boldsymbol{t}^T\boldsymbol{X}}) = \psi(\boldsymbol{t}^T\boldsymbol{t}) = \psi(t_1^2 + \cdots + t_d^2), \tag{2.2.23}$$

where $\phi_{\boldsymbol{X}}$ is the characteristic function of $\boldsymbol{X}$. The function ψ is known as the *characteristic generator* of the spherical distribution. We use the notation $\boldsymbol{X} \sim S_d(\psi)$ for d-dimensional spherically distributed random vectors hereafter. The following theorem characterizes spherical distributions.

Theorem 2.2.11 $\boldsymbol{X} \sim S_d(\psi)$ if and only if $\boldsymbol{X}$ has the stochastic representation

$$\boldsymbol{X} \stackrel{d}{=} R\boldsymbol{U}, \tag{2.2.24}$$

where $\boldsymbol{U}$ is uniformly distributed on the unit sphere (with respect to Euclidean distance) in $\mathbb{R}^d$ and $R \geq 0$ is a random variable independent of $\boldsymbol{U}$.

Example 2.2.12 Multivariate normal distributions are spherical. Suppose $\boldsymbol{X}$ has the multivariate normal distribution $\mathcal{N}(\boldsymbol{0}, I_d)$. Then

$$\phi_{\boldsymbol{X}}(\boldsymbol{t}) = \mathrm{E}(e^{i\boldsymbol{t}^T\boldsymbol{X}}) = e^{-\frac{1}{2}\boldsymbol{t}^T\boldsymbol{t}}.$$

Hence $\boldsymbol{X} \sim S_d(\psi)$ where $\psi(t) = e^{-\frac{1}{2}t}$.

Definition 2.2.13 If a random vector $\boldsymbol{X}$ has the following stochastic representation

$$\boldsymbol{X} \stackrel{d}{=} \boldsymbol{\mu} + A\boldsymbol{Y},$$

where $\boldsymbol{\mu} \in \mathbb{R}^d$ is a constant, A is a $d \times k$ matrix and $\boldsymbol{Y} \sim S_k(\psi)$, then $\boldsymbol{X}$ is said to have an *elliptical distribution*. We denote by $\boldsymbol{X} \sim \mathcal{E}_d(\boldsymbol{\mu}, \Sigma, \psi)$, where $\Sigma = AA^T$.

Like spherical distributions, elliptical distributions have an important stochastic representation as well. This characterization is particularly useful when constructing elliptically distributed random vectors.

Theorem 2.2.14 $\boldsymbol{X} \sim \mathcal{E}_d(\boldsymbol{\mu}, \Sigma, \psi)$ if and only if there exist R, A and $\boldsymbol{U}$ such that

$$\boldsymbol{X} \stackrel{d}{=} \boldsymbol{\mu} + RA\boldsymbol{U}, \tag{2.2.25}$$

where $R \geq 0$ is a random variable independent of $\boldsymbol{U}$, A is a $d \times k$ matrix such that $AA^T = \Sigma$ and $\boldsymbol{U}$ is uniformly distributed on the unit sphere in $\mathbb{R}^k$.

Linear combinations of elliptical random vectors are still elliptical, with the same characteristic generator. If $\boldsymbol{X} \sim \mathcal{E}_d(\boldsymbol{\mu}, \Sigma, \psi)$, then for any $B \in \mathbb{R}^{k\times d}$ and $\boldsymbol{c} \in \mathbb{R}^d$ we have

$$B\boldsymbol{X} + \boldsymbol{c} \sim \mathcal{E}_k(B\boldsymbol{\mu} + \boldsymbol{c}, B\Sigma B^T, \psi). \tag{2.2.26}$$

Additionally, the marginal distributions of an elliptical distribution are elliptical with the same characteristic generator as well. Consider a partition $\boldsymbol{X} = (\boldsymbol{X}_1^T, \boldsymbol{X}_2^T)^T$, where $\boldsymbol{X}_1 = (X_1, \ldots, X_k)^T$ and $\boldsymbol{X}_2 = (X_{k+1}, \ldots, X_d)^T$, $2 \leq k \leq d-1$. Similarly we can partition $\boldsymbol{\mu}$ and Σ:

$$\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}.$$

Then we have that $\boldsymbol{X}_1 \sim \mathcal{E}_k(\boldsymbol{\mu}_1, \Sigma_{11}, \psi)$ and $\boldsymbol{X}_2 \sim \mathcal{E}_{d-k}(\boldsymbol{\mu}_2, \Sigma_{22}, \psi)$.

The tail properties of elliptical distributions are closely connected with the regular variation of the random variable R in the representation (2.2.25). The relationship will be given in Example 3.2.4 in the next chapter. The upper tail dependence function of a multivariate elliptical distribution is derived explicitly in [21]. We will discuss the multivariate upper tail dependence and exponent functions, upper tail and extremal dependence parameters for elliptical distributions in Section 4.1.

Chapter 3

Multivariate Regular Variation

The theory of regularly varying functions is an important analysis tool for discussing the tail dependence and properties of heavy tailed distributions. It is also closely related to extreme value theory. This chapter discusses multivariate regular variation and its tail dependence. In Section 3.1, we review the theory of regular variation. In Section 3.2, the main theorem in Section 3.1 is made applicable to higher dimensions. The primary resources for Section 3.1 and 3.2 are [6, 28, 32]. More detail and further references can be found there. In Section 3.3, the tail dependence of multivariate regularly varying distributions is analyzed by the intensity measure, which provides a link between tail analysis and multivariate regular variation.

3.1 Regular Variation

A real function $H : \mathbb{R}_+ \to \mathbb{R}_+$ is said to be *regularly varying* at ∞ with index $\rho \in \mathbb{R}$ (written $H \in \mathbf{RV}_\rho$) if

$$\lim_{t\to\infty} \frac{H(ct)}{H(t)} = c^\rho.$$

ρ is called the *exponent of variation*.

If $\rho = 0$, H is a *slowly varying* function, usually denoted by L; that is, L is a positive function on $(0, \infty)$ with property

$$\lim_{t\to\infty} \frac{L(ct)}{L(t)} = 1, \text{ for every } c > 0. \tag{3.1.1}$$

In this dissertation, we consider distributions with regular varying right tails,

$$\overline{F}_i(t) := \Pr\{X_i > t\} = t^{-\beta} L(t),\ t \geq 0, \tag{3.1.2}$$

where $\beta > 0$ is the marginal *heavy-tail index*.

Generally speaking, regularly varying functions behave asymptotically like power functions. Additionally, $\mathbf{RV}_0$ is the set of slowly varying functions. If $H \in \mathbf{RV}_\infty$, then

$$\lim_{t\to\infty} \frac{H(ct)}{H(t)} = c^\infty = \begin{cases} 0, & \text{if } 0 < c < 1; \\ 1, & \text{if } c = 1; \\ \infty, & \text{if } c > 1. \end{cases}$$

Similarly, $H \in \mathbf{RV}_{-\infty}$ means

$$\lim_{t\to\infty} \frac{H(ct)}{H(t)} = c^{-\infty} = \begin{cases} \infty, & \text{if } 0 < c < 1; \\ 1, & \text{if } c = 1; \\ 0, & \text{if } c > 1. \end{cases}$$

3.1.1 Maximum Domains of Attraction

In this section, we illustrate the relationship between extreme value theory and regular variation. More information about extreme value theory can be found in [8, 28].

Consider a sequence of i.i.d. random variables $X_n,\ n = 1, 2, \ldots$ with a common distribution function F. These variables may be interpreted as various kinds of financial losses in risk management. The most traditional models in extreme value theory are concerned with the *block maxima* $M_n := \max(X_1, \ldots, X_n)$. Suppose that M_n converges in distribution under an appropriate normalization. Then there exist two sequences of real constants $\{c_n\}$ and $\{d_n\}$, where $c_n > 0$ for all n, such that

$$\lim_{n\to\infty} \Pr\{(M_n - d_n)/c_n \leq x\} = \lim_{n\to\infty} F^n(c_n x + d_n) = H(x) \tag{3.1.3}$$

for some non-degenerate distribution function $H(x)$.

Definition 3.1.1 (Maximum Domain of Attraction) If (3.1.3) holds for some non-degenerate d.f. H, then F is said to be in the *maximum domain of attraction* of H, written $F \in \mathrm{MDA}(H)$.

Theorem 3.1.2 If $F \in \mathrm{MDA}(H)$ for some non-degenerate d.f. $H(x)$, then H must be a *generalized extreme value* (GEV) distribution H_θ which is defined as

$$H_\theta(x) := \begin{cases} \exp(-(1+\theta x)^{-1/\theta}), & \text{if } \theta \neq 0; \\ \exp(-e^{-x}), & \text{if } \theta = 0. \end{cases}$$

where $1+\theta x > 0$. The parameter θ is called the *shape* parameter of the GEV distribution.

There are three types of GEV distributions specified by different names according to the value of θ. When $\theta > 0$ the distribution is a Fréchet distribution; distributions in this class include the t, F, inverse gamma and loggamma distributions. When $\theta = 0$ it is a Gumbel distribution; examples include the gamma, chi-squared, standard Weibull and Gumbel itself. When $\theta < 0$ it is a Weibull distribution; the beta distribution belongs to this class.

Example 3.1.3 (Pareto Distribution) Assume that $\{X_n\}$ have Pareto distribution with $F(x) = 1 - \frac{1}{(1+x)^\alpha}$, $\alpha > 0$, $x \geq 0$. Choose the normalizing sequences to be $c_n = \frac{1}{\alpha} n^{1/\alpha}$ and $d_n = n^{1/\alpha} - 1$. Then from (3.1.3) we have

$$F^n(c_n x + d_n) = \left(1 - \frac{1}{n}\left(1 + \frac{x}{\alpha}\right)^{-\alpha}\right)^n, \ 1 + \frac{x}{\alpha} \geq n^{-1/\alpha}.$$

Hence

$$\lim_{n\to\infty} F^n(c_n x + d_n) = \exp\left(-\left(1 + \frac{x}{\alpha}\right)^{-\alpha}\right), \ 1 + \frac{x}{\alpha} \geq 0.$$

This shows that $F \in \mathrm{MDA}(H_{1/\alpha})$.

The limiting behavior of minima can be obtained via maxima using the identity

$$\min(X_1, \ldots, X_n) = -\max(-X_1, \ldots, -X_n). \tag{3.1.4}$$

It is shown that if the d.f. of $-X_n$, $1-F(-x)$, is in $\text{MDA}(H_\theta)$, then the GEV distribution for minima is $1 - H_\theta(-x)$.

The following theorem shows the relationship between different types of extreme value distributions and regularly varying functions. There are only results for the Fréchet and Weibull classes. The characterization of Gumbel class is more complicated.

Theorem 3.1.4 1. (Fréchet MDA) For $\theta > 0$,

$$F \in \text{MDA}(H_\theta) \Longleftrightarrow \bar{F}(x) = x^{-1/\theta} L(x), \tag{3.1.5}$$

for some L slowly varying at ∞.

2. (Weibull MDA) For $\theta < 0$,

$$F \in \text{MDA}(H_{1/\theta}) \Longleftrightarrow x_F < \infty \text{ and } \bar{F}(x_F - x^{-1}) = x^{1/\theta} L(x), \tag{3.1.6}$$

where $x_F := \sup\{x \in \mathbb{R} : F(x) < 1\}$ is the *right endpoint* and L is some slowly varying function at ∞.

In Example 3.1.3, the survival function of Pareto distribution can be written as $\bar{F}(x) = \frac{1}{(1+x)^\alpha} = x^{-\alpha} L(x)$, where $L(x) = (1 + \frac{1}{x})^{-\alpha}$ is a slowly varying function. Thus we can also use Theorem 3.1.4 to prove that $F \in \text{MDA}(H_{1/\alpha})$.

3.1.2 Properties of Regularly Varying Functions

Now we give some integral and differential properties of regularly varying functions, assuming that all functions are locally integrable. We can see that a regularly varying function with index ρ behaves roughly similar to x^ρ. The proofs can be found in [32].

Theorem 3.1.5 (Karamata's Theorem)

1. Let $\rho \geq -1$ and $H \in \mathbf{RV}_\rho$. Then $\int_0^x H(t)dt \in \mathbf{RV}_{\rho+1}$ and

$$\lim_{x\to\infty} \frac{xH(x)}{\int_0^x H(t)dt} = \rho + 1. \tag{3.1.7}$$

   If $\rho < -1$ and $H \in \mathbf{RV}_\rho$, then $\int_x^\infty H(t)dt \in \mathbf{RV}_{\rho+1}$ and

$$\lim_{x\to\infty} \frac{xH(x)}{\int_x^\infty H(t)dt} = -\rho - 1. \tag{3.1.8}$$

2. If H satisfies that

$$\lambda := \lim_{x\to\infty} \frac{xH(x)}{\int_0^x H(t)dt} \in \mathbb{R}_+, \tag{3.1.9}$$

   then $H \in \mathbf{RV}_{\lambda-1}$. If H satisfies that

$$\lambda := \lim_{x\to\infty} \frac{xH(x)}{\int_x^\infty H(t)dt} \in \mathbb{R}_+, \tag{3.1.10}$$

   then $H \in \mathbf{RV}_{-\lambda-1}$.

Corollary 3.1.6 (The Karamata Representation)

1. A function L is slowly varying if and only if L has the following form,

$$L(x) = c(x) \exp\left\{\int_1^x \frac{\varepsilon(t)}{t} dt\right\}, \ x > 0, \tag{3.1.11}$$

   where $c, \varepsilon : \mathbb{R}_+ \to \mathbb{R}_+$ such that $\lim_{x\to\infty} c(x)$ is finite and positive *and* $\lim_{t\to\infty} \varepsilon(t) = 0$.

2. A function H is regularly varying with index ρ if and only if H has the representation below,

$$H(x) = c(x) \exp\left\{\int_1^x \frac{\rho(t)}{t} dt\right\}, \ x > 0, \tag{3.1.12}$$

   where $\lim_{x\to\infty} c(x)$ is finite and positive *and* $\lim_{t\to\infty} \rho(t) = \rho$.

Theorem 3.1.7 Let $H : \mathbb{R}_+ \to \mathbb{R}_+$ be absolutely continuous with density h so that $H(x) = \int_0^x h(t)dt$.

1. If H satisfies
$$\lim_{x\to\infty} \frac{xh(x)}{H(x)} = \rho, \tag{3.1.13}$$
then $H \in \mathbf{RV}_\rho$.

2. If $H \in \mathbf{RV}_\rho$, $\rho \in \mathbb{R}$, and h is monotone, then (3.1.13) holds. In addition, if $\rho \neq 0$, then $|h| \in \mathbf{RV}_{\rho-1}$.

The following proposition provides further properties of regularly varying functions.

Proposition 3.1.8

1. Suppose $H \in \mathbf{RV}_\rho$, $\rho \in \overline{\mathbb{R}}$. Then $\lim_{x\to\infty} \frac{H(x)}{\log x} = \rho$ and
$$\lim_{x\to\infty} H(x) = \begin{cases} 0, & \text{if } \rho < 0 \\ \infty, & \text{otherwise.} \end{cases}$$

2. (Potter bounds) Suppose $H \in \mathbf{RV}_\rho$, $\rho \in \mathbb{R}$. For any $\varepsilon > 0$, there exists t_0 such that when $x \geq 1$ and $t \geq t_0$ we have
$$(1-\varepsilon)x^{\rho-\varepsilon} < \frac{H(tx)}{H(t)} < (1+\varepsilon)x^{\rho+\varepsilon}. \tag{3.1.14}$$

3. Suppose $H \in \mathbf{RV}_\rho$, $\rho \in \mathbb{R}$. Let $\{a_n\}$ and $\{b_n\}$ be two positive sequences such that $\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = \infty$. For $c \in \mathbb{R}_+$, if $b_n \sim ca_n$ as $n \to \infty$, then $H(b_n) \sim c^\rho H(a_n)$ as $n \to \infty$.

4. Suppose $H_1 \in \mathbf{RV}_{\rho_1}$ and $H_2 \in \mathbf{RV}_{\rho_2}$, $\rho_2 < \infty$. If $\lim_{x\to\infty} H_2(x) = \infty$, then $H_1 \circ H_2 \in \mathbf{RV}_{\rho_1\rho_2}$.

5. Suppose $H \in \mathbf{RV}_\rho$ is an increasing function with $H(\infty) = \infty$, $\rho \in [0, \infty]$. Then $H^{\leftarrow} \in \mathbf{RV}_{\rho^{-1}}$.

6. Suppose $H_1, H_2 \in \mathbf{RV}_\rho$ are increasing, $\rho \in \mathbb{R}_+$. Then for $c \in [0, \infty]$, $H_1(x) \sim cH_2(x)$ as $x \to \infty$ if and only if $H_1^{\leftarrow}(x) \sim c^{-1/\rho} H_2^{\leftarrow}(x)$ as $x \to \infty$.

7. Suppose $H \in \mathbf{RV}_\rho$, $\rho \neq 0$. Then there exists an absolutely continuous and strictly monotone function H^* such that $H(x) \sim H^*(x)$ as $x \to \infty$.

3.1.3 Vague Convergence

In this section we introduce the weak convergence of probability measures on metric spaces. Since the results will be used on random vectors, we use $\mathbb{R}^d$ as the metric space and give the definitions. In general, $\mathbb{R}^d$ can be replaced by any locally compact topological space with a countable base.

Definition 3.1.9 (Radon Measure on $\mathbb{R}^d$) A *measure* $\mu : \mathcal{E} \to [0, \infty]$ is a map defined on $\mathcal{E} \subset \mathbb{R}^d$ such that

1. $\mu(\emptyset) = 0$ and $\mu(A) \geq 0$ for all $A \in \mathcal{E}$.

2. If $\{A_n, n = 1, 2, \ldots\}$ are mutually exclusive sets in $\mathcal{E}$, then the σ-additivity property holds:

$$\mu\left(\cup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i).$$

The measure μ is called *Radon* if $\mu(K) < \infty$ for any compact set $K \subset \mathbb{R}^d$.

Define $\mathcal{M}_+(\mathbb{R}^d) := \{\mu : \mu \text{ is a non-negative Radon measure on } \mathcal{E}\}$, and $\mathcal{C}_K^+(\mathbb{R}^d) := \{f : f \text{ is continuous and positive with compact support on } \mathbb{R}^d\}$, where a function with compact support vanishes outside of a compact set.

Definition 3.1.10 Suppose that $\mu_n \in \mathcal{M}_+(\mathbb{R}^d)$, $n \in \mathbb{N}$. If for all $f \in \mathcal{C}_K^+(\mathbb{R}^d)$ we have

$$\mu_n(f) := \int_{\mathbb{R}^d} f(x)\mu_n(dx) \to \mu_0(f) := \int_{\mathbb{R}^d} f(x)\mu_0(dx) \quad \text{as } n \to \infty, \tag{3.1.15}$$

then we say that μ_n converges *vaguely* to μ_0, written $\mu_n \xrightarrow{v} \mu_0$.

For the regular variation of distribution functions, the following theorem is important. The generalization to higher dimensions will be given in the next section.

Theorem 3.1.11 Let X be a non-negative random variable with d.f. F. The following statements are equivalent:

1. $\bar{F} \in \mathbf{RV}_{-\alpha}$, where $\alpha > 0$ is the heavy-tail index.
2. There exists a sequence $\{b_n\}$ with $b_n \to \infty$ such that

$$n\bar{F}(b_n x) \to x^{-\alpha},\ x > 0, \quad \text{as } n \to \infty. \tag{3.1.16}$$

3. There exists a sequence $\{b_n\}$ with $b_n \to \infty$ such that

$$\mu_n(\cdot) := n \Pr\left\{\frac{X}{b_n} \in \cdot\right\} \xrightarrow{v} \nu_\alpha(\cdot), \quad \text{as } n \to \infty \tag{3.1.17}$$

in $\mathcal{M}_+(0, \infty]$, where $\nu_\alpha(x, \infty] = x^{-\alpha}$.

3.2 Multivariate Regular Variation

Here we discuss the regular variation of functions and measures in high dimensions. The goal is to generalize the results of Theorem 3.1.11.

A set $C \in \mathbb{R}^d$ is called a *cone* if for any $\boldsymbol{x} \in C$ and any $t > 0$, $t\boldsymbol{x} \in C$. Suppose that $\mathbf{1} = (1, \ldots, 1) \in C$ and $h \geq 0$ is a measurable function defined on C. Then h is said to be *multivariate regularly varying (MRV)* with limit function $\lambda(\boldsymbol{x}) > 0$ if

$$\lim_{t\to\infty} \frac{h(t\boldsymbol{x})}{h(t\mathbf{1})} = \lambda(\boldsymbol{x}), \text{ for all } \boldsymbol{x} \in C. \tag{3.2.1}$$

One of the restriction arguments for regular variation is that the sets involved (cones) need to be bounded away from the origin. This is satisfied when we are focusing on the tail probabilities, which consider the probability of a neighborhood of infinity. However, we also need the neighborhoods of infinity to be relatively compact sets (with no mass on the boundary). By means of the *one-point uncompactification*, we can make semi-infinite intervals compact. Suppose that $\mathbb{E}$ is a compact set and $x \in \mathbb{E}$. The compact sets of $\mathbb{E}\backslash\{x\}$ are the compact sets $K \subset \mathbb{E}$ such that $x \notin K$. In the following theorem, we use the punctured space with a one-point uncompactification $\mathbb{R}^d\backslash\{0\}$. More detail and equivalent properties are given in [13, 31, 32].

Theorem 3.2.1 (Multivariate Regular Variation) Let F be the distribution function of a random vector $\boldsymbol{X}$, then the following statements are equivalent:

1. F is an MRV distribution.

2. There exists a Radon measure $\mu \in \mathcal{M}_+(\mathbb{R}^d\backslash\{0\})$, called the *intensity measure*, and a common normalization sequence $\{b_n\}$ with $b_n \to \infty$ such that

$$n \Pr\left\{\frac{\boldsymbol{X}}{b_n} \in \cdot\right\} \xrightarrow{v} \mu(\cdot), \quad n \to \infty. \tag{3.2.2}$$

3. There exists a Radon (intensity) measure μ on $\overline{\mathbb{R}}^d\backslash\{\mathbf{0}\}$ such that

$$\lim_{t\to\infty} \frac{1 - F(t\boldsymbol{x})}{1 - F(t\mathbf{1})} = \lim_{t\to\infty} \frac{\Pr\{\boldsymbol{X}/t \in [\mathbf{0}, \boldsymbol{x}]^c\}}{\Pr\{\boldsymbol{X}/t \in [\mathbf{0}, \mathbf{1}]^c\}} = \mu([\mathbf{0}, \boldsymbol{x}]^c), \tag{3.2.3}$$

for all continuous points $\boldsymbol{x}$ of μ, where $\mu([\mathbf{0},\boldsymbol{x}]^c) = c\int_{\mathbb{S}_+^{d-1}} \max_{1\le j\le d}(u_j/x_j)^\alpha\, S(d\boldsymbol{u})$ for some constant c and a probability measure S on $\mathbb{S}_+^{d-1} := \{\boldsymbol{x}\in\mathbb{R}_+^d : ||\boldsymbol{x}|| = 1\}$.

4. There exists a Radon (intensity) measure μ on $\overline{\mathbb{R}}^d\backslash\{\mathbf{0}\}$ such that for every Borel set $B\subset\overline{\mathbb{R}}^d\backslash\{\mathbf{0}\}$ bounded away from the origin that satisfies $\mu(\partial B) = 0$, we have

$$\lim_{t\to\infty}\frac{\Pr\{\boldsymbol{X}\in tB\}}{\Pr\{||\boldsymbol{X}|| > t\}} = \mu(B). \tag{3.2.4}$$

Remark 3.2.2 The intensity measure μ in (3.2.4) depends on the choice of norm $||\cdot||$, but the intensity measures for any two norms are proportional, thus equivalent. If $\boldsymbol{X}$ is MRV with $\mu([\mathbf{0},\mathbf{1}]^c) > 0$, where $[\mathbf{0},\mathbf{1}]^c$ denotes the complement of $[\mathbf{0},\mathbf{1}]$ in $\overline{\mathbb{R}}^d$, then for every relatively compact set $B\subset\overline{\mathbb{R}}^d\backslash\{\mathbf{0}\}$ with $\mu(\partial B) = 0$, we have

$$\lim_{t\to\infty}\frac{\Pr\{\boldsymbol{X}\in tB\}}{\Pr\{\boldsymbol{X}\in t[\mathbf{0},\mathbf{1}]^c\}} = \tilde{\mu}(B), \tag{3.2.5}$$

where $\tilde{\mu}(B) = \mu(B)/\mu([\mathbf{0},\mathbf{1}]^c)$.

For any non-negative MRV random vector $\boldsymbol{X}$, its non-degenerate univariate margins X_i have regularly varying right tails. All the extremal dependence information of an MRV vector $\boldsymbol{X}$ is encoded in the intensity measure μ, which can be further decomposed into the rank-invariant tail dependence and marginal heavy-tail index using the copula approach. The next theorem gives a property of random vectors with MRV tails when multiplied by a thin-tailed scalar random variable. We will need this result in Section 4.1.

Theorem 3.2.3 Suppose that $\boldsymbol{X}$ is a non-negative random vector satisfying (3.2.2) with exponent of variation $-\alpha$. Let $Y\ge 0$ be a random variable with a moment greater than α, that is, $\exists\, \varepsilon > 0$ such that the expected value $E(Y^{\alpha(1+2\varepsilon)}) < \infty$. Then

$$n \Pr\left\{\frac{Y\boldsymbol{X}}{b_n} \in \cdot\right\} \xrightarrow{v} E(Y^\alpha)\mu.$$

In particular, if $d = 1$, we have

$$\lim_{n\to\infty} \frac{\Pr\{YZ > x\}}{\Pr\{Z > x\}} = E(Y^\alpha).$$

The following example shows that for an elliptical random vector $\boldsymbol{X}$ which has the stochastic representation $\boldsymbol{X} \stackrel{d}{=} \boldsymbol{\mu} + RA\boldsymbol{U}$ given in (2.2.25), $\boldsymbol{X}$ is MRV if and only if R has a regularly varying right tail. More detail can be found in [17, 33].

Example 3.2.4 Assume that $\boldsymbol{X} \stackrel{d}{=} \boldsymbol{\mu} + RA\boldsymbol{U} \sim \mathcal{E}_d(\boldsymbol{\mu}, \Sigma, \psi)$ and each element of Σ is positive. If R has a regularly varying right tail with heavy-tail index $\alpha > 0$, then the upper and lower tail dependence parameter between X_i and X_j is given by

$$\lambda_L(X_i, X_j) = \lambda_U(X_i, X_j) = \frac{\int_{\pi/2-\arcsin\rho_{ij}}^{\pi/2} \cos^\alpha(t)\, dt}{\int_0^{\pi/2} \cos^\alpha(t)\, dt}, \tag{3.2.6}$$

where ρ_{ij} is the (i,j)th element of the correlation matrix $[\Delta(\Sigma)]^{-1}\Sigma[\Delta(\Sigma)]^{-1}$ and $\Delta(\Sigma) := D[\sqrt{\sigma_{11}}, \ldots, \sqrt{\sigma_{dd}}]$.

3.3 Tail Dependence of MRV Distributions

In this section we derive the main result of this dissertation from Theorem 3.2.1. Using the intensity measure we are able to calculate tail dependence functions without taking the marginal transforms required for the copula method.

Theorem 3.3.1 Consider a non-negative random vector $\boldsymbol{X} = (X_1, \ldots, X_d)$ with MRV d.f. F and continuous margins $F_1, \ldots, F_d$. Let C^F and μ denote, respectively, the copula and the intensity measure of F. If the margins are tail equivalent (i.e. $\bar{F}_i(x)/\bar{F}_j(x) \to 1$ as $x \to \infty$ for any $i \neq j$), then the upper tail dependence function $b^*(\,\cdot\,; C^F)$ exists and

1. $b^*(\boldsymbol{w}; C^F) = \frac{\mu(\prod_{i=1}^d [w_i^{-1/\alpha}, \infty])}{\mu([1,\infty]\times\overline{\mathbb{R}}_+^{d-1})}$ and $a^*(\boldsymbol{w}; C^F) = \frac{\mu((\prod_{i=1}^d [0, w_i^{-1/\alpha}])^c)}{\mu(([0,1]\times\overline{\mathbb{R}}_+^{d-1})^c)}$;
2. $\frac{\mu([\boldsymbol{w},\infty])}{\mu([\mathbf{0},\mathbf{1}]^c)} = \frac{b^*((w_1^{-\alpha},\ldots,w_d^{-\alpha}); C^F)}{a^*((1,\ldots,1); C^F)}$, and $\frac{\mu([\mathbf{0},\boldsymbol{w}]^c)}{\mu([\mathbf{0},\mathbf{1}]^c)} = \frac{a^*((w_1^{-\alpha},\ldots,w_d^{-\alpha}); C^F)}{a^*((1,\ldots,1); C^F)}$.

Proof. Since each F_i, $i \in I$, is regularly varying, from (3.1.2) and (3.1.1) we have that $\bar{F}_i(x) = L_i(x)/x^\alpha$ for $x \geq 0$. To estimate $\bar{F}_i^{-1}(w_i u)$ when $u \to 0^+$ for fixed $w_i > 0$, consider

$$\bar{F}_i(w_i^{1/\alpha} x) = \frac{L_i(w_i^{1/\alpha} x)}{w_i x^\alpha} = w_i^{-1} \bar{F}_i(x) g_i(w_i, x),$$

where $g_i(w_i, x) := L_i(w_i^{1/\alpha} x)/L_i(x) \to 1$ as $x \to \infty$. Substitute $\bar{F}_i(x) = w_i u$ into above expression and take $\bar{F}_i^{-1}$ on both sides, we obtain that

$$\bar{F}_i^{-1}(w_i u) = w_i^{-1/\alpha} \bar{F}_i^{-1}(u g_i(w_i, \bar{F}_i^{-1}(w_i u))).$$

Asymptotically, $\bar{F}_i^{-1}(w_i u) \approx w_i^{-1/\alpha} \bar{F}_i^{-1}(u)$ as $u \to 0^+$.

For any fixed $\boldsymbol{w} = (w_1, \ldots, w_d)$ with $w_i > 0$, $i \in I$, consider

$$\begin{aligned} b^*(\boldsymbol{w}; C^F) &= \lim_{u \to 0^+} \frac{\Pr\{F_i(X_i) > 1 - w_i u, \forall i \in I\}}{\Pr\{F_1(X_1) > 1 - u\}} = \lim_{u \to 0^+} \frac{\Pr\{X_i > \bar{F}_i^{-1}(w_i u), \forall i \in I\}}{\Pr\{X_1 > \bar{F}_1^{-1}(u)\}} \\ &= \lim_{u \to 0^+} \frac{\Pr\{X_i > w_i^{-1/\alpha} \bar{F}_i^{-1}(u g_i(w_i, \bar{F}_i^{-1}(w_i u))), \forall i \in I\}}{\Pr\{X_1 > \bar{F}_1^{-1}(u)\}}. \qquad (3.3.1) \end{aligned}$$

Since $\bar{F}_i(x)/\bar{F}_1(x) \to 1$ as $x \to \infty$, we have from Proposition 3.1.8 that $\bar{F}_i^{-1}(u)/\bar{F}_1^{-1}(u) \to 1$ as $u \to 0^+$. For any small $\epsilon > 0$, when u is sufficiently small,

$$1 - \epsilon < g_i(w_i, \bar{F}_i^{-1}(w_i u)) < 1 + \epsilon, \text{ for all } i \in I.$$

Thus, when u is sufficiently small,

$$\frac{\bar{F}_i^{-1}(u(1-\epsilon))}{\bar{F}_1^{-1}(u)} \geq \frac{\bar{F}_i^{-1}(u g_i(w_i, \bar{F}_i^{-1}(w_i u)))}{\bar{F}_1^{-1}(u)} \geq \frac{\bar{F}_i^{-1}(u(1+\epsilon))}{\bar{F}_1^{-1}(u)}. \qquad (3.3.2)$$

Since $\bar{F_1}^{-1}(u)$ is regularly varying at 0, or more precisely, $\bar{F_1}^{-1}(uc)/\bar{F_1}^{-1}(u) \to c^{-1/\alpha}$ as $u \to 0^+$ for any $c > 0$, then by taking the limits in (3.3.2) we have

$$(1-\epsilon)^{-1/\alpha} \geq \frac{\bar{F}_i^{-1}(ug_i(w_i, \bar{F}_i^{-1}(w_i u)))}{\bar{F_1}^{-1}(u)} \geq (1+\epsilon)^{-1/\alpha}, \text{ as } u \to 0^+,$$

for any small $\epsilon > 0$. That is, when u is sufficiently small,

$$(1-\epsilon)^{-1/\alpha}\bar{F_1}^{-1}(u) \geq \bar{F}_i^{-1}(ug_i(w_i, \bar{F}_i^{-1}(w_i u))) \geq (1+\epsilon)^{-1/\alpha}\bar{F_1}^{-1}(u), \ \forall \epsilon > 0, i \in I.$$

Combining these inequalities with (3.3.1), then for any $\epsilon > 0$ we have

$$\begin{aligned} & \lim_{u\to 0^+} \frac{\Pr\{X_i > w_i^{-1/\alpha}(1-\epsilon)^{-1/\alpha}\bar{F_1}^{-1}(u), \forall i \in I\}}{\Pr\{X_1 > \bar{F_1}^{-1}(u)\}} \\ \leq \ & b^*(\boldsymbol{w}; C^F) \leq \lim_{u\to 0^+} \frac{\Pr\{X_i > w_i^{-1/\alpha}(1+\epsilon)^{-1/\alpha}\bar{F_1}^{-1}(u), \forall i \in I\}}{\Pr\{X_1 > \bar{F_1}^{-1}(u)\}}, \end{aligned}$$

which implies, after substituting $t = \bar{F_1}^{-1}(u)$,

$$b^*(\boldsymbol{w}; C^F) = \lim_{t\to\infty} \frac{\Pr\{X_i > w_i^{-1/\alpha} t, i \in I\}}{\Pr\{X_1 > t\}}. \tag{3.3.3}$$

Set $A = \prod_{i=1}^d [w_i^{-1/\alpha}, \infty]$ and $B = [1,\infty] \times \overline{\mathbb{R}}_+^{d-1}$, then (3.2.4) implies

$$b^*(\boldsymbol{w}; C^F) = \frac{\mu(A)}{\mu(B)} = \frac{\mu(\prod_{i=1}^d [w_i^{-1/\alpha}, \infty])}{\mu([1,\infty] \times \overline{\mathbb{R}}_+^{d-1})}.$$

Similar to (3.3.1), it follows from (2.2.16) that

$$\begin{aligned} a^*(\boldsymbol{w}; C^F) &= \lim_{u\to 0^+} \frac{\Pr\{F_i(X_i) > 1 - w_i u, \exists i \in I\}}{\Pr\{F_1(X_1) > 1-u\}} \\ &= \lim_{u\to 0^+} \frac{\Pr\{X_i > w_i^{-1/\alpha}\bar{F}_i^{-1}(ug_i(w_i, \bar{F}_i^{-1}(w_i u))), \exists i \in I\}}{\Pr\{X_1 > \bar{F_1}^{-1}(u)\}}. \end{aligned}$$

Borrowing from the derivation of (3.3.3) then using (3.2.4), we have

$$a^*(\boldsymbol{w}; C^F) = \lim_{t\to\infty} \frac{\Pr\{X_i > w_i^{-1/\alpha} t, \exists i \in I\}}{\Pr\{X_1 > t\}} = \frac{\mu((\prod_{i=1}^d [0, w_i^{-1/\alpha}])^c)}{\mu([1,\infty] \times \overline{\mathbb{R}}_+^{d-1})}.$$

The expressions in part 2 can be easily obtained from part 1. □

Observe that the rescaled intensity measure $\tilde{\mu}(B) = \mu(B)/\mu([\mathbf{0}, \mathbf{1}]^c)$ in (3.2.5) satisfies that $\tilde{\mu}(([\mathbf{0}, \mathbf{1}])^c) = 1$. The intensity measures μ and $\tilde{\mu}$ are uniquely determined by each other. In addition, Theorem 3.2.1 shows that the upper tail dependence function and the intensity measure for multivariate regular variation are also uniquely determined by each other. Therefore, the tail dependence function and intensity measure are equivalent in the sense that the Radon measure generated by the tail dependence function is a rescaled version of the intensity measure with marginal scaling functions being of Pareto type. In contrast, the tail dependence function describes the rank-invariant extremal dependence extracted from the intensity measure.

In general, $\bar{F}_i(x)/\bar{F}_j(x) \to r_{ij}$ as $x \to \infty$ for any $i \neq j$. If $0 < r_{ij} < \infty$, then Theorem 3.3.1 still holds by properly adjusting the marginal scaling constants; for example, substitute X_i by $c_{1i}^{1/\alpha} X_i$ for $i = 2, \ldots, d$. If $r_{ij} = 0$ or $r_{ij} = \infty$, then some margins have heavier tails than others and more subtle separate marginal scalings are needed to derive the limiting results.

Chapter 4

Applications

4.1 Heavy-Tailed Scale Mixtures of Multivariate Distributions

One broad family of multivariate distributions has the form:

$$(X_1, \dots, X_d) := (RT_1, \dots, RT_d), \tag{4.1.1}$$

where $(T_1, \dots, T_d)$ has the joint distribution $G(t_1, \dots, t_d)$ with some finite moments, and the scale variable R, independent of $(T_1, \dots, T_d)$, has a regularly varying right tail at infinity with survival function of form (3.1.2). We call such distributions "heavy-tailed scale mixtures". The class of distributions of (4.1.1) has a variety of interpretations in different applications, including, for example, multivariate elliptical distributions and various multivariate Pareto distributions as special cases.

As we showed in Theorem 3.3.1, the upper tail dependence function represents a rescaled version of the intensity measure. The link presented in Theorem 3.3.1 allows us on one hand to develop (via tail dependence functions) tractable parametric models for the intensity measure. On the other hand we can derive (via multivariate regular variation) tail dependence functions in the situations where copulas have neither explicit analytic expressions nor closure properties for their conditional distributions. We illustrate the latter part in this section for the random vectors of the form (4.1.1). We also give a monotonicity property of the tail dependence parameters with respect to the heavy-tail indices.

4.1.1 Tail Dependence

Let $T_{i+} := T_i \vee 0,\ i \in I$, and assume in this section that $0 < E(T_{i+}^{\alpha+\epsilon}) < \infty$ for some $\epsilon > 0$. Since R is regularly varying with survival function (3.1.2), then by Theorem 3.2.3, RT_{i+}, $i \in I$, is also regularly varying with survival function $\Pr\{RT_{i+} > r\} = E(T_{i+}^{\alpha})L_i(r)/r^{\alpha}$, for some L such that $L_i(r)/L_j(r) \to 1$ as $r \to \infty$ for $i \neq j$. Consider

$$(Y_1, \ldots, Y_d)^T := D[(E(T_{1+}^{\alpha}))^{-1/\alpha} T_{1+}, \ldots, (E(T_{d+}^{\alpha}))^{-1/\alpha} T_{d+}](R, \ldots, R)^T. \qquad (4.1.2)$$

Since Y_i is strictly increasing in $X_i \geq 0$, $i \in I$, $(Y_1, \ldots, Y_d)$ in (4.1.2) and $(X_1, \ldots, X_d)$ in (4.1.1) have the same upper tail dependence function. It can be shown that $(Y_1, \ldots, Y_d)$ is regularly varying, and $\Pr\{Y_i > r\}/\Pr\{Y_j > r\} \to 1$ as $r \to \infty$. Thus, by Theorem 3.3.1, the expression of the upper tail dependence function of $(X_1, \ldots, X_d)$ reduces to the determination of the intensity measure of $(Y_1, \ldots, Y_d)$ in (4.1.2).

Proposition A.1 in [5] presents a general formula for the intensity measure of a random affine transform of a random vector that is regularly varying. Applying this formula to (4.1.2), the intensity measure μ of (4.1.2) is given by, for any Borel measurable subset $B \subseteq \overline{\mathbb{R}}_+^d$,

$$\mu(B) = E(\nu(D[(E(T_{1+}^{\alpha}))^{1/\alpha} T_{1+}^{-1}, \ldots, (E(T_{d+}^{\alpha}))^{1/\alpha} T_{d+}^{-1}](B))) \qquad (4.1.3)$$

where ν is the intensity measure of $(R, \ldots, R)$. Using (3.2.3), we have for any non-negative $\boldsymbol{w} = (w_1, \ldots, w_d)$,

$$\nu([\mathbf{0}, \boldsymbol{w}]^c) = \lim_{t \to \infty} \frac{\Pr\{R > t \wedge_{i=1}^d w_i\}}{\Pr\{R > t\}} = \vee_{i=1}^d w_i^{-\alpha}.$$

Using the inclusion-exclusion relation and the fact that $\sum_{\emptyset \neq S \subseteq I}(-1)^{|S|-1} \vee_{i \in S} w_i = \wedge_{i \in I} w_i$ for all non-negative $w_1, \ldots, w_d$, we also have

$$\nu([\boldsymbol{w}, \boldsymbol{\infty}]) = \wedge_{i=1}^{d} w_i^{-\alpha}.$$

Substitute these two expressions into (4.1.3), then

$$\mu([\mathbf{0}, \boldsymbol{w}]^c) = E\left(\bigvee_{i=1}^{d} \frac{w_i^{-\alpha} T_{i+}^{\alpha}}{E(T_{i+}^{\alpha})}\right), \text{ and } \mu([\boldsymbol{w}, \boldsymbol{\infty}]) = E\left(\bigwedge_{i=1}^{d} \frac{w_i^{-\alpha} T_{i+}^{\alpha}}{E(T_{i+}^{\alpha})}\right),$$

which lead to the expression of the upper tail dependence function.

In summary, we give the following theorem.

Theorem 4.1.1 Consider a random vector $\boldsymbol{X} = (RT_1, \ldots, RT_d)$ in the form of (4.1.1) with d.f. F and continuous margins $F_1, \ldots, F_d$. Assume that $0 < E(T_{i+}^{\alpha+\epsilon}) < \infty$, $i \in I$, for some $\epsilon > 0$, then the upper tail dependence and exponent functions are given by

$$b^*(\boldsymbol{w}; C^F) = E\left(\bigwedge_{i=1}^{d} \frac{w_i T_{i+}^{\alpha}}{E(T_{i+}^{\alpha})}\right) \text{ and } a^*(\boldsymbol{w}; C^F) = E\left(\bigvee_{i=1}^{d} \frac{w_i T_{i+}^{\alpha}}{E(T_{i+}^{\alpha})}\right).$$

Applying Theorem 4.1.1 to (2.2.15) and (2.2.17), we obtain the following corollary.

Corollary 4.1.2 Making the same assumptions as Theorem 4.1.1, the upper tail and extremal dependence parameters are given by

$$\tau_J = \frac{E\left(\bigwedge_{i=1}^{d} (E(T_{i+}^{\alpha}))^{-1} T_{i+}^{\alpha}\right)}{E\left(\bigwedge_{i \in J} (E(T_{i+}^{\alpha}))^{-1} T_{i+}^{\alpha}\right)} \text{ for } \emptyset \neq J \subseteq I \text{ and } \gamma = \frac{E\left(\bigwedge_{i=1}^{d} (E(T_{i+}^{\alpha}))^{-1} T_{i+}^{\alpha}\right)}{E\left(\bigvee_{i=1}^{d} (E(T_{i+}^{\alpha}))^{-1} T_{i+}^{\alpha}\right)}.$$

We illustrate our main results using multivariate elliptical distributions. Let Σ be a $d \times d$ positive semi-definite matrix, and $\boldsymbol{U} = (U_1, \ldots, U_m)$ be uniformly distributed on the unit sphere in $\mathbb{R}^m$. Consider the following stochastic representation,

$$(X_1, \ldots, X_d)^T = (\mu_1, \ldots, \mu_d)^T + RA(U_1, \ldots, U_m)^T,$$

where A is an $d \times m$ matrix with $AA^T = \Sigma$ and R is a non-negative random variable independent of $\boldsymbol{U}$. By Theorem 2.2.14, $(X_1, \dots, X_d)$ has an elliptical distribution.

Let $(T_1, \dots, T_d)^T = A(U_1, \dots, U_m)^T$, then we have

$$(X_1, \dots, X_d) = (\mu_1, \dots, \mu_d) + R(T_1, \dots, T_d),$$

which is a scale mixture of $(T_1, \dots, T_d)$.

In the bivariate case, let

$$\Sigma = \begin{bmatrix} \sigma_{11}^2 & \sigma_{12} \\ \sigma_{12} & \sigma_{22}^2 \end{bmatrix} \text{ and } A = \begin{bmatrix} \sigma_{11} & 0 \\ \frac{\sigma_{12}}{\sigma_{11}} & \sigma_{22}\sqrt{1-\rho^2} \end{bmatrix},$$

where $\rho = \frac{\sigma_{12}}{\sigma_{11}\sigma_{22}}$. Thus,

$$T_1 = \sigma_{11} U_1, \; T_2 = \frac{\sigma_{12}}{\sigma_{11}} U_1 + \sigma_{22}\sqrt{1-\rho^2}U_2.$$

Observe that, marginally, RT_1 and RT_2 have one-dimensional elliptical distributions, and $\sigma_{11}^{-1} RT_1$ and $\sigma_{22}^{-1} RT_2$ have the same distribution ([11] Chapter 2). Assume that R has a regularly varying right tail with heavy-tail index $\alpha > 0$. Since (X_1, X_2) and (RT_1, RT_2) have the same tail dependence parameter, we have

$$\tau_1 = \lim_{t \to \infty} \Pr\{\sigma_{22}^{-1} RT_2 > t \mid \sigma_{11}^{-1} RT_1 > t\}.$$

Obviously, (T_1, T_2) has a bounded support, thus it follows from Corollary 4.1.2 that

$$\tau_1 = \frac{E(\sigma_{11}^{-1} T_{1+} \wedge \sigma_{22}^{-1} T_{2+})^\alpha}{E(\sigma_{11}^{-1} T_{1+})^\alpha} = \frac{E(U_{1+} \wedge (\rho U_1 + \sqrt{1-\rho^2}U_2)_+)^\alpha}{E(U_{1+})^\alpha}, \tag{4.1.4}$$

where $U_{1+} := U_1 \vee 0$ and $(\rho U_1 + \sqrt{1-\rho^2}U_2)_+ := (\rho U_1 + \sqrt{1-\rho^2}U_2) \vee 0$.

If $\rho = 1$, then trivially $\tau_1 = 1$. Suppose that $\rho < 1$. Transferring to polar coordinates, we have

$$U_1 = \cos\Theta, \; U_2 = \sin\Theta, \text{ and } \rho U_1 + \sqrt{1-\rho^2}U_2 = \sin(\Theta + \theta_0),$$

where Θ is uniformly distributed on $[0, 2\pi]$ and $\theta_0 = \tan^{-1} \frac{\rho}{\sqrt{1-\rho^2}}$. Clearly,

$$E(U_{1+}^{\alpha}) = \frac{1}{2\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^{\alpha} \theta d\theta = \frac{1}{\pi} \int_{0}^{\frac{\pi}{2}} \cos^{\alpha} \theta d\theta = \frac{1}{\pi} \int_{0}^{1} \frac{u^{\alpha}}{\sqrt{1-u^2}} du.$$

To calculate the quantity in the numerator of (4.1.4), consider the boundary case when

$$u_1 = \rho u_1 + \sqrt{1-\rho^2} u_2,$$

which leads to the solution $\theta_1 = \tan^{-1} \frac{u_2}{u_1} = \tan^{-1} \frac{1-\rho}{\sqrt{1-\rho^2}}$. That is,

$$\cos\theta \geq \sin(\theta + \theta_0) \geq 0 \quad \text{if } -\theta_0 \leq \theta \leq \theta_1$$
$$0 \leq \cos\theta \leq \sin(\theta + \theta_0) \quad \text{if } \theta_1 \leq \theta \leq \tfrac{\pi}{2}.$$

Simple trigonometric arguments show that $\theta_0 + 2\theta_1 = \frac{\pi}{2}$. Thus,

$$\begin{aligned} E(U_{1+} \wedge (\rho U_1 + \sqrt{1-\rho^2} U_2)_+)^{\alpha} &= \frac{1}{2\pi} \int_{-\theta_0}^{\theta_1} \sin^{\alpha}(\theta + \theta_0) d\theta + \frac{1}{2\pi} \int_{\theta_1}^{\frac{\pi}{2}} \cos^{\alpha} \theta d\theta \\ &= \frac{1}{2\pi} \int_{0}^{\frac{\pi}{2}-\theta_1} \sin^{\alpha} \theta d\theta + \frac{1}{2\pi} \int_{\theta_1}^{\frac{\pi}{2}} \cos^{\alpha} \theta d\theta. \end{aligned}$$

Since $\cos\theta_1 = \frac{\sqrt{1-\rho^2}}{\sqrt{2-2\rho}} = \left(\frac{1+\rho}{2}\right)^{\frac{1}{2}}$, we obtain that

$$E(U_{1+} \wedge (\rho U_1 + \sqrt{1-\rho^2} U_2)_+)^{\alpha} = \frac{1}{\pi} \int_{0}^{\left(\frac{1+\rho}{2}\right)^{\frac{1}{2}}} \frac{u^{\alpha}}{\sqrt{1-u^2}} du.$$

Hence,

$$\tau_1 = \frac{\int_0^{\left(\frac{1+\rho}{2}\right)^{\frac{1}{2}}} \frac{u^{\alpha}}{\sqrt{1-u^2}} du}{\int_0^1 \frac{u^{\alpha}}{\sqrt{1-u^2}} du},$$

which is identical to the parameter obtained in [17, 33].

4.1.2 Monotonicity

An intriguing issue is whether or not these tail dependence parameters are monotone in response to changes in the heavy-tail index α. It was demonstrated by the numerical results in [33] that the tail dependence parameter τ_1 is decreasing in α. We show that this is indeed the case using the ratio-of-moments expressions obtained in Corollary 4.1.2.

A non-negative random variable X is said to be smaller than a non-negative random variable Y in the hazard rate ordering, denoted by $X \leq_{hr} Y$, if the hazard rate of X is larger than that of Y. A detailed discussion on the hazard rate ordering can be found, for example, in Chapter 1 of [34], from which, the following result also holds.

Lemma 4.1.3 Let X and Y be non-negative random variables. If $X \leq_{hr} Y$, then $\frac{EX^\alpha}{EY^\alpha}$ is decreasing in α.

Proof. Theorem 1.B.12 in [34] states that if $X \leq_{hr} Y$, then

$$Eg_2(X)Eg_1(Y) \leq Eg_1(X)Eg_2(Y), \tag{4.1.5}$$

for all non-negative real functions g_1 and g_2 satisfying that $g_1(\cdot)$ is increasing and $\frac{g_2(\cdot)}{g_1(\cdot)}$ is increasing. Consider the situation that $\alpha_1 \leq \alpha_2$, $g_1(x) = x^{\alpha_1}$ and $g_2(x) = x^{\alpha_2}$ for $x \geq 0$. Then (4.1.5) reduces to

$$EX^{\alpha_2}EY^{\alpha_1} \leq EX^{\alpha_1}EY^{\alpha_2}.$$

The monotonicity of $\frac{EX^\alpha}{EY^\alpha}$ follows.

Proposition 4.1.4 Let RT_i be the ith component of (4.1.1), where T_i and R satisfy the regularity conditions specified in Theorem 4.1.2. Suppose that $E(T_{i+}^\alpha) = E(T_{j+}^\alpha)$, $i \neq j$.

1. If $\wedge_{i=1}^d T_{i+} \leq_{hr} \wedge_{i\in J} T_{i+}$, then τ_J is decreasing in α.

2. If $\wedge_{i=1}^d T_{i+} \leq_{hr} \vee_{i=1}^d T_{i+}$, then γ is decreasing in α.

Proof. If $E(T_{i+}^{\alpha}) = E(T_{j+}^{\alpha})$, $i \neq j$, then from Corollary 4.1.2,

$$\tau_J = \frac{E(\wedge_{i=1}^{d} T_{i+})^{\alpha}}{E(\wedge_{i\in J} T_{i+})^{\alpha}}, \text{ for all } \emptyset \neq J \subset I, \text{ and } \gamma = \frac{E(\wedge_{i=1}^{d} T_{i+})^{\alpha}}{E(\vee_{i=1}^{d} T_{i+})^{\alpha}}.$$

The monotone properties follow from Lemma 4.1.3.

Notice that the inequality (4.1.5) resembles the property of total positivity of order 2, and in fact, Proposition 4.1.4 can be established directly by using the theory of total positivity.

To show that τ_1 of a bivariate elliptical distribution is decreasing in α, we need, by virtue of Lemma 4.1.3, to establish that $U_{1+} \geq_{hr} U_{1+} \wedge (\rho U_1 + \sqrt{1-\rho^2}U_2)_+$. From [34] (condition 1.B.3 in page 16), it is sufficient to show that

$$\frac{\Pr\{U_{1+} > t\}}{\Pr\{U_{1+} \wedge (\rho U_1 + \sqrt{1-\rho^2}U_2)_+ > t\}} \text{ is increasing in } t \in [0, s], \tag{4.1.6}$$

where $s \leq 1$ is the right endpoint of the support of $U_{1+} \wedge (\rho U_1 + \sqrt{1-\rho^2}U_2)_+$. Again, using the polar coordinate system, we have for any $t \in [0, s]$,

$$\begin{aligned} \Pr\{U_{1+} > t\} &= \frac{\cos^{-1} t}{\pi}, \\ \Pr\{U_{1+} \wedge (\rho U_1 + \sqrt{1-\rho^2}U_2)_+ > t\} &= \frac{(\cos^{-1} t - \sin^{-1} t)_+}{2\pi}. \end{aligned}$$

It is easy to verify that $\frac{1}{2}\frac{\cos^{-1} t}{\cos^{-1} t - \sin^{-1} t}$ is increasing in $t \in [0, s]$, and then (4.1.6) follows. Therefore, τ_1 is decreasing in α.

4.2 Tail Approximation of Value-at-Risk

In this section a general tail approximation method is given for the Value-at-Risk of any norm of any random vector with a multivariate regularly varying distribution. The main result is derived using the equivalence of the intensity measure of multivariate regular variation and the tail dependence function of the underlying copula (Theorem 3.3.1). In particular, we extend the tail approximation discussed in [10] for Archimedean copulas. The explicit tail approximations for random vectors with Archimedean copulas and Pareto distributions of Marshall-Olkin type are also presented to illustrate the results. The results obtained in the section are also included in [36].

Value-at-Risk (VaR) is one of the most widely used risk measures in financial risk management [28]. Given a non-negative random variable X representing loss, the VaR at confidence level p, $0 < p < 1$, is defined as the p-th quantile of the loss distribution:

$$\mathrm{VaR}_p(X) := \inf\{t \in \mathbb{R} : \Pr\{X > t\} \leq 1 - p\}.$$

In risk analysis for multivariate portfolios, we are more interested in calculating the VaR for aggregated data than for a single loss variable. For a non-negative random loss vector $\boldsymbol{X} = (X_1, \ldots, X_d)$ representing various losses in a multivariate portfolio, we need to calculate $\mathrm{VaR}_p(||\boldsymbol{X}||)$ where the norm $|| \cdot ||$ determines the manner in which the data are aggregated. Calculating $\mathrm{VaR}_p(||\boldsymbol{X}||)$ is in general a difficult problem, but the tail estimates for $\mathrm{VaR}_p(||\boldsymbol{X}||)$ as $p \to 1$ are often tractable for various multivariate loss distributions. Good tail approximations of $\mathrm{VaR}_p(||\boldsymbol{X}||)$ as $p \to 1$ can be used to accurately estimate the risk, as measured by the VaR, for extreme losses.

Tail asymptotics for $\mathrm{VaR}_p(\sum_{i=1}^d X_i)$, as $p \to 1$, for loss vectors with Archimedean copulas and regularly varying margins are obtained in [1, 2, 3, 10, 22]. These tail estimates for the VaR of sums are, asymptotically, linear functions of the VaR of the univariate margin, with a proportionality constant that depends on the tail dependence of the underlying Archimedean copula and the marginal heavy-tail index. These asymptotic relations can also be used to analyze the structural properties of the VaR of aggregated data from multivariate portfolios, such as the subadditivity properties of the VaR and how the dependence and marginal parameters would affect extreme risks. In this section, we develop a general and more unified approach to derive the tail asymptotics of $\mathrm{VaR}_p(||\boldsymbol{X}||)$, as $p \to 1$, for loss random vectors that have multivariate regularly varying distributions. Our method is based on the link between multivariate regular variation and tail dependence functions of copulas. The tail estimates previously obtained in the literature can be obtained from our tail asymptotics as special cases.

Consider a non-negative MRV random vector $\boldsymbol{X} = (X_1, \dots, X_d)$ with intensity measure μ and margins $F_1, \dots, F_d$ that are tail equivalent with heavy-tail index $\beta > 0$. Without loss of generality, we use F_1 to define the following limit:

$$q_{||\cdot||}(\beta, b^*) := \lim_{t\to\infty} \frac{\Pr\{||\boldsymbol{X}|| > t\}}{\overline{F}_1(t)}, \tag{4.2.1}$$

where b^* denotes the upper tail dependence function of $\boldsymbol{X}$. This limiting constant depends on the intensity measure μ, which in turn depends on the heavy-tail index β, tail dependence function b^* and norm $||\cdot||$.

Theorem 4.2.1 If margins $F_1, \dots, F_d$, $d \geq 2$, are continuous and the partial derivative $\partial^d b^*(\boldsymbol{v})/\partial v_1 \cdots \partial v_d$ exists almost everywhere, then $q_{||\cdot||}(\beta, b^*)$ has the following representation

$$q_{||\cdot||}(\beta, b^*) = (-1)^d \int_W \frac{\partial^d b^*(\boldsymbol{w}^{-\beta})}{\partial w_1 \cdots \partial w_d}\, d\boldsymbol{w} = \int_{W^{-1/\beta}} \frac{\partial^d b^*(\boldsymbol{v})}{\partial v_1 \cdots \partial v_d}\, d\boldsymbol{v}, \tag{4.2.2}$$

where $W = \{\boldsymbol{w} \geq \boldsymbol{0} : ||\boldsymbol{w}|| > 1\}$ and $W^{-1/\beta} = \{\boldsymbol{v} \geq \boldsymbol{0} : ||\boldsymbol{v}^{-1/\beta}|| > 1\}$.

Proof. It follows from (3.2.4) that

$$\frac{\Pr\{||\boldsymbol{X}|| > t\}}{\Pr\{X_1 > t\}} = \frac{\Pr\{\boldsymbol{X} \in tW\}}{\Pr\{\boldsymbol{X} \in tW_1\}} \to \frac{\mu(W)}{\mu(W_1)}, \quad \text{as } t \to \infty,$$

where $W = \{\boldsymbol{w} : ||\boldsymbol{w}|| > 1\}$, and $W_1 = \{\boldsymbol{w} : w_1 > 1\}$. Let $\tilde{\mu}(\cdot) := \mu(\cdot)/\mu(W_1)$, then by Theorem 3.3.1 (1), $\tilde{\mu}(\cdot)$ is the measure generated by $b^*(\boldsymbol{w}^{-\beta})$. Since margins are continuous, $\tilde{\mu}(\cdot)$ is absolutely continuous with respect to the Lebesgue measure, it follows from the Radon-Nikodym Theorem that

$$q_{||\cdot||}(\beta, b^*) = \frac{\mu(W)}{\mu(W_1)} = \int_W \tilde{\mu}'(\boldsymbol{w})\, d\boldsymbol{w},$$

where $\tilde{\mu}'(\boldsymbol{w}) = (-1)^d \frac{d^d}{dw_1\cdots dw_d} \frac{\mu([\boldsymbol{w},\boldsymbol{\infty}])}{\mu([1,\infty]\times\overline{\mathbb{R}}_+^{d-1})} \geq 0$ is the Radon-Nikodym derivative of the intensity measure $\tilde{\mu}$ with respect to the Lebesgue measure. Therefore, using Theorem 3.3.1 (1) we obtain the first expression in (4.2.2). The second expression in (4.2.2) follows upon substitution. □

Remark 4.2.2

1. It follows from the non-negativity of the Radon-Nikodym derivative that

$$0 \leq (-1)^d \frac{d^d}{dw_1\cdots dw_d} \frac{\mu([\boldsymbol{w},\boldsymbol{\infty}])}{\mu([1,\infty]\times\overline{\mathbb{R}}_+^{d-1})} = \frac{\partial^d b^*(\boldsymbol{v})}{\partial v_1\cdots\partial v_d}, \text{ with } v_i = w_i^{-\beta}, 1 \leq i \leq d,$$

which implies that $\partial^d b^*(\boldsymbol{v})/\partial v_1\cdots\partial v_d \geq 0$ for all $\boldsymbol{v} \geq \boldsymbol{0}$. It is easy to see that $W^{-1/\beta} \subseteq W^{-1/\beta'}$ for any $\beta \leq \beta'$, and thus by (4.2.2), $q_{||\cdot||}(\beta, b^*)$ is non-decreasing in β. This extends Theorem 2.5 in [2] to multivariate regular variation with respect to any norm.

2. If we choose $||\cdot||$ to be the ℓ_∞-norm, i.e., $||\boldsymbol{w}|| = \max\{w_i, i = 1,\ldots,d\}$, then $W^{-1/\beta} = \{\boldsymbol{w} \geq \boldsymbol{0} : \wedge_{i=1}^d w_i < 1\}$. In this case, $q_{||\cdot||}(\beta, b^*)$ is independent of β and depends on the tail dependence function only.

The limiting proportionality constant $q_{||\cdot||}(\beta, b^*)$ takes particularly tractable forms for Archimedean copulas as shown in the following corollaries.

Corollary 4.2.3 (Archimedean Copula) Consider a random vector $(X_1, \dots, X_d)$ which satisfies the assumptions of Theorem 4.2.1. Assume that $(X_1, \dots, X_d)$ has an Archimedean copula $C(u_1, \dots, u_d) = \psi^{-1}(\sum_{i=1}^{d} \psi(u_i))$, where the generator ψ is regularly varying at 1 with heavy-tail index $\alpha > 0$. Then

$$q_{||\cdot||}(\beta, b^*) = \int_{W^{-1/\beta}} \frac{\partial^d}{\partial v_1 \cdots \partial v_d} \sum_{S \subseteq I, S \neq \emptyset} \left[(-1)^{|S|-1} \Big(\sum_{j \in S} v_j^{\alpha} \Big)^{\frac{1}{\alpha}} \right] dv_1 \cdots dv_d, \qquad (4.2.3)$$

where $W^{-1/\beta} = \{\boldsymbol{v} : ||\boldsymbol{v}^{-\frac{1}{\beta}}|| > 1\}$.

Proof. From [19], the upper tail dependence function $b^*(\boldsymbol{v})$ can be expressed by the upper exponent functions of the margins:

$$b^*(\boldsymbol{v}) = \sum_{S \subseteq I, S \neq \emptyset} (-1)^{|S|-1} a_S^*(v_i, i \in S),$$

where $a_S^*(v_i, i \in S) = \lim_{v_i \to 0, i \notin S} a^*(\boldsymbol{v})$ is the upper exponent function of the margin of C with component indices in S. It is given in [15] that the upper exponent function of Archimedean copula is $a^*(\boldsymbol{v}) = (\sum_{j=1}^{d} v_j^{\alpha})^{\frac{1}{\alpha}}$. Therefore, $a_S^*(v_i, i \in S) = (\sum_{j \in S} v_j^{\alpha})^{\frac{1}{\alpha}}$ and the result follows. □

Corollary 4.2.4 (Archimedean Survival Copula) Consider a random vector $(X_1, \dots, X_d)$ satisfying the assumptions of Theorem 4.2.1. Assume that $(-X_1, \dots, -X_d)$ has an Archimedean copula $C(u_1, \dots, u_d) = \phi^{-1}(\sum_{i=1}^{d} \phi(u_i))$, where the generator ϕ is regularly varying at 0 with heavy-tail index $\alpha > 0$. Then

$$q_{||\cdot||}(\beta, b^*) = \int_{B} \frac{\partial^d}{\partial x_1 \cdots \partial x_d} \Big(\sum_{i=1}^{d} x_i^{-\alpha\beta} \Big)^{-\frac{1}{\alpha}} dx_1 \cdots dx_d, \qquad (4.2.4)$$

where $B = \{\boldsymbol{x} : ||\boldsymbol{x}^{-1}|| > 1\}$.

Proof. It is given in [19] that the upper tail dependence function of $(-X_1, \ldots, -X_d)$ is $b^*(\boldsymbol{w}) = (\sum_{j=1}^d w_j^{-\alpha})^{-\frac{1}{\alpha}}$. Therefore,

$$q_{||\cdot||}(\beta, b^*) = (-1)^d \int_W \frac{\partial^d}{\partial w_1 \cdots \partial w_d} b^*(\boldsymbol{w}^{-\beta})\, d\boldsymbol{w} = (-1)^d \int_W \frac{\partial^d}{\partial w_1 \cdots \partial w_d} \big(\sum_{j=1}^d w_j^{\alpha\beta}\big)^{-\frac{1}{\alpha}}\, d\boldsymbol{w}$$

where $W = \{\boldsymbol{w} : ||\boldsymbol{w}|| > 1\}$. If we substitute x_i^{-1} for w_i, then we obtain (4.2.4). □

Remark 4.2.5 1. In Corollary 4.2.3, if we choose $||\cdot||$ to be the ℓ_∞-norm, then the explicit expression of the limiting constant can be easily computed for Archimedean copulas. Take the bivariate case for example,

$$q_{||\cdot||}(\beta, b^*) = \int_{W^{-1/\beta}} \frac{\partial^2}{\partial v_1 \partial v_2} [v_1 + v_2 - (v_1^\alpha + v_2^\alpha)^{\frac{1}{\alpha}}]\, dv_1 dv_2$$

where $W^{-1/\beta} = \{\boldsymbol{v} : v_1^{-\frac{1}{\beta}} \vee v_2^{-\frac{1}{\beta}} > 1\} = \{\boldsymbol{v} : 0 < v_1 \wedge v_2 < 1\}$. Hence,

$$\begin{aligned} q_{||\cdot||}(\beta, b^*) &= \left(\int_0^\infty \int_0^1 + \int_0^1 \int_0^\infty - \int_0^1 \int_0^1\right) \frac{\partial^2}{\partial v_1 \partial v_2} [v_1 + v_2 - (v_1^\alpha + v_2^\alpha)^{\frac{1}{\alpha}}]\, dv_1 dv_2 \\ &= 2[v_1 + v_2 - (v_1^\alpha + v_2^\alpha)^{\frac{1}{\alpha}}]\ |_{v_1=0}^1 |_{v_2=0}^\infty - [v_1 + v_2 - (v_1^\alpha + v_2^\alpha)^{\frac{1}{\alpha}}]\ |_{v_1=0}^1 |_{v_2=0}^1 \\ &= 2^{\frac{1}{\alpha}} \end{aligned}$$

which only depends on the tail dependence of the Archimedean copula.

2. In Corollary 4.2.4, if we choose $||\cdot||$ to be the ℓ_1-norm, i.e., $||\boldsymbol{x}|| = \sum_{i=1}^d x_i$ for non-negative x_i, $i = 1, \ldots, d$, then the result is reduced to the limiting constant obtained in Proposition 2.2 of [10]. Using the ℓ_1-norm is the most common way to aggregate data. The monotonicity property of $q_{||\cdot||}(\beta, b^*)$ with respect to α for $d \geq 2$ is also given in [10]: $q_{||\cdot||}(\beta, b^*)$ is increasing in α when $\beta > 1$; $q_{||\cdot||}(\beta, b^*)$ is decreasing in α when $\beta < 1$; $q_{||\cdot||}(\beta, b^*) = d$ when $\beta = 1$. It is further shown that for $\alpha > 0$ and $\beta > 1$ ($\beta < 1$), VaR is asymptotically subadditive (or superadditive),

i.e., $\mathrm{VaR}_p(\sum_{i=1}^d X_i) < \sum_{i=1}^d \mathrm{VaR}_p(X_i)$ (or $\mathrm{VaR}_p(\sum_{i=1}^d X_i) > \sum_{i=1}^d \mathrm{VaR}_p(X_i)$) for p near 1.

Heavy-tailed scale mixtures of multivariate exponential distributions also have explicit tail dependence functions [26]. Thus the limiting constant $q_{||\cdot||}(\beta, b^*)$ can be calculated using Theorem 4.2.1 for these distributions. Consider the following multivariate Pareto distribution of Marshall-Olkin type:

$$\boldsymbol{X} = (X_1, \ldots, X_n) = \left(\frac{T_1}{Z}, \ldots, \frac{T_n}{Z}\right), \tag{4.2.5}$$

where Z follows a gamma distribution with shape parameter (Pareto index) $\beta > 0$ and scale parameter 1. The random vector $(T_1, \ldots, T_n)$, independent of Z, has a multivariate Marshall-Olkin distribution [27]. Obtaining the limit $q_{||\cdot||}(\beta, b^*)$ for such $\boldsymbol{X}$ is straightforward using Theorem 3.1 of [26]. Here we give the bivariate case as an example.

Example 4.2.6 (Bivariate Pareto Distribution of Marshall-Olkin Type)

The Marshall-Olkin distribution with rate parameters $\{\lambda_1, \lambda_2, \lambda_{12}\}$ is the joint distribution of $T_1 := E_1 \wedge E_{12}$, $T_2 := E_2 \wedge E_{12}$, where $\{E_S, S \subseteq \{1, 2\}\}$ is a sequence of independent exponentially distributed random variables, with E_S having mean $1/\lambda_S$. In the reliability context, T_1, T_2 can be viewed as the lifetime of two components operating in a random shock environment where a fatal shock governed by the Poisson process with rate λ_S destroys all the components with indices in $S \subseteq \{1, 2\}$ simultaneously. In credit-risk modeling, T_1, T_2 can be viewed as the time-to-default of various different counterparties or types of counterparty, for which the Poisson shocks might be a variety of underlying economic events [9].

Let $R = 1/Z$ in (4.2.5), then R has an inverse gamma distribution with shape parameter $\beta > 0$ and scale parameter 1. It is known that R is regularly varying with heavy-tail index β and its survival function is given by

$$\overline{F}(r) = \Pr\{R > r\} = 1 - \frac{\Gamma(\beta, \frac{1}{r})}{\Gamma(\beta)}, \quad r > 0, \ \beta > 0,$$

where $\Gamma(\beta, \frac{1}{r}) = \int_{\frac{1}{r}}^{\infty} t^{\beta-1} e^{-t} dt$ is the upper incomplete gamma function and $\Gamma(\beta) = \int_0^{\infty} t^{\beta-1} e^{-t} dt$ is the gamma function. The classical Breiman Theorem (see, e.g., [31]) implies that the i-th margin X_i in (4.2.5) is regularly varying with heavy-tail index β.

From [26], the bivariate distribution of (RT_1, RT_2) is regularly varying, and its upper tail dependence function is

$$b^*(w_1, w_2) = E\left(\frac{w_1 T_1^{\beta}}{E(T_1^{\beta})} \bigwedge \frac{w_2 T_2^{\beta}}{E(T_2^{\beta})}\right). \tag{4.2.6}$$

Since T_i is exponentially distributed with mean $1/(\lambda_i + \lambda_{12})$ for $i = 1, 2$, we have $ET_i^{\beta} = \beta!(1/(\lambda_i + \lambda_{12}))^{\beta}$ for any positive integer β, thus

$$b^*(w_1, w_2) = \frac{1}{\beta!} E[w_1^{1/\beta}(\lambda_1 + \lambda_{12})T_1 \wedge w_2^{1/\beta}(\lambda_2 + \lambda_{12})T_2]^{\beta}.$$

Consider

$$\begin{aligned}
& \Pr\left\{[w_1^{1/\beta}(\lambda_1 + \lambda_{12})T_1 \wedge w_2^{1/\beta}(\lambda_2 + \lambda_{12})T_2]^{\beta} > t\right\} \\
= & \Pr\left\{(\lambda_1 + \lambda_{12})T_1 > (t/w_1)^{\frac{1}{\beta}}, (\lambda_2 + \lambda_{12})T_2 > (t/w_2)^{\frac{1}{\beta}}\right\} \\
= & \Pr\left\{T_1 > (\lambda_1 + \lambda_{12})^{-1}(t/w_1)^{\frac{1}{\beta}}, T_2 > (\lambda_2 + \lambda_{12})^{-1}(t/w_2)^{\frac{1}{\beta}}\right\} \\
= & e^{-\left(\frac{\lambda_1}{(\lambda_1+\lambda_{12})w_1^{1/\beta}} + \frac{\lambda_2}{(\lambda_2+\lambda_{12})w_2^{1/\beta}} + \frac{\lambda_{12}}{(\lambda_1+\lambda_{12})w_1^{1/\beta} \wedge (\lambda_2+\lambda_{12})w_2^{1/\beta}}\right)t^{\frac{1}{\beta}}}.
\end{aligned}$$

Thus, for any positive integer β,

$$b^*(w_1,w_2) = \frac{1}{\beta!}\int_0^\infty e^{-\left(\frac{\lambda_1}{(\lambda_1+\lambda_{12})w_1^{1/\beta}}+\frac{\lambda_2}{(\lambda_2+\lambda_{12})w_2^{1/\beta}}+\frac{\lambda_{12}}{(\lambda_1+\lambda_{12})w_1^{1/\beta}\wedge(\lambda_2+\lambda_{12})w_2^{1/\beta}}\right)t^{\frac{1}{\beta}}}dt$$

$$= \frac{1}{\beta!}\int_0^\infty \beta s^{\beta-1} e^{-\left(\frac{\lambda_1}{(\lambda_1+\lambda_{12})w_1^{1/\beta}}+\frac{\lambda_2}{(\lambda_2+\lambda_{12})w_2^{1/\beta}}+\frac{\lambda_{12}}{(\lambda_1+\lambda_{12})w_1^{1/\beta}\wedge(\lambda_2+\lambda_{12})w_2^{1/\beta}}\right)s}ds$$

$$= \left(\frac{\lambda_1}{(\lambda_1+\lambda_{12})w_1^{1/\beta}}+\frac{\lambda_2}{(\lambda_2+\lambda_{12})w_2^{1/\beta}}+\frac{\lambda_{12}}{(\lambda_1+\lambda_{12})w_1^{1/\beta}\wedge(\lambda_2+\lambda_{12})w_2^{1/\beta}}\right)^{-\beta}. \tag{4.2.7}$$

Observing that b^* is differentiable almost everywhere, the limiting constant follows upon substitution of b^* into (4.2.2). □

Remark 4.2.7 In Example 4.2.6, if we set $\lambda_{12} = 0$, then $T_i = E_i$, $i = 1, 2$. In this case, it is known that the survival copula of (RT_1, RT_2) is a bivariate Clayton copula. On the other hand, equation (4.2.7) becomes $b^*(w_1, w_2) = (w_1^{-1/\beta} + w_2^{-1/\beta})^{-\beta}$. This is the upper tail dependence function of $(-X_1, -X_2)$ from the bivariate case of Corollary 4.2.4, where the dependence parameter is $1/\beta$.

The next theorem, an application of Theorem 4.2.1, provides a method of obtaining VaR approximations in higher dimensions based on VaR in one dimension.

Theorem 4.2.8 Consider a non-negative random vector $\boldsymbol{X} = (X_1, \ldots, X_d)$ with MRV distribution function F and continuous margins $F_1, \ldots, F_d$, $d \geq 2$. Assume that the margins are tail equivalent with heavy-tail index $\beta > 0$, and the tail dependence function $b^* > 0$. Then

$$\lim_{p\to 1}\frac{\mathrm{VaR}_p(||\boldsymbol{X}||)}{\mathrm{VaR}_p(X_1)} = q_{||\cdot||}(\beta, b^*)^{\frac{1}{\beta}}. \tag{4.2.8}$$

Proof. Let G be the distribution function of $||\boldsymbol{X}||$. From (4.2.1), $\frac{\overline{G}(t)}{\overline{F}_1(t)} \to q_{||\cdot||}(\beta, b^*)$ as $t \to \infty$, i.e. $\overline{F}_1(t) \approx q_{||\cdot||}(\beta, b^*)^{-1}\overline{G}(t)$, where '$\approx$' denotes tail equivalence. Hence, we have $t \approx \overline{F}_1^{-1}(q_{||\cdot||}(\beta, b^*)^{-1}\overline{G}(t))$.

Define $u := \overline{G}(t)$. Then $\overline{G}^{-1}(u) \approx \overline{F}_1^{-1}(q_{||\cdot||}(\beta, b^*)^{-1}u)$. Since F_1 is regularly varying at ∞ with heavy-tail index $\beta > 0$, we have from Proposition 2.6 of [32] that $\overline{F}_1^{-1}(t)$ is regularly varying at 0, or more precisely, $\overline{F}_1^{-1}(uc)/\overline{F}_1^{-1}(u) \to c^{-\frac{1}{\beta}}$ as $u \to 0^+$ for any $c > 0$. Thus $\overline{F}_1^{-1}(q_{||\cdot||}(\beta, b^*)^{-1}u)/\overline{F}_1^{-1}(u) \to q_{||\cdot||}(\beta, b^*)^{\frac{1}{\beta}}$. Therefore, $\overline{G}^{-1}(u) \approx q_{||\cdot||}(\beta, b^*)^{\frac{1}{\beta}}\overline{F}_1^{-1}(u)$, i.e., $\lim_{u\to 0^+} \overline{G}^{-1}(u)/\overline{F}_1^{-1}(u) = q_{||\cdot||}(\beta, b^*)^{\frac{1}{\beta}}$. Replace u by $1 - p$, then we have (4.2.8). □

This result gives an asymptotic estimate of $\mathrm{VaR}_p(||\boldsymbol{X}||)$. Value at Risk does not generally have a closed form expression. The closed forms of tail dependence functions occur more often than for explicit copula expressions. This is because we can obtain tail dependence functions from either explicit copula expressions or the closure properties of the related conditional distributions whose parent distributions do not have explicit copulas. Using Theorems 4.2.1 and 4.2.8, when the confidence level is close to 1 and the tail dependence functions exist, we can take $q_{||\cdot||}(\beta, b^*)^{\frac{1}{\beta}}\mathrm{VaR}_p(X_1)$ as an approximation to the VaR of the normed (aggregated) loss. Explicit forms of VaR approximations for Archimedean copulas and Pareto distributions of Marshall-Olkin type can be derived directly from our results. As mentioned in Remark 4.2.5, the additivity properties of VaR under Archimedean copulas depend on the heavy-tail index of the margins and the copula generators. Whether similar structural properties hold for multivariate regular variation and how they hold are topics worth further investigation.

4.3 Simulations

For bivariate elliptical distributions, we proved in Section 4.1 that τ_1 is always decreasing in the heavy-tail index α. This property is illustrated in the following Example 4.3.1.

To generate an elliptical random vector $\boldsymbol{X} \sim \mathcal{E}_3(\mathbf{0}, \Sigma, \psi)$, we have the following steps:

- Generate a random variable R which follows the Pareto distribution with d.f. $F(x) = 1 - \frac{1}{(1+x)^\alpha}$, $x > 0$, $\alpha > 0$.

- Generate a 3-dimensional random vector $\boldsymbol{U}$ which follows the uniform distribution on the unit sphere. Specifically, we generate two random variables u, v which are uniformly distributed on (0,1). Then let

$$\boldsymbol{U} = (1 - 2u, 2\sqrt{u(1-u)}\cos(2\pi u), 2\sqrt{u(1-u)}\sin(2\pi u)).$$

  This method of generating uniformly distributed data points on a sphere is given in [12]. Figure 1 below shows 1000 uniformly distributed data on the unit sphere.

- Take the Cholesky decomposition of $\Sigma = AA^T$, where Σ is a given $d \times d$ matrix, A is a $d \times d$ lower triangular matrix with positive diagonal entries.

- Set $\boldsymbol{X} = R\boldsymbol{T}$, where $\boldsymbol{T} = A\boldsymbol{U}$.

Example 4.3.1 Let $\Sigma = \begin{bmatrix} 8 & 6 & 7 \\ 6 & 8 & 7 \\ 7 & 7 & 8 \end{bmatrix}$. From (2.2.7) we have

$$\tau_1 = \frac{E\left(\bigwedge_{i=1}^{3} (E(T_{i+}^\alpha))^{-1} T_{i+}^\alpha\right)}{E\left((E(T_{1+}^\alpha))^{-1} T_{1+}^\alpha\right)}.$$

It is verified in Figure 2 that τ_1 is decreasing in α.

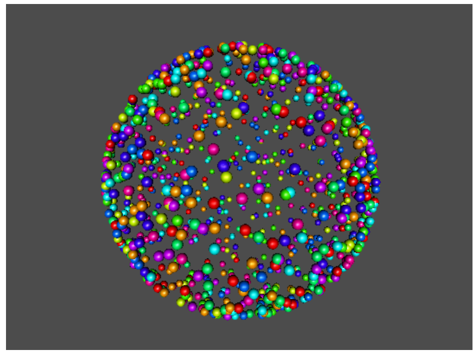

Figure 1: 1000 uniformly distributed data points generated on a sphere

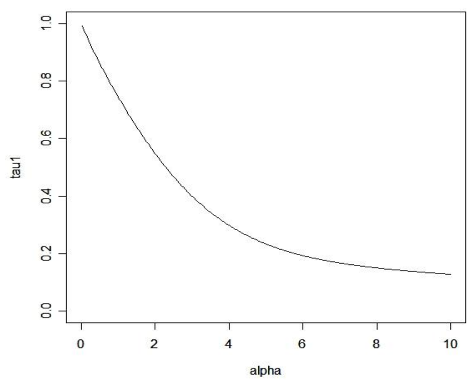

Figure 2: τ_1 decreasing in α

For higher dimensional cases, we have obtained some numerical results. Here we give two conjectures based on our simulation results.

Conjecture 4.3.2 For a 3-dimensional elliptical random vector, the upper tail dependence parameter τ_{12} increases as σ_{12} decreases (compare Σ and Σ_1 below); τ_{12} increases as σ_{13} or σ_{23} increase (compare Σ and Σ_2), where σ_{ij} is the (i, j) element of Σ. This phenomenon should be true in general for higher dimensions.

For example, let $\Sigma = \begin{bmatrix} 6 & 4 & 1 \\ 4 & 4 & 1 \\ 1 & 1 & 2 \end{bmatrix}$, $\Sigma_1 = \begin{bmatrix} 6 & 3 & 1 \\ 3 & 4 & 1 \\ 1 & 1 & 2 \end{bmatrix}$ and $\Sigma_2 = \begin{bmatrix} 6 & 4 & 2 \\ 4 & 4 & 1 \\ 2 & 1 & 2 \end{bmatrix}$. From Figure 3 we can see that $\tau_{12}(\Sigma) < \tau_{12}(\Sigma_1)$, $\tau_{12}(\Sigma) < \tau_{12}(\Sigma_2)$.

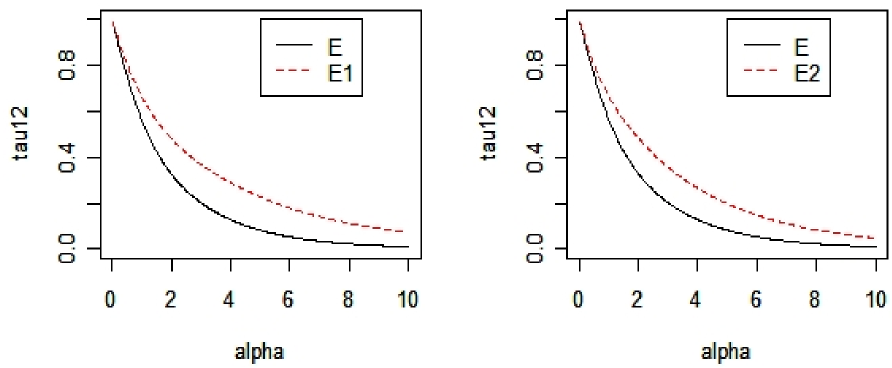

Figure 3: Comparison on τ_{12}

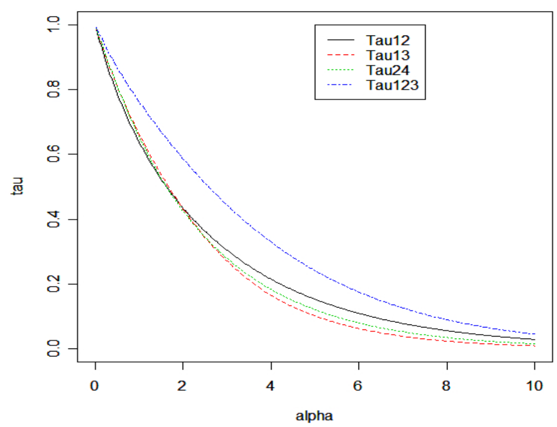

Figure 4: τ_J decreasing in α

Conjecture 4.3.3 All the tail dependence parameters τ_J are decreasing in the heavy tail index α.

For example, choose the dispersion matrix as $\Sigma = \begin{bmatrix} 8 & 6 & 6 & 5 \\ 6 & 8 & 7 & 6 \\ 6 & 7 & 8 & 3 \\ 5 & 6 & 3 & 8 \end{bmatrix}$ for a 4-dimensional elliptical distribution.

From Figure 4 we can see that the four upper tail dependence parameters (randomly chosen from all possible τ_J) are all decreasing in α.

Chapter 5

Concluding Remarks

In this dissertation, we discussed the copula method and the theory of multivariate regular variation. We studied the tail dependence of regularly varying distributions. After giving the overview we investigated the relationships among copulas and regular variation; leading to an equivalence of the two methods. Tail dependence is defined by copulas, but the copula of a random vector is difficult to obtain in high dimensions. Using the equivalence that we established, we are able to express tail dependence in terms of the intensity measure which scales well with the dimension. In addition, we showed in Chapter 4 that the general method developed in this dissertation is convenient in tail analysis and has many applications. We gave explicit tail dependence expressions for heavy-tailed scale mixtures of multivariate distributions such as elliptical distributions; we also gave a tail approximation for high dimensional Value-at-Risk. These are new contributions, among others in this dissertation, to probability and risk management. Particularly, we note that the Archimedean copulas, which are widely used in financial risk management, have tractable properties in many aspects.

We also investigated the monotonicity properties of tail dependence parameters with respect to heavy-tail indices. This is useful because there exist methods for approximating the heavy tail parameters. Any monotonicity property will help determine the tail behavior. However, this problem was not fully resolved. We gave two reasonable conjectures in the last section, based on simulation results. Moreover, the tail dependence discussed in this dissertation does not depend on the order of the variables. For example, the tail dependence of a number of stocks does not depend on how the stocks are listed. There are other types of dependence based on the nature and cause. For example, some dependence might be unidirectional [24]. There could be a causal link between the variables so that one variable taking an extreme value will cause the other variables to do so as well. Further research on these new topics is needed.

Bibliography

[1] Albrecher, H., Asmussen, S. and Kortschak, D. (2006). Tail asymptotics for the sum of two heavy-tailed dependent risks. *Extremes*, 9:107-130.

[2] Alink, S., Löwe, M. and Wüthrich, M. V. (2004). Diversification of aggregate dependent risks. *Insurance: Math. Econom.*, 35:77-95.

[3] Alink, S., Löwe, M. and Wüthrich, M. V. (2007). Diversification for general copula dependence. *Statistica Neerlandica*, 61:446-465.

[4] Basrak, B., Davis, R.A. and Mikosch, T. (2002). A characterization of multivariate regular variation. *Ann. Appl. Probab.*, 12:908-920.

[5] Basrak, B., Davis, R.A. and Mikosch, T. (2002). Regular variation of GARCH processes. *Stoch. Proc. Appl.*, 99:95-116.

[6] Bingham, N. H., Goldie, C. M. and Teugels, J. L. (1987). *Regular Variation*. Cambridge University Press, Cambridge, UK.

[7] Buhl, C., Reich, C. and Wegmann, P. (2002). Extremal dependence between return risk and liquidity risk: An analysis for the Swiss Market. Technical report, Department of Finance, University of Basel.

[8] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997). *Modelling Extremal Events for Insurance and Finance*. Springer, New York.

[9] Embrechts, P., Lindskog, F. and McNeil, A. (2003). Modeling dependence with copulas and applications to risk management.*Handbook of Heavy Tailed Distributions in Finance*, ed. S. Rachev, Elsevier, Chapter 8, pp. 329-384.

[10] Embrechts, P., Neslehová, J. and Wüthrich, M. V., (2009). Additivity properties for Value-at-Risk under Archimedean dependence and heavy-tailedness. *Insurance: Mathematics and Economics*, 44(2):164-169.

[11] Fang, K. T., Kotz, S. and Ng, K. W. (1990). *Symmetric Multivariate and Related Distributions*. Chapman & Hall, London.

[12] Fang, K. T. and Wang Y. (1994) *Number-Theoretic Methods in Statistics*. Chapman & Hall, London.

[13] Finkenstädt, B. and Rootzén, H. (2004). *Extreme Values in Finance, Telecommunications, and the Environment*. Chapman & Hall/CRC, New York.

[14] Frahm, G. (2006). On the extreme dependence coefficient of multivariate distributions. *Probability & Statistics Letter*, 76:1470-1481.

[15] Genest, C. and Rivest, L.-P. (1989). A characterization of Gumbel's family of extreme value distributions. *Statistics and Probability Letters*, 8:207-211.

[16] Hartmann, P., Straetmans, S., and de Vries, C. (2005). Banking system stability: A cross-Atlantic perspective, ECB Working Paper # 527 (`http://www.ecb.int/pub/`).

[17] Hult, H. and Lindskog, F. (2002). Multivariate extremes: Aggregation and dependence in elliptical distributions. *Adv. Appl. Probab*, 34:587-608.

[18] Joe, H. (1997). *Multivariate Models and Dependence Concepts*. Chapman & Hall, London.

[19] Joe, H., Li, H. and Nikoloulopoulos, A.K. (2008). Tail dependence functions and vine copulas. *Journal of Multivariate Analysis*, 101(1):252-270.

[20] Kimberling, C. H. (1974). A probabilistic interpretation of complete monotonicity. *Aequationes Math*, 10:12-164.

[21] Klüppelberg, C., Kuhn, G. and Peng, L. (2008). Semi-parametric models for the multivariate tail dependence function-the asymptotically dependent. *Scandinavian of Journal of Statistics*, 35(4):701-718.

[22] Kortschak, D. and Albrecher, H. (2009). Asymptotic results for the sum of dependent non-identically distributed random variables. *Methodol. Comput. Appl. Probab.* 11:279-306.

[23] Kruskal, W. H. (1958). Ordinal measures of association. *J. Amer. Statist. Assoc.*, 53:814-861.

[24] Kousky, C. and Cooke, R. M. (2009). Climate change and risk management: challenges for insurance, adaptation, and loss estimation. RFF Discussion Paper No. 09-03-REV (`http://ssrn.com/paper=1346387`).

[25] Li, H. (2007). Tail dependence comparison of survival Marshall-Olkin copulas. *Methodology and Computing in Applied Probability*, 10(1):39–54.

[26] Li, H. and Sun, Y. (2009). Tail dependence for heavy-tailed scale mixtures of multivariate distributions. *Journal of Applied Probability*, 46(4):925-937.

[27] Marshall, A. W. and Olkin, I. (1967). A multivariate exponential distribution. *J. Amer. Statist. Assoc.*, 2:84-98.

[28] McNeil, A. J., Frey, R., Embrechts, P. (2005). *Quantitative Risk Management: Concepts, Techniques, and Tools*. Princeton University Press, Princeton, New Jersey.

[29] Nelsen, R. (2006). *An Introduction to Copulas*. Springer, New York.

[30] Nikoloulopoulos, A.K., Joe, H. and Li, H. (2009). Extreme value properties of multivariate t copulas. *Extremes*, 12(2):129-148.

[31] Resnick, S. (1987). *Extreme Values, Regular Variation, and Point Processes*. Springer, New York.

[32] Resnick, S. (2007). *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling*. Springer, New York.

[33] Schmidt, R. (2002). Tail dependence for elliptically contoured distributions. *Mathematical Methods of Operations Research*, 55:301-327.

[34] Shaked, M. and Shanthikumar, J. G. (2007). *Stochastic Orders*. Springer, New York.

[35] Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris*, 8:229–231.

[36] Sun, Y. and Li, H. (2010). Tail Approximation of Value-at-Risk under Multivariate Regular Variation, *Statistics & Probability Letters*, (under review).

[37] Whitehouse, M. (2005). How a formula ignited market that burned some big investors. *The Wall Street Journal*, September 12, 2005.

[38] Widder, D. V. (1941). *The Laplace Transform*. Princeton University Press, Princeton.