TECHNISCHE UNIVERSITÄT WIEN
Vienna University of Technology

DIPLOMARBEIT

Multivariate Extremes and Dependence Structures: A Theoretical Background for Modelling

Ausgeführt am Institut für

Wirtschaftsmathematik

der Technischen Universität Wien

unter der Anleitung von

Univ.-Prof. Dr. Uwe Schmock

durch

Johannes Heiny, B.Sc.

Leystraße 168

1020 Wien

Österreich

Wien, am 21. November 2013

(Unterschrift)

DIPLOMA THESIS

Multivariate Extremes and Dependence Structures: A Theoretical Background for Modelling

Author:

Johannes Heiny, B.Sc.

Supervisor:

Univ.-Prof. Dr. Uwe Schmock

Abstract

Many fields of modern science have to deal with events which are rare but of outstanding importance. *Extreme value theory* is a practical and useful mathematical tool for modelling events which occur with very small probability. In a wide variety of applications these extreme events have an inherently multivariate character. This thesis provides an overview of the relevant theoretical results for modelling multivariate extremes and their dependence structures. We study *multivariate extreme value distributions* (MEVDs) and characterise their *maximum domain of attraction* (MDA). We state the relationships between four equivalent representations of MEVDs which can be used as a basis for estimation. Moreover we look at *tail dependence coefficients* and provide information about the underlying dependence. The central result is the multivariate extension of the Fisher–Tippett theorem, which basically says that the maximum domain of attraction of a MEVD is characterised by the univariate MDAs of its margins and a so-called *copula domain of attraction* (CDA) of its copula. We construct explicit examples of copulas which are in no CDA and describe models for the tail of a multivariate distribution function. In order to facilitate model building, some methods to construct new extreme value copulas from known ones are presented.

Key words: *multivariate extreme value distribution, extreme value copula, copula domain of attraction, extreme value theory, tail dependence.*

Acknowledgements

This thesis would not have been possible without the help and support of the people around me. First of all I want to thank my supervisor, Prof. Uwe Schmock, for challenging me to explore the fascinating theory of extremes in depth and for providing an opportunity to exchange some ideas with practitioners.

I have always received invaluable support of my parents, Helga and Franz Heiny, who created a loving environment for my brother and me.

Furthermore, I also want to thank my brother Michael for cheering me up whenever needed.

Special thanks go to my friend Arpad Pinter for taking the time to carefully read this work.

Johannes Heiny

Introduction

Extreme value theory (EVT) is a branch of statistics that deals with unusual or rare events and gives a scientific approach to pure guesswork. It provides a set of tools that help us to deduce reasonable conclusions from sparse data.

Classical statistical analyses often give primacy to averages and central moments. To force data to fit a certain model, extremes are often labelled as outliers and even ignored. Such an approach might be acceptable when seeking information about everyday events, but fails miserably here.

EVT tries to model events that occur with a relatively small probability but have decisive consequences. For the layman such surprising phenomena seem to follow no rule, but "careful analysis has helped to discover distributions that acceptably model these extreme events." However, it is no silver bullet as the following experts pointed out.

> "*There is always going to be an element of doubt, as one is extrapolating into areas one doesn't know about. But what EVT is doing is making the best use of whatever data you have about extreme phenomena.*" (Richard Smith)

> "*The key message is that EVT cannot do magic but it can do a whole lot better than empirical curve fitting and guesswork. My answer to the sceptics is that if people aren't given well founded methods like EVT, they'll just use dubious ones instead.*" (Jonathan Tawn)

Many applications require the prediction of rare events, often outside the range of available data. EVT provides a framework that enables such extrapolation. But one has to be careful. Relatively small model changes probably have astonishingly huge effects on extrapolation. Thus, we have to compare different parameter estimates and measures of uncertainty when working with parametric models. The models for the tail of a multivariate distribution function we will present are parametric. In general, the results provided by extreme value analyses are subject to the following restrictions (cf. [3]):

- As in many other statistical applications, the models are developed using asymptotic arguments. Therefore, great care is required when interpreting the results for finite samples.
- The models are based on idealised circumstances, which may not be reasonable, exact or verifiable.

The statistical procedures obtained through EVT have many applications in hydrology, reliability theory, finance and insurance. Think for instance of the value–at–risk and catastrophic claims.

This thesis is about multivariate extremes. Most of the EVT literature treats the one-dimensional case. The reason might be the lack of a standard definition of order in a multi-dimensional space like $\mathbb{R}^d$. In this work we consider the vector of componentwise maxima. The aim of the thesis is to present the theory of multivariate extremes from an applied mathematical point of view. This means that on the one hand the reader can get a thorough understanding of the mathematical aspects. We discuss the critical assumptions and the wide range of possible outcomes. We also included extensive motivation and instructive examples for the achieved results and definitions. On the other hand the reader is able to find very practical ideas and help on how to set up a model in Chapter 6.

We now describe the contents of this thesis in more detail. Chapter 1 introduces the concept of copulas and summarises some properties. Every copula gives rise to a survival copula. Readers who are familiar with copulas can skip this chapter and start with Chapter 2.

In Chapter 2 we formulate the multivariate version of the maximum domain of attraction (MDA). We work in a d-dimensional space. *Multivariate extreme value distributions* (MEVDs) arise as the limit distributions of properly normalised componentwise maxima. A MEVD can be standardised in order to work with simpler margins. By correcting a result in [14], we explicitly state this standardisation for Fréchet margins. The class of multivariate extreme value distributions coincides with the class of *max-stable* distribution functions with non-degenerate margins. This first characterisation is not very helpful in practice. Therefore we provide three other representations of MEVDs, which can be used as a basis for statistical modelling and estimation. We state the relationships between these equivalent representations and illustrate them in a comprehensive example of the parametric Gumbel model.

In Chapter 3 we study conditional probabilities of high quantile exceedances. *Tail dependence coefficients* are scalar measures of dependence in the tails of a bivariate distribution. They provide a basic classification of the existing extremal dependence and can be expressed via copulas. We look at two kinds of tail dependence coefficients and show how they complement each other. The bivariate normal distribution with correlation less than 1 is asymptotically independent, meaning that extreme events of the components happen independently.

Chapter 4 focuses on the limits of copulas of componentwise maxima of independent identically distributed (iid) random sequences, which can naturally be considered to be appropriate for the dependence structure between extreme events in the components. They are called *extreme value copulas* (EV copulas). Similar to MEVDs, the class of EV copulas does not permit a finite-dimensional parametrisation. Using the results of Chapter 2, we again find many equivalent descriptions of EV copulas. A MEVD can be split into its margins, which are univariate EVDs, and its copula, which has to be an EV copula. The central *multivariate extension of the Fisher–Tippett theorem* basically says that the maximum domain of attraction of a MEVD is characterised by the univariate MDAs of its margins and a so-called *copula domain of attraction* (CDA) of its copula. An example illustrates how a verification of these two conditions can be approached. The tail dependence coefficients of a copula provide a lot of information about the corresponding EV copula. From a theoretical point of view, asymptotically dependent distributions are more interesting than asymptotically independent ones because they lead to EV copulas different from the independence copula.

Section 4.4 is of crucial importance for a thorough understanding of the copula domain of attraction. It shows that there do exist copulas which are in no CDA. From the definition of the CDA this is far from obvious. Nevertheless this issue is not addressed in the vast majority of EVT literature. The author of this thesis constructed two different kinds of examples of a copula C such that $C \notin \mathrm{CDA}(C_0)$ for every EV copula C_0. Especially the first example is very intuitive and instructive. The second example is more general and forms a basis for further probabilistic counterexamples.

Motivated by univariate EVT, Chapter 5 focuses on multivariate threshold exceedances and offers a fully parametric model for the tail of a distribution. Moreover, we condition a random vector to exceed high thresholds and study the arising copulas, the so-called *threshold copulas*. We illustrate an alternative way to find the corresponding EV copula.

Chapter 6 is devoted to the practical considerations and challenges the user of MEVT has to face. First, we present an approach based on the generalised extreme value distribution and block maxima. It turns out that the availability of a wide variety of EV copulas is crucial for its implementation. Therefore some methods to construct new extreme value copulas from known ones are described. *Product copulas* and *nested copulas* are useful to introduce asymmetry. Copula estimation and testing for extremes are discussed. Finally, we consider violations of the standard assumptions and highlight the strengths and limitations of EVT.

Appendix A briefly summarises the most important results of univariate extreme value theory.

Johannes Heiny

Contents

Chapter 1

Copulas

The aim of this short chapter and Appendix A is to gather important results on dependence modelling and univariate extreme value theory, which will be referred to in the main chapters numerous times.

This thesis deals with extremes of random variables with values in $\mathbb{R}^d$, where $d \in \mathbb{N}$. A copula describes the dependence structure of a multivariate distribution function (df).

Definition 1.1 (copula). A d-dimensional **copula** is a df on $[0,1]^d$ with standard uniform marginal distributions.

Theorem 1.2 (Sklar). [12, p. 186]
Let F be a d-dimensional df with margins $F_1, \dots, F_d$. Then there exists a copula C such that for all $x_1, \dots, x_d \in \bar{\mathbb{R}} = [-\infty, \infty]$,

$$F(x_1, \dots, x_d) = C(F_1(x_1), \dots, F_d(x_d)). \tag{1.1}$$

If the margins are continuous, then C is unique. Otherwise C is uniquely determined on Ran $F_1 \times \cdots \times$ Ran F_d.
Conversely, for a copula C and continuous margins $F_1, \dots, F_d$, the function F defined in (1.1) is a d-dimensional df with margins $F_1, \dots, F_d$.

Definition 1.3 (generalised inverse). The **generalised inverse** of a non-decreasing function $f : \mathbb{R} \to \mathbb{R}$ is defined by

$$f^{\leftarrow}(t) := \inf \{x \in \mathbb{R} : f(x) \geq t\}$$

with the convention $\inf \varnothing := \infty$.

From (1.1) one obtains with $x_i = F^{\leftarrow}(u_i)$ that

$$C(u_1, \dots, u_d) = F(F_1^{\leftarrow}(u_1), \dots, F_d^{\leftarrow}(u_d)), \quad (u_1, \dots, u_d) \in [0,1]^d. \tag{1.2}$$

Let $\boldsymbol{X} = (X_1, \dots, X_d) \sim F$ with margins $F_1, \dots, F_d$. Since

$$F(F_1^{\leftarrow}(u_1), \dots, F_d^{\leftarrow}(u_d)) = \mathbb{P}(X_1 \leq F_1^{\leftarrow}(u_1), \dots, X_d \leq F_d^{\leftarrow}(u_d)), \quad (u_1, \dots, u_d) \in [0,1]^d,$$

we can interpret $C(u_1, \dots, u_d)$ as the joint probability that all components of $\boldsymbol{X}$ stay at or below their lower u_i-quantiles of their margins.

Definition 1.4 (copula of F). Let the random vector $\boldsymbol{X} = (X_1, \ldots, X_d)$ have the joint df F with continuous marginal dfs $F_1, \ldots, F_d$. The copula of F (or $\boldsymbol{X}$) is the df C of $(F_1(X_1), \ldots, F_d(X_d))$.

Strictly increasing transformations do not change the copula. Let $\boldsymbol{X}$ be a d-dimensional random vector with continuous margins and copula C, and let $T_1, \ldots, T_d$ be strictly increasing functions. Then $(T_1(X_1), \ldots, T_d(X_d))$ also has copula C.

Remark 1.5. The above result justifies the transformation of margins in a multivariate extreme value (MEV) analysis. Suppose we prefer working with the continuous marginal dfs $G_1, \ldots, G_d$ over the given strictly increasing dfs $F_1, \ldots, F_d$. By applying the strictly increasing transformations $G_i^{\leftarrow} \circ F_i$ to the components of $\boldsymbol{X}$, we could consider the vector

$$(G_1^{\leftarrow} \circ F_1(X_1), \ldots, G_d^{\leftarrow} \circ F_d(X_d)),$$

which not only has the required margins, but also the same copula as $\boldsymbol{X}$. A popular choice for G_i is the standard Fréchet df, which we define below. An application can be found in Theorem 2.5 later on.

For $\alpha > 0$ the df of the Fréchet distribution Φ_α is given by

$$\Phi_\alpha(x) = \begin{cases} 0 & \text{if } x \leq 0, \\ \exp\{-x^{-\alpha}\} & \text{if } x > 0. \end{cases} \tag{1.3}$$

If $\alpha = 1$ we have the standard Fréchet distribution Φ_1. This distribution will appear numerous times in this thesis.

Proposition 1.6. [13, Theorem 2.10.7]
Every d-dimensional copula C is Lipschitz continuous. More precisely,

$$|C(u_1, \ldots, u_d) - C(v_1, \ldots, v_d)| \leq \sum_{i=1}^{d} |u_i - v_i|, \quad (u_1, \ldots, u_d), (v_1, \ldots, v_d) \in [0,1]^d. \tag{1.4}$$

The following bounds hold for every copula. The lower bound will be refined for the class of extreme value copulas in (2.8).

Proposition 1.7 (Fréchet bounds). [12, Theorem 5.7]
Every d-dimensional copula C fulfills the inequalities

$$\max\left\{\sum_{i=1}^{d} u_i + 1 - d, 0\right\} \leq C(u_1, \ldots, u_d) \leq \min\{u_1, \ldots, u_d\}, \quad (u_1, \ldots, u_d) \in [0,1]^d. \tag{1.5}$$

The upper Fréchet bound is called comonotonicity copula. It represents the case of complete dependence. Let U be a standard uniform random variable. Then the components of the d-dimensional random vector $(U, \ldots, U)$ are completely dependent and its copula is the upper bound in (1.5). The lower bound in (1.5), however, only is a copula if $d = 2$.

1.1 Survival copulas

Definition 1.8 (survival copula). Let C be a d-dimensional copula and let the random vector $\boldsymbol{U} \sim C$. The **survival copula** $\widehat{C}$ of C is the df of the random vector $\mathbf{1} - \boldsymbol{U}$.

Remark 1.9. Since $\mathbf{1} - \boldsymbol{U}$ has standard uniform margins, the df $\widehat{C}$ is a proper copula.

For a d-dimensional df F with margins $F_1, \ldots, F_d$ and copula C, let $(X_1, \ldots, X_d) \sim F$. Define

$$\bar{F}(\boldsymbol{x}) := \mathbb{P}(X_1 > x_1, \ldots, X_d > x_d), \quad \boldsymbol{x} = (x_1, \ldots, x_d) \in \mathbb{R}^d,$$

and $\bar{F}_i := 1 - F_i$ for $i = 1, \ldots, d$. The survival copula $\widehat{C}$ of C satisfies

$$\bar{F}(x_1, \ldots, x_d) = \widehat{C}(\bar{F}_1(x_1), \ldots, \bar{F}_d(x_d)), \quad \boldsymbol{x} \in \mathbb{R}^d. \tag{1.6}$$

In case of continuous margins

$$\bar{F}(\boldsymbol{x}) = \mathbb{P}(1 - F_1(X_1) \leq \bar{F}_1(x_1), \ldots, 1 - F_d(X_d) \leq \bar{F}_d(x_d)),$$

where $1 - F_i(X_i) \sim U(0, 1)$. Therefore $\widehat{C}$ is the df of $(1 - F_1(X_1), \ldots, 1 - F_d(X_d))$.

For a two-dimensional copula C and its corresponding survival copula $\widehat{C}$, we have the useful relationship

$$\widehat{C}(u_1, u_2) = u_1 + u_2 - 1 + C(1 - u_1, 1 - u_2), \quad (u_1, u_2) \in [0, 1]^2. \tag{1.7}$$

Chapter 2

Multivariate Extremes

Many applications that require the prediction of rare events involve multivariate data. In the univariate case our intuition tells us how to define records or extreme values. In the multivariate case, however, we lack a natural notion of order as several different concepts of ordering are possible.

A particular combination of the components, rather than the individual components themselves, might be of interest. In general these components influence each other, which explains the importance of modelling their dependence structure. Unfortunately, as Coles [3] points out, "some multivariate processes have a strength of dependence that weakens at high levels, to the extent that the most extreme events are near-independent."

In this chapter we consider the vector of componentwise maxima and describe all possible limiting multivariate extreme value distributions. The reader is assumed to be familiar with the basic results from univariate extreme value theory. A brief overview can be found in the Appendix. From now on $\boldsymbol{X}$ is d-dimensional random vector with joint distribution function (df) F and margins $F_1, \ldots, F_d$. Furthermore, $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots$ is a corresponding iid sequence. The vector $\boldsymbol{X}_i$ has d (univariate) components:

$$\boldsymbol{X}_i = (X_{1,i}\ , \ldots, X_{d,i})^\top, \quad i \in \mathbb{N}.$$

2.1 The model

Our aim is to model the statistical behaviour of

$$\boldsymbol{M}_n := \left(\max_{1 \le i \le n} X_{1,i}\ , \ldots, \max_{1 \le i \le n} X_{d,i} \right)^\top, \quad n \in \mathbb{N},$$

the vector of componentwise block maxima. The vector $\boldsymbol{M}_n$ does not necessarily correspond to any of the observations $\boldsymbol{X}_i$. The df of the random vector $\boldsymbol{M}_n$ is F^n, the nth power of F.

Remark 2.1. Without loss of generality we can assume all margins to be non-degenerate. For a degenerate random variable $X_{d,i}$, i.e. there exists $x_0 \in \mathbb{R}$ with $\mathbb{P}(X_{d,i} = x_0) = 1$, the df F factorises into $F = G \cdot F_d$, where G denotes the joint df of the first $d-1$ components.

Notation: For vectors $\boldsymbol{x} = (x_1, \ldots, x_d)^\top$ and $\boldsymbol{y} = (y_1, \ldots, y_d)^\top \in \mathbb{R}^d$ arithmetic operations and relations are understood componentwise.

- $\boldsymbol{x}\,\boldsymbol{y} = (x_1 y_1, \ldots, x_d y_d)^\top$.
- $\frac{\boldsymbol{x}}{\boldsymbol{y}} = (\frac{x_1}{y_1}, \ldots, \frac{x_d}{y_d})^\top$ where $\boldsymbol{y} \neq \boldsymbol{0}$.
- $\boldsymbol{x} + \boldsymbol{y} = (x_1 + y_1, \ldots, x_d + y_d)^\top$.
- $\boldsymbol{x} \leq \boldsymbol{y}$ means $x_i \leq y_i$ for all $i \in \{1, \ldots, d\}$.
- $\boldsymbol{x} < \boldsymbol{y}$ means $x_i < y_i$ for all $i \in \{1, \ldots, d\}$.
- $\boldsymbol{x} \not< \boldsymbol{y}$ means there exists an $i \in \{1, \ldots, d\}$ with $x_i \geq y_i$.

Multivariate extreme value distributions are defined analogously to the one-dimensional case.

Definition 2.2 (MEVD, MDA). Let F be a d-dimensional df. Suppose there exist a d-dimensional df G with non-degenerate margins, and normalising sequences $\boldsymbol{a}_n > \boldsymbol{0}$ and $\boldsymbol{b}_n \in \mathbb{R}^d$ for $n \in \mathbb{N}$ such that

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{\boldsymbol{M}_n - \boldsymbol{b}_n}{\boldsymbol{a}_n} \leq \boldsymbol{x}\right) = \lim_{n\to\infty} F^n(\boldsymbol{a}_n \boldsymbol{x} + \boldsymbol{b}_n) = G(\boldsymbol{x}) \tag{2.1}$$

for all continuity points $\boldsymbol{x} \in \mathbb{R}^d$ of G, then G is called a **multivariate extreme value distribution** (MEVD) and F is said to be in the **maximum domain of attraction** of G. We use the notation $F \in \mathrm{MDA}(G)$.

Remark 2.3.

1. If the df G is no multivariate extreme value distribution, $\mathrm{MDA}(G) := \varnothing$.
2. The convergence in distribution of the random vector $\frac{\boldsymbol{M}_n - \boldsymbol{b}_n}{\boldsymbol{a}_n}$ in (2.1) implies the convergence of its margins. Consequently, the margin F_i is in the MDA of a univariate extreme value distribution (EVD) and a possible choice for the normalising sequences of F_i in the *Fisher–Tippett theorem* (Theorem A.1) is $a_n := a_{i,n}$ and $b_n := b_{i,n}$, $n \in \mathbb{N}$.
3. A MEVD G necessarily has univariate EVDs as margins. Since the marginal dfs of G are continuous (see Theorem $A.1$) and since the copula of G is continuous by Proposition (1.6), G is continuous too. By Sklar's theorem, G has a unique copula.

Example 2.4 (Fréchet margins). Let $d = 2$ and F be a df with margins $F_1(x) = F_2(x) = \exp(-x^{-1}) = \Phi_1(x)$, for $x > 0$, the standard Fréchet df. Then

$$\mathbb{P}\left(\frac{M_{1,n}}{n} \leq x\right) = \mathbb{P}\left(\frac{M_{2,n}}{n} \leq x\right) = F_2^n(nx) = \exp(-x^{-1}), \quad x > 0,$$

which shows $\boldsymbol{a}_n = (n, n)^\top$ and $\boldsymbol{b}_n = (0, 0)^\top$, $n \in \mathbb{N}$. Next we consider three possible choices for the joint df F.

(a) Let $F_a(x_1, x_2) = F_1(x_1)F_2(x_2)$ for $(x_1, x_2) \in \mathbb{R}^2$. By calculating

$$F_a^n(\boldsymbol{a}_n \boldsymbol{x} + \boldsymbol{b}_n) = F_1^n(nx_1)F_2^n(nx_2) = F_a(x_1, x_2), \quad n \in \mathbb{N},$$

we have found a first MEVD.

(b) Let $F_b(x_1, x_2) = \exp\left\{-\left(x_1^{-\alpha} + x_2^{-\alpha}\right)^{1/\alpha}\right\}$, $(x_1, x_2) \in \mathbb{R}^2$ and $1 \leq \alpha < \infty$. This defines a df because the underlying copula is the well known Gumbel copula. Again

$$F_b^n(nx_1, nx_2) = \exp\left\{-n^{-1}\left(x_1^{-\alpha} + x_2^{-\alpha}\right)^{1/\alpha}\right\}^n = F_b(x_1, x_2), \quad n \in \mathbb{N},$$

hence F_b is a MEVD. In general, for a MEVD F it holds $F \in \text{MDA}(F)$, cf. Theorem 4.9 later on.

(c) Let $F_c(x_1, x_2) = F_1(x_1)F_2(x_2)[1 + (1 - F_1(x_1))(1 - F_2(x_2))]$, $\boldsymbol{x} > \boldsymbol{0}$, which is a df. Now

$$\begin{aligned} F_c^n(nx_1, nx_2) &= F_1(x_1)F_2(x_2)[1 + (1 - F_1(nx_1))(1 - F_2(nx_2))]^n \\ &\to F_1(x_1)F_2(x_2) = F_a(x_1, x_2) \end{aligned}$$

for $n \to \infty$, thus $F_c \in \text{MDA}(F_a)$.

Taking EV distributions as margins simplified the calculations a lot. In fact, we did not have to evaluate a limit in *(a)* and *(b)*, a property shared by all max-stable dfs. ∘

The next result shows that one can easily standardise the problem in order to work with simpler margins of G. Example 2.4 motivates to use standard Fréchet margins. The standard Fréchet df is $\Phi_1(x) = \exp(-x^{-1})$ for $x > 0$, see also (1.3).

Suppose G is a multivariate df with continuous margins $G_1, \ldots, G_d$ and $\boldsymbol{Y} \sim G$. According to Remark 1.5 the vector

$$(\Phi_1^{\leftarrow}(G_1(Y_1)), \ldots, \Phi_1^{\leftarrow}(G_d(Y_d)))$$

has df

$$\widetilde{G}(\boldsymbol{x}) = G(G_1^{\leftarrow}(\Phi_1(x_1)), \ldots, G_d^{\leftarrow}(\Phi_1(x_d))), \quad \boldsymbol{x} \geq \boldsymbol{0}, \tag{2.2}$$

and standard Fréchet margins. Now set the function $U_i := \frac{1}{1-F_i}$, $i \in \{1, \ldots, d\}$. Since U_i has range $[1, \infty]$, the generalised inverse $U_i^{\leftarrow}$ has domain $[1, \infty]$. By a direct calculation one can show that $U_i^{\leftarrow}(x) = F_i^{\leftarrow}(1 - 1/x)$ for $x \geq 1$. The vector $(U_1(X_1), \ldots, U_d(X_d))$ has df

$$\widehat{F}(\boldsymbol{x}) = F(U_1^{\leftarrow}(x_1), \ldots, U_d^{\leftarrow}(x_d)), \quad \boldsymbol{x} \in [1, \infty]^d.$$

Theorem 2.5. (cf. [14, Proposition 5.10])
Given a d-dimensional df G and using the notation from above, (1) implies (2), where

(1) G is a MEVD,

(2) $\widetilde{G}$ is a MEVD with standard Fréchet margins.

For a d-dimensional df F an even more general result is: *(i)* implies *(ii)*, where

(i) $F \in \text{MDA}(G)$,

(ii) $\widehat{F} \in \text{MDA}(\widetilde{G})$ and $\widehat{F}_i \in \text{MDA}(\Phi_1)$ for $i = 1, \ldots, d$.

Proof. This proof is an extended version of the proof given in [14]. Suppose we know that *(i)* $\Rightarrow$ *(ii)* holds. Since every MEVD G satisfies $G \in \mathrm{MDA}(G)$, cf. Theorem 4.9, we can take $F = G$ in *(i)* to conclude with *(ii)* that $\widetilde{G}$ is a MEVD with standard Fréchet margins. Therefore, we only need to show *(i)* $\Rightarrow$ *(ii)*.

Let $F \in \mathrm{MDA}(G)$ with normalising sequences $\boldsymbol{a}_n > \boldsymbol{0}$ and $\boldsymbol{b}_n \in \mathbb{R}^d$ for $n \in \mathbb{N}$. Let $i \in \{1, \dots, d\}$. Convergence of margins means

$$\lim_{n\to\infty} F_i^n(a_{i,n}x + b_{i,n}) = G_i(x), \quad x \in \mathbb{R}. \tag{2.3}$$

We derive an equivalent limit relation to (2.3) for all $x \in \mathbb{R}$. For this aim we first consider only $x \in \mathbb{R}$ with $G_i(x) \in (0,1)$. We start by taking logarithms and obtain

$$\lim_{n\to\infty} n \cdot \log F_i(a_{i,n}x + b_{i,n}) = \log G_i(x). \tag{2.4}$$

It follows that $F_i(a_{i,n}x + b_{i,n}) \longrightarrow 1$ as $n \to \infty$. By using $\log(1+y) \sim y$ for $y \to 0$, meaning that the quotient of the two expressions approaches 1, we get

$$\log F_i(a_{i,n}x + b_{i,n}) \sim F_i(a_{i,n}x + b_{i,n}) - 1, \quad n \to \infty.$$

Accordingly (2.4) is equivalent to

$$\lim_{n\to\infty} n\,\bar{F}_i(a_{i,n}x + b_{i,n}) = -\log G_i(x) \tag{2.5}$$

for all $x \in \mathbb{R}$ with $G_i(x) \in (0,1)$. In the case of an x such that $G_i(x) = 0$ we can write equation (2.3) as

$$\lim_{n\to\infty} F_i^n(a_{i,n}x + b_{i,n}) = \lim_{n\to\infty} \left(1 - \frac{n\,\bar{F}_i(a_{i,n}x + b_{i,n})}{n}\right)^n = 0.$$

Since $\left(1 - \frac{K}{n}\right)^n \to \exp(-K)$ for every $K > 0$ when $n \to \infty$, there exists for every $K > 0$ an $n_K \in \mathbb{N}$ such that $m\,\bar{F}_i(a_{i,m}x + b_{i,m}) > K$ for every $m > n_K$. In other words, this means that $n\,\bar{F}_i(a_{i,n}x + b_{i,n}) \to \infty$ which establishes (2.5), where we used the convention $\log 0 := -\infty$. Similarly one sees that (2.5) also holds for x with $G_i(x) = 1$.

Taking reciprocals in (2.5) gives us for all $x \in \mathbb{R}$ that

$$\frac{1}{n\,\bar{F}_i(a_{i,n}x + b_{i,n})} = \frac{U_i(a_{i,n}x + b_{i,n})}{n} \longrightarrow \frac{1}{-\log G_i(x)} \quad \text{as } n \to \infty. \tag{2.6}$$

Next we invert both sides of (2.6):

$$\begin{aligned}\left(\frac{U_i(a_{i,n}\cdot + b_{i,n})}{n}\right)^{\leftarrow}(x) &= \inf\left\{t \;:\; \frac{U_i(a_{i,n}t + b_{i,n})}{n} \geq x\right\} \\ &= \inf\left\{t \;:\; a_{i,n}t + b_{i,n} \geq U_i^{\leftarrow}(nx)\right\} \\ &= \frac{U_i^{\leftarrow}(nx) - b_{i,n}}{a_{i,n}}, \quad x > 0,\end{aligned}$$

and

$$\left(\frac{1}{-\log G_i(\cdot)}\right)^{\leftarrow}(x) = G_i^{\leftarrow}(\Phi_1(x)), \quad x > 0.$$

Hence we get from (2.6) that

$$\lim_{n\to\infty} \frac{U_i^{\leftarrow}(nx) - b_{i,n}}{a_{i,n}} = G_i^{\leftarrow}(\Phi_1(x)), \quad x > 0.$$

We are ready to compute the limiting distribution for $\widehat{F}$. We find that

$$\begin{aligned}\lim_{n\to\infty} \widehat{F}^n(n\boldsymbol{x}) &= \lim_{n\to\infty} \mathbb{P}\left(M_{i,n} \le U_i^{\leftarrow}(nx_i)\ ,\ i = 1,\dots,d\right)\\ &= \lim_{n\to\infty} \mathbb{P}\left(\frac{M_{i,n} - b_{i,n}}{a_{i,n}} \le \frac{U_i^{\leftarrow}(nx_i) - b_{i,n}}{a_{i,n}}\ ,\ i = 1,\dots,d\right)\\ &= \widetilde{G}(\boldsymbol{x})\end{aligned}$$

for $\boldsymbol{x} > \boldsymbol{0}$, completing the proof. □

Remark 2.6. Theorem 2.5 is a corrected version of Theorem 5.10 in [14], which unfortunately is a frequently cited result in EVT literature. The problem is that in [14] the equivalence of (1) and (2) is claimed. We show the opposite by a counterexample. Let

$$H(x_1, x_2) = N(x_1)N(x_2), \quad (x_1, x_2) \in \mathbb{R}^2,$$

be the df of two independent standard normal random variables, where N denotes the standard normal df. The df H obviously is no MEVD. Now we transform the margins of H according to formula (2.2) and get

$$\begin{aligned}\widetilde{H}(x_1, x_2) &= H\left(N^{\leftarrow}(\Phi_1(x_1)), N^{\leftarrow}(\Phi_1(x_2))\right)\\ &= N(N^{\leftarrow}(\Phi_1(x_1)))N(N^{\leftarrow}(\Phi_1(x_2)))\\ &= \Phi_1(x_1)\Phi_1(x_2), \quad (x_1, x_2) \in \mathbb{R}^2.\end{aligned}$$

Since $\widetilde{H}$ is a MEVD with standard Fréchet margins, we conclude that (2) does not imply (1).

However, with the additional assumption $F_i \in \mathrm{MDA}(G_i)$ for $i \in \{1,\dots,d\}$ it is possible to show the other implication *(ii)* ⇒ *(i)*. In this case the normalising sequences can be chosen as $\boldsymbol{a}_n = \boldsymbol{n}$ and $\boldsymbol{b}_n = \boldsymbol{0} \in \mathbb{R}^d$ for $n \in \mathbb{N}$. Using (2.6) and the definition of $\widetilde{G}$, one finds

$$\begin{aligned}\lim_{n\to\infty} F^n(n\boldsymbol{x}) &= \lim_{n\to\infty} \widehat{F}^n\left(n\frac{U_1(nx_1)}{n}, \dots, n\frac{U_d(nx_d)}{n}\right)\\ &= \widetilde{G}\left(\frac{1}{-\log G_1(x_1)}, \dots, \frac{1}{-\log G_d(x_d)}\right)\\ &= \widetilde{G}\left(\Phi_1^{\leftarrow}(G_1(x_1)), \dots, \Phi_1^{\leftarrow}(G_d(x_d))\right)\\ &= G(\boldsymbol{x}), \quad \boldsymbol{x} \in \mathbb{R}^d.\end{aligned}$$

If the function $\widetilde{G}$ in formula (2.2) is a MEVD, it has an interesting property (cf. Example 2.4 *(a)* and *(b)*):

$$\widetilde{G}^n(n\boldsymbol{x}) = \widetilde{G}(\boldsymbol{x}), \quad \text{for } n \in \mathbb{N} \text{ and } \boldsymbol{x} > \boldsymbol{0}.$$

This basically follows from the facts that the standard Fréchet distribution function satisfies $\Phi_1^n(nx) = \Phi_1(x)$ for $n \in \mathbb{N}$ and $x \in \mathbb{R}$, and that the copula of a MEVD has the stability property (4.6). Setting $\boldsymbol{x} = \frac{1}{n}\boldsymbol{y}$ we then get $\widetilde{G}(\boldsymbol{y}) = \widetilde{G}^{1/n}(\frac{1}{n}\boldsymbol{y})$ and by the continuity of the MEVD $\widetilde{G}$,

$$\widetilde{G}(\boldsymbol{y}) = \widetilde{G}^t(t\boldsymbol{y}), \quad \text{for } t \in \mathbb{R}_+ \text{ and } \boldsymbol{y} > \boldsymbol{0}. \tag{2.7}$$

In particular, $\widetilde{G}^t$ is a df for all real $t > 0$.

Definition 2.7 (max-id). A df F is is called **max-infinitely divisible** (max-id) if F^t is a df for all $t > 0$.

In contrast to univariate dfs, not every multivariate df is max-id.

Example 2.8 (bivariate case). Let the bivariate df F satisfy

$$F(0,0) = 0, \quad F(1,0) = F(0,1) = \frac{1}{4},$$

and suppose there exists a random vector $Z \sim F^{1/3}$. We calculate

$$\mathbb{P}\left(Z \in \{(0,1] \times (-\infty,0]\} \cup \{(-\infty,0] \times (0,1]\}\right) = F^{1/3}(1,0) + F^{1/3}(0,1) = 1.2599 > 1,$$

a contradiction. $F^{1/3}$ is no proper df and F is not max-id. ◦

Resnick provides a nice criterion for a two-dimensional df F to be max-id, [14, p. 254].

Proposition 2.9. *A df F on $\mathbb{R}^2$ with continuous density $F_{x_1x_2}$ and F_{x_i} the partial derivative of F with respect to x_i, $i \in \{1,2\}$, is max-id if and only if $F_{x_1}F_{x_2} \leq F_{x_1x_2}$ a.s. on $\mathbb{R}^2$.*

Remark 2.10. Applying this result to the bivariate normal distribution with correlation ρ reveals max infinite divisibility if and only if $\rho \geq 0$, see [14].

A stronger property than max infinite divisibility is max stability.

Definition 2.11 (max-stable). A d-dimensional df F is is called **max-stable** if there exist functions $\boldsymbol{\alpha} : \mathbb{R} \to (0,\infty)^d$ and $\boldsymbol{\beta} : \mathbb{R} \to \mathbb{R}^d$ such that for every $t > 0$

$$F^t(\boldsymbol{x}) = F(\boldsymbol{\alpha}(t)\boldsymbol{x} + \boldsymbol{\beta}(t)), \quad \boldsymbol{x} \in \mathbb{R}^d.$$

Note that every univariate non-extreme value distribution is max-id but not necessarily max-stable. If G is a MEVD, the df $\widetilde{G}$ defined in (2.2) is max-stable with $\boldsymbol{\alpha}(t) = t^{-1}$ and $\boldsymbol{\beta}(t) = 0$ (see (2.7)).

We round off this section with a characterisation of all MEV distributions and a property of their dependence structure.

Theorem 2.12. The class of multivariate extreme value distributions coincides with the class of max-stable dfs with non-degenerate margins.

Proof. Consult [14, p. 264]. □

Every MEVD G is positively quadrant dependent, i.e.

$$G(\boldsymbol{x}) \geq G_1(x_1) \cdots G_d(x_d), \quad \boldsymbol{x} \in \mathbb{R}^d. \tag{2.8}$$

A proof of this interesting result can be found in Remark 4.14. The lower bound for G in (2.8) is a refinement of the lower Fréchet bound in Proposition 1.7. Moreover, the independence of the components (i.e. $G(\boldsymbol{x}) = G_1(x_1) \cdots G_d(x_d)$) is equivalent to the existence of a point $\boldsymbol{y} \in \mathbb{R}^d$ with $G_i(y_i) \in (0,1)$ for all $i = 1, \ldots, d$ such that $G(\boldsymbol{y}) = G_1(y_1) \cdots G_d(y_d)$. Complete dependence (i.e. $G(\boldsymbol{x}) = \min\{G_1(x_1), \ldots, G_d(x_d)\}$) is equivalent to the existence of a point $\boldsymbol{y} \in \mathbb{R}^d$ with $0 < G_1(y_1) = \ldots = G_d(y_d) < 1$ such that $G(\boldsymbol{y}) = G_1(y_1)$. The reference for the last two statements is [1, p. 266]

2.2 Representations of multivariate extreme value distributions

Unfortunately the characterisation of MEV distributions in Theorem 2.12 fails to be very useful in practice. Therefore this section is concerned with more explicit representations of MEV distributions, which form a starting point for model building and estimation procedures. The most modern approach via extreme value copulas, however, will be discussed in a separate chapter.

In the whole section G is a MEV distribution with margins $G_1, \ldots, G_d$. According to the Fisher–Tippett theorem, the marginal dfs belong to the three parameter generalised extreme value distribution family (see Definition A.2). Hence, it remains to shed light on the possible dependence structures, which cannot be captured in a finite-dimensional parametric family.

2.2.1 Stable tail dependence function ℓ

Definition 2.13 (stable tail dependence function). [1, p. 257]
The **stable tail dependence function** of a d-dimensional MEVD G is defined by

$$\ell(\boldsymbol{y}) = -\log G(G_1^{\leftarrow}(\mathrm{e}^{-y_1}), \ldots, G_d^{\leftarrow}(\mathrm{e}^{-y_d})), \quad \boldsymbol{y} \geq \mathbf{0}. \tag{2.9}$$

Given ℓ and univariate EVDs $G_1, \ldots, G_d$ we can recover the MEVD G through

$$G(\boldsymbol{x}) = \exp\left\{-\ell(-\log G_1(x_1), \ldots, -\log G_d(x_d))\right\}, \quad \boldsymbol{x} \in \mathbb{R}^d, \tag{2.10}$$

hence ℓ contains all the information about the underlying dependence. The function ℓ has properties (L1)–(L4) (cf. [1, p. 257]). We provide some heuristic arguments to illustrate (L1) and (L1), and prove (L3).

(L1) $\ell(t\boldsymbol{y}) = t\ell(\boldsymbol{y})$ for $t > 0$ and $\boldsymbol{y} \in [0, \infty]^d$.

Observe that, with the convention $\frac{1}{0} := \infty$, ℓ can be expressed by $\widetilde{G}$ of the previous section (see (2.2)):

$$\ell(\boldsymbol{y}) = -\log \widetilde{G}(\boldsymbol{y}^{-1}). \tag{2.11}$$

Using equation (2.7) we see that

$$\ell(t\boldsymbol{y}) = -\log \widetilde{G}(t^{-1}\boldsymbol{y}^{-1}) = -\log \widetilde{G}^t(\boldsymbol{y}^{-1}) = t\ell(\boldsymbol{y}).$$

(L2) $\ell(\boldsymbol{e}_i) = 1$ for $i = 1, \ldots, d$, where $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_d$ is the canonical basis of $\mathbb{R}^d$.

This property follows from (2.9). To verify it in a special case one could take the df $\widetilde{G}$ with standard Fréchet margins.

(L3) $\max\{y_1, \ldots, y_d\} \leq \ell(y) \leq y_1 + \cdots + y_d, \quad \boldsymbol{y} \in [0, \infty]^d.$

From Remark 3.10 we know that the MEVD $\widetilde{G}$ with Fréchet margins is positively quadrant dependent. Therefore we have

$$\Phi_1(x_1) \cdots \Phi_1(x_d) \leq \widetilde{G}(\boldsymbol{x}) \leq \min\{\Phi_1(x_1), \ldots, \Phi_1(x_d)\}, \quad \boldsymbol{x} \in \mathbb{R}^d,$$

and because of (2.11), we conclude that (L3) is an immediate consequence. Note that the bounds $\max\{y_1, \ldots, y_d\}$ and $y_1 + \cdots + y_d$ correspond to complete dependence (comonotonic copula) and independence, respectively.

(L4) ℓ is convex.

2.2.2 Spectral measure H

This section is based on the presentation in [1, Chapter 8]. MEV distributions are characterised by finite measures on the unit simplex

$$\boldsymbol{\mathcal{S}}_d := \{\boldsymbol{w} \in [0, \infty]^d \ : \ w_1 + \cdots + w_d = 1\}.$$

Proposition 2.14 (spectral measure). *For a d-dimensional extreme value distribution G with margins $G_1, \ldots, G_d$ there exists a finite measure H on $\boldsymbol{\mathcal{S}}_d$ with*

$$\int_{\boldsymbol{\mathcal{S}}_d} w_i \,\mathrm{d}H(\boldsymbol{w}) = 1, \quad i = 1, \ldots, d,$$

such that

$$G(\boldsymbol{x}) = \exp\left\{\int_{\boldsymbol{\mathcal{S}}_d} \min_{1\leq i\leq n} \{w_i \log G_i(x_i)\} \,\mathrm{d}H(\boldsymbol{w})\right\}, \quad \boldsymbol{x} \in \mathbb{R}^d. \tag{2.12}$$

The measure H is called **spectral measure** *and $H(\boldsymbol{\mathcal{S}}_d) = d$.*

Because of (2.12) and (2.9), the stable tail dependence function is uniquely determined by H,

$$\ell(\boldsymbol{y}) = \int_{\boldsymbol{\mathcal{S}}_d} \max_{1\leq i\leq n} \{w_i y_i\} \,\mathrm{d}H(\boldsymbol{w}), \quad \boldsymbol{y} \geq \boldsymbol{0}. \tag{2.13}$$

Conversely, given ℓ the calculation of H is possible but complicated, mainly due to the fact that H can have atoms.

Remark 2.15 (bivariate case). In the bivariate case, where the stable tail dependence function is $\ell(y_1, y_2) = \int_{[0,1]} \max\{wy_1, (1-w)y_2\} \,\mathrm{d}H(w)$, one achieves this by the following formulas from [1, p. 264]:

$$H(\{0\}) = \lim_{y_1\to\infty} \frac{\partial\ell}{\partial y_2}(y_1, y_2),$$
$$H(\{1\}) = \lim_{y_2\to\infty} \frac{\partial\ell}{\partial y_1}(y_1, y_2).$$

A density of H on the interior of the unit interval is

$$h(w) = -\frac{1}{w(1-w)} \frac{\partial^2 \ell}{\partial y_1 \partial y_2}(1-w, w), \quad w \in (0,1).$$

Extremal coefficients

The extremal coefficients θ_V, defined below, provide a first idea of the dependence structure of the MEV distribution G. They are finite dimensional and thus cannot tell the full story. However, the extremal coefficients are easily derived from ℓ and have an intuitive interpretation. Let $V \neq \varnothing$ be a subset of $\{1, \ldots, d\}$ and define $\boldsymbol{e}_V = \sum_{i \in V} \boldsymbol{e}_i$, a sum of canonical basis vectors. The extremal coefficients are the numbers $\theta_V := \ell(\boldsymbol{e}_V)$, where V runs through all possible non-empty subsets of $\{1, \ldots, d\}$. They have an interesting property. Let $p \in (0,1)$ and $\boldsymbol{Y} \sim G$, then θ_V satisfies

$$\begin{aligned} \mathbb{P}(Y_i \leq G_i^{\leftarrow}(p),\ \forall i \in V) &= \exp\left\{ \int_{\boldsymbol{\mathcal{S}}_d} \min_{i \in V}\{w_i \log p\} \, \mathrm{d}H(\boldsymbol{w}) \right\} \\ &= \exp\left\{ \log p \int_{\boldsymbol{\mathcal{S}}_d} \max_{i \in V}\{w_i\} \, \mathrm{d}H(\boldsymbol{w}) \right\} \\ &= p^{\theta_V}, \end{aligned}$$

where the first equality followed from (2.12) and the last from (2.13). In particular for the set $V = \{1, \ldots, d\}$, we have $p^{\theta_V} = C_G(p, \ldots, p)$, where C_G is the unique copula of G. Beirlant et al. point out that "stronger dependence corresponds to smaller extremal coefficients". It holds $1 \leq \theta_V \leq |V|$, where $|V|$ is the cardinality of the set V.

2.2.3 Pickands dependence function A in the bivariate case

For $d = 2$, the **Pickands dependence function** A is defined as the restriction of the stable tail dependence function to the unit simplex:

$$A(t) := \ell(1-t, t), \quad t \in [0,1]. \tag{2.14}$$

Because of property (L1) we get $\ell((y_1+y_2)^{-1}\boldsymbol{y}) = (y_1+y_2)^{-1}\ell(\boldsymbol{y})$. It follows

$$\ell(\boldsymbol{y}) = (y_1+y_2) A\left(\frac{y_2}{y_1+y_2}\right), \quad \boldsymbol{y} \geq \boldsymbol{0}, \tag{2.15}$$

and consequently with (2.10)

$$G(\boldsymbol{x}) = \exp\left\{ \log(G_1(x_1)G_2(x_2)) A\left(\frac{\log G_2(x_2)}{\log(G_1(x_1)G_2(x_2))}\right)\right\}, \quad \boldsymbol{x} \in \mathbb{R}^2.$$

Properties (L3) and (L4) translate into

(A1) $\max(1-t, t) \leq A(t) \leq 1 \quad$ for $t \in [0,1]$,

(A2) A is convex.

Any function $A : [0,1] \to \mathbb{R}$ satisfying (A1) and (A2) is a valid Pickands dependence function (see [8]).

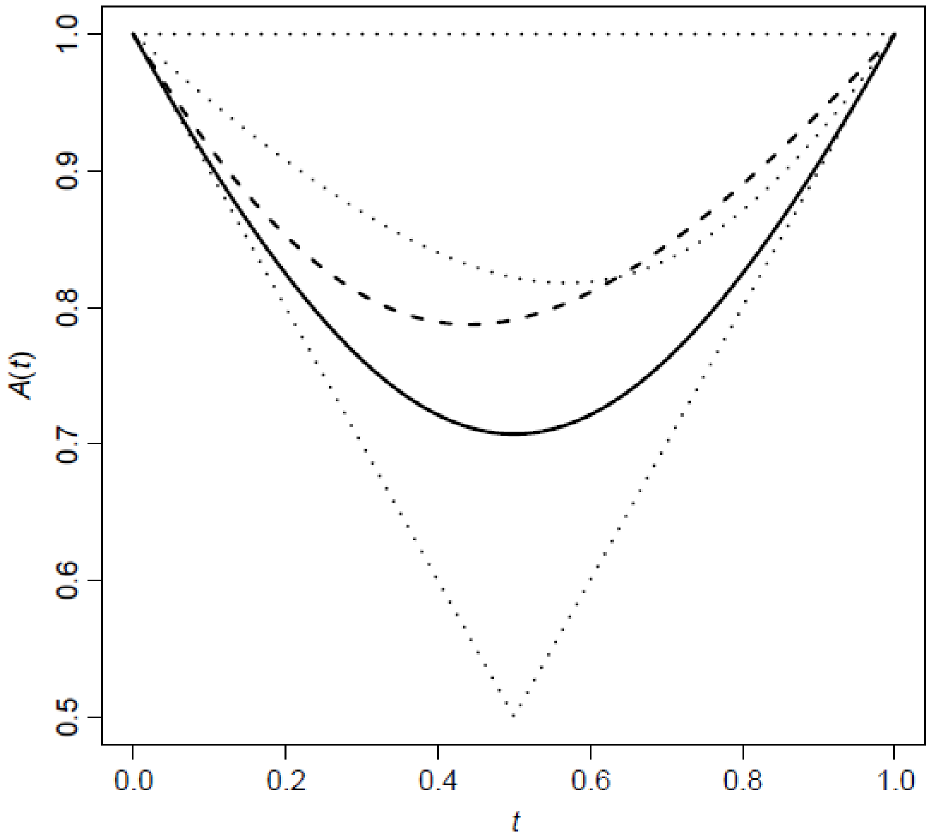

Figure 2.1: *Some A-functions and the triangle region in (A1).* $A(t) = 1$ *corresponds to independence, and* $A(t) = \max(1-t, t)$ *to complete dependence.*

Remark 2.16. Properties (L1)–(L4) are sufficient to characterise a stable tail dependence function if and only if $d = 2$.

In [8] and [9] Gudendorf and Segers compare various nonparametric estimators for A. The main challenge for nonparametric estimation techniques in MEVT in general is to fulfill functional constraints such as (A1) and (A2) (cf. Figure 2.2). The alternative is to consider parametric sub-families (e.g. for A, ℓ, H or G), which enables us to apply maximum-likelihood methods, and the obtained results always define a MEVD. "In this way only a small subset of the complete class of limit distributions for G is obtained, but by careful choice it is possible to ensure that a wide sub-class of the entire limit family is approximated" [3, p. 146].

2.3 A comprehensive example

We have seen that there exists a great variety of equivalent descriptions of multivariate extreme value distributions. Now we study the parametric Gumbel model in detail.

Let $1 < \alpha < \infty$ and G_1, G_2 be arbitrary univariate extreme value distributions. The function $A(t) := ((1-t)^\alpha + t^\alpha)^{1/\alpha},\ t \in [0,1]$, fulfills (A1) and (A2). Therefore it is a Pickands

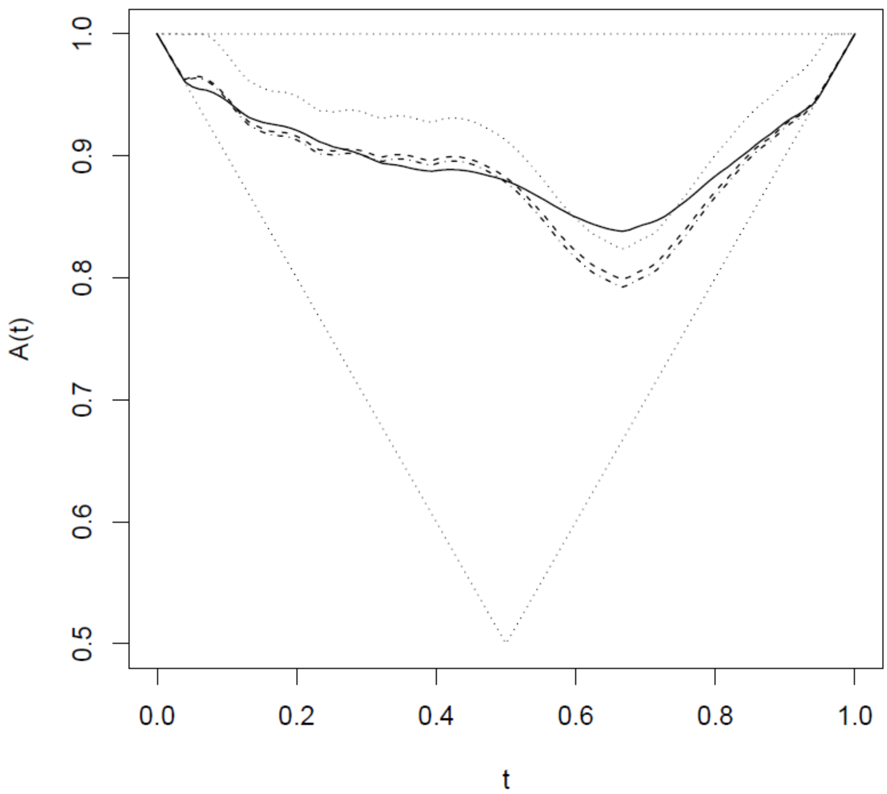

Figure 2.2: *Plots of various nonparametric estimates for a Pickands dependence function, which are discussed in [8]. Estimates by Pickands (dotted line), Deheuvels (dashed line), Hall–Tajvidi (dot-dashed line) and Capéràa–Fougères–Genest (solid line). In general these estimates are, however, not convex which requires additional work for convexifying.*

dependence function and characterises a two-dimensional EVD. From (2.15) we get the corresponding stable tail dependence function ℓ,

$$\ell(y_1, y_2) = (y_1^\alpha + y_2^\alpha)^{\frac{1}{\alpha}}, \quad \boldsymbol{y} \geq \boldsymbol{0}.$$

Given ℓ and the margins G_1 and G_2 one calculates

$$G(x_1, x_2) = \exp\left\{-\left[(-\log G_1(x_1))^\alpha + (-\log G_2(x_2))^\alpha\right]^{\frac{1}{\alpha}}\right\}, \quad \boldsymbol{x} \in \mathbb{R}^2.$$

The copula of G is the Gumbel copula, which explains the model's name. Now we turn to the extremal coefficients. Trivially $\theta_{\{1\}} = \theta_{\{2\}} = 1$ and $\theta_{\{1,2\}} = \ell(1,1) = 2^{1/\alpha}$. A large α implies strong extremal dependence. This is due to the fact that

$$G(G_1^{\leftarrow}(p), G_2^{\leftarrow}(p)) = p^{\theta_{\{1,2\}}}, \quad p \in (0,1)$$

increases for increasing values of α. Finally, we use Remark 2.15 to obtain an expression for the spectral measure H on $[0,1]$. We get

$$H(\{0\}) = H(\{1\}) = 0.$$

On $(0,1)$ a density of H is given by

$$\begin{aligned} h(w) &= -\frac{1}{w(1-w)} \frac{\partial^2 \ell}{\partial y_1 \partial y_2}(1-w, w) \\ &= (\alpha - 1)(w(1-w))^{\alpha-2} \left((1-w)^\alpha + w^\alpha\right)^{\frac{1}{\alpha}-2}, \quad w \in (0,1). \end{aligned}$$

Chapter 3

Tail Dependence Coefficients

Tail dependence coefficients are scalar measures of dependence in the tails of a bivariate distribution. They are defined via conditional probabilities of high quantile exceedances and can be expressed in terms of copulas. Thus, tail dependence coefficients do not depend on the margins.

In this chapter $\boldsymbol{X} = (X_1, X_2)$ is a two-dimensional random vector with df F, copula C and margins F_1 and F_2.

3.1 Tail dependence coefficient of the first kind

Definition 3.1 (upper tail dependence coefficient). [4, Definition 2]
The **(upper) tail dependence coefficient** of X_1 and X_2 is

$$\lambda_u := \lambda_u(X_1, X_2) := \lim_{q \to 1^-} \mathbb{P}(X_2 > F_2^{\leftarrow}(q) | X_1 > F_1^{\leftarrow}(q)), \tag{3.1}$$

provided a limit $\lambda_u \in [0, 1]$ exists. If $\lambda_u \in (0, 1]$, then X_1 and X_2 are said to be **asymptotically dependent** (or **tail dependent**) in the upper tail. If $\lambda_u = 0$, they are **asymptotically independent**.

Analogously, the lower tail dependence coefficient is

$$\lambda_l := \lim_{q \to 0^+} \mathbb{P}(X_2 \leq F_2^{\leftarrow}(q) | X_1 \leq F_1^{\leftarrow}(q)), \tag{3.2}$$

if a limit $\lambda_l \in [0, 1]$ exists.

Remark 3.2. If X_1 and X_2 are independent, $\lambda_u = 0$. Whereas in case of complete dependence $\lambda_u = 1$.

The upper tail dependence coefficient λ_u can be written in terms of the survival copula $\widehat{C}$ of $\boldsymbol{X}$ (see Definition 1.8):

$$\lambda_u = \lim_{q \to 1^-} \frac{1 - 2q + C(q, q)}{1 - q} = \lim_{q \to 0^+} \frac{\widehat{C}(q, q)}{q}. \tag{3.3}$$

Hence, λ_u is independent of the margins F_1 and F_2.

Furthermore, using the middle expression in (3.3), one obtains for a differentiable copula C that

$$\lambda_u = 2 - \lim_{q\to 1^-} \frac{1 - C(q,q)}{1-q} = 2 - \lim_{q\to 1^-} \frac{\mathrm{d}}{\mathrm{d}q} C(q,q). \tag{3.4}$$

Example 3.3. The bivariate Gumbel copula

$$C_\alpha(u_1, u_2) = \exp\left\{ - \left[(-\log u_1)^\alpha + (-\log u_2)^\alpha\right]^{\frac{1}{\alpha}} \right\}, \quad 1 \le \alpha < \infty,$$

has $\lambda_u = 2 - 2^{1/\alpha}$. ◦

Next we are going to derive the upper tail dependence coefficient of the non-explicit Gauss copula. Some notation: $\boldsymbol{U} = (U_1, U_2) \sim C$ for a two-dimensional copula C.

Lemma 3.4. Suppose C is a two-dimensional and differentiable copula. Then

$$\mathbb{P}(U_2 \le u_2 | U_1 = u_1) = \frac{\partial}{\partial u_1} C(u_1, u_2)$$

for all $(u_1, u_2) \in [0,1]^2$.

Proof. C is increasing and continuous in both arguments. For $\delta > 0$ such that $u_1 + \delta \le 1$ we find

$$\frac{C(u_1+\delta, u_2) - C(u_1, u_2)}{\delta} = \frac{\mathbb{P}(u_1 \le U_1 \le u_1 + \delta, U_2 \le u_2)}{\delta} = \mathbb{P}(U_2 \le u_2 | u_1 \le U_1 \le u_1 + \delta)$$

and for $\delta < 0$ such that $u_1 + \delta \ge 0$ we find

$$\frac{C(u_1+\delta, u_2) - C(u_1, u_2)}{\delta} = \frac{-\mathbb{P}(u_1 + \delta \le U_1 \le u_1, U_2 \le u_2)}{\delta} = \mathbb{P}(U_2 \le u_2 | u_1 + \delta \le U_1 \le u_1).$$

Taking the limit $\delta \to 0^+$ yields the desired result. □

Due to Lemma 3.4 and the identity

$$\frac{\mathrm{d}}{\mathrm{d}q} C(q,q) = \frac{\partial}{\partial u_1} C(q,q) + \frac{\partial}{\partial u_2} C(q,q) \tag{3.5}$$

equation (3.4) simplifies to

$$\lambda_u = \lim_{q\to 1^-} \left(\mathbb{P}(U_2 > q | U_1 = q) + \mathbb{P}(U_1 > q | U_2 = q) \right). \tag{3.6}$$

Example 3.5 (Gauss copula). Let $\boldsymbol{X} = (X_1, X_2)$ have a bivariate normal distribution with correlation ρ. The components $X_i \sim N(\mu_i, \sigma_i^2)$ so that $F_i(x) = \Phi\left(\frac{x-\mu_i}{\sigma_i}\right)$, $i = 1, 2$. The copula C_ρ^{Ga} of $\boldsymbol{X}$ is unique. We evaluate its upper tail dependence coefficient λ_u. If $\rho = -1$, then $\lambda_u = 0$. If $\rho = 1$, then $\lambda_u = 1$. Therefore assume $\rho \in (-1, 1)$ and let $(U_1, U_2) \sim C_\rho^{\mathrm{Ga}}$. For the first summand in (3.6) one finds

$$\begin{aligned}\mathbb{P}(U_2 > q | U_1 = q) &= \mathbb{P}(F_2^{-1}(U_2) > F_2^{-1}(q) | F_1^{-1}(U_1) = F_1^{-1}(q)) \\ &= \mathbb{P}(X_2 > \mu_2 + \sigma_2 \Phi^{-1}(q) | X_1 = \mu_1 + \sigma_1 \Phi^{-1}(q))\end{aligned}$$

Because of $\frac{X_2 - \mu_2}{\sigma_2} | (X_1 = \mu_1 + \sigma_1 x) \sim N(\rho x, 1 - \rho^2)$ it follows that

$$\lim_{q \to 1^-} \mathbb{P}(U_2 > q | U_1 = q) = \lim_{x \to \infty} \mathbb{P}\left(\frac{X_2 - \mu_2}{\sigma_2} > x | X_1 = \mu_1 + \sigma_1 x \right)$$
$$= \lim_{x \to \infty} \bar{\Phi}\left(x \frac{\sqrt{1-\rho}}{\sqrt{1+\rho}} \right) = 0$$

and since the same applies to the second summand in (3.6), we get that $\lambda_u = 0$ for all $\rho < 1$. Extreme events for X_1 and X_2 happen independently unless the correlation is 1. ◦

Remark 3.6. In contrast to the Gauss copula, the t-copula shows positive tail dependence. For details consult [12, p. 211].

The bivariate normal distribution illustrates that even strong correlation $\rho < 1$ does not necessarily imply asymptotic dependence. In general, there exists a variety of copulas (e.g. Gauss copula) which do not have positive upper tail dependence, but nevertheless feature some kind of dependence between $F_1(X_1)$ and $F_2(X_2)$ in their tails.

This means that the range of possibilities in the asymptotic independence case is extremely wide. For instance if U_1, U_2 are independent standard uniform random variables, both pairs (U_1, U_2) and $(U_1, 1 - U_1)$ are asymptotically independent. As a consequence, a need to complement the information gained by λ_u with additional dependence measures arises.

3.2 Tail dependence coefficient of the second kind

Definition 3.7 (tail dependence coefficient of the second kind). [3, p. 164]
If the following limit exists, the **tail dependence coefficient of the second kind** is

$$\bar{\lambda}_u := \bar{\lambda}_u(X_1, X_2) := \lim_{q \to 1^-} \frac{2 \log(1-q)}{\log \mathbb{P}(F_1(X_1) > q, F_2(X_2) > q)} - 1. \tag{3.7}$$

Remark 3.8. For ease of notation we may omit the index and write $\lambda, \bar{\lambda}$ instead of $\lambda_u, \bar{\lambda}_u$. The upper tail dependence coefficient will be referred to as tail dependence coefficient.

Coles [3, p. 164] states the following properties of $\bar{\lambda}$:

1. $-1 \leq \bar{\lambda} \leq 1$.
2. In case of asymptotic dependence $\bar{\lambda} = 1$.
3. For independent random variables $\bar{\lambda} = 0$.
4. For asymptotically independent random variables $\bar{\lambda}$ increases with the strength of dependence at extreme levels.

It makes a lot of sense to consider the pair $(\lambda, \bar{\lambda})$. Three cases for the pair $(\lambda, \bar{\lambda})$ can be distinguished.

- $\lambda = 0,\ \bar{\lambda} \in [-1, 1)$
  Here $\bar{\lambda}$ provides additional information about the dependence, although we have asymptotic independence.
- $\lambda = (0, 1],\ \bar{\lambda} = 1$
  Under asymptotic dependence λ measures its strength.
- $\lambda = 0,\ \bar{\lambda} = 1$
  This is a pathological case. By mistake, Coles (2001) claims that it does not exist, while Demarta (2007) even provides an example [4, Example 2.9].

Example 3.9 (Gauss copula C_ρ^{Ga}). The tail dependence coefficient of the second kind $\bar{\lambda}$ exists and equals the correlation ρ. Hence $(\lambda, \bar{\lambda}) = (0, \rho)$ for $\rho < 1$ and $(\lambda, \bar{\lambda}) = (1, 1)$ for $\rho = 1$. ∘

For the Gumbel copula C_α the coefficients are $(\lambda, \bar{\lambda}) = (2 - 2^{1/\alpha}, 1)$.

3.3 The tail dependence coefficient in context

We now consider the tail dependence coefficient λ of MEV distributions. Let G be a bivariate EVD and C its copula. Using the middle expression in (3.4), $1 - q \approx \log q$ for q close to 1 and $C(q, q) = q^{\ell(1,1)}$, one finds

$$\lambda = 2 - \lim_{q \to 1^-} \frac{\log C(q, q)}{\log q} = 2 - \ell(1, 1), \tag{3.8}$$

where ℓ is the stable tail dependence function of G. In terms of the Pickands dependence function: $\lambda = 2\left(1 - A\left(\frac{1}{2}\right)\right)$.

Remark 3.10. Note that if $\lambda = 0$, then $A\left(\frac{1}{2}\right) = 1$ and by convexity $A(t) = 1$ (see Figure 3.1), which implies C being the independence copula.

One calculates the lower tail dependence coefficient in (3.2):

$$\lambda_l = \lim_{q \to 0^+} \frac{C(q, q)}{q} = \lim_{q \to 0^+} \frac{q^{2A(1/2)}}{q} = \begin{cases} 0 & \text{if } A\left(\frac{1}{2}\right) > \frac{1}{2}\,, \\ 1 & \text{if } A\left(\frac{1}{2}\right) = \frac{1}{2}\,. \end{cases}$$

Apart from perfect dependence, copulas of MEV distributions are asymptotically independent in the lower tails.

Finally, we (again) investigate the bivariate $\widetilde{G}$ (i.e. a bivariate EVD on $(0, \infty)^2$ with standard Fréchet margins). Then, due to (2.10), property (L1) of the stable tail dependence function ℓ, and (3.8),

$$\widetilde{G}(x, x) = \exp\{-\ell(x^{-1}, x^{-1})\} = \exp\{-x^{-1}\ell(1, 1)\} = \exp\{-x^{-1}(2 - \lambda)\}, \quad \forall x > 0.$$

Hence, λ characterises the df $\widetilde{G}$ on its diagonal.

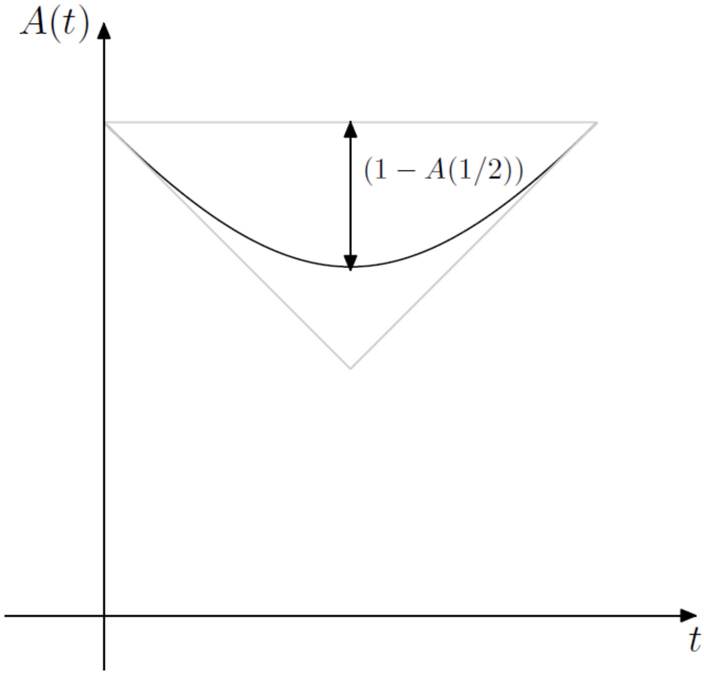

Figure 3.1: *The tail dependence coefficient is connected to the A-function via* $\lambda = 2(1 - A(1/2))$, *source: [8].*

Chapter 4

Extreme Value Copulas

Basically, multivariate extreme value theory deals with the dependence structure of extreme events. The most natural way of studying the inference between components of a random vector, without worrying about its marginal distributions, is to consider its copula.

Our aim is to derive limiting copulas for the normalised vector $\boldsymbol{M}_n$ of componentwise block maxima. We use the notation from Chapter 2:

$$\boldsymbol{M}_n = \left(\max_{1 \le i \le n} X_{1,i}\ , \ldots, \max_{1 \le i \le n} X_{d,i} \right)^{\top}, \quad n \in \mathbb{N},$$

where $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots$ is a sequence of iid d-dimensional random vectors with joint distribution function (df) F, copula C and margins $F_1, \ldots, F_d$.

4.1 Multivariate extension of the Fisher–Tippett theorem

Lemma 4.1. For every $n \in \mathbb{N}$, the copula C_n of componentwise maxima $\boldsymbol{M}_n$ is given by

$$C_n(\boldsymbol{u}) = C^n(\boldsymbol{u}^{1/n}), \quad \boldsymbol{u} \in [0,1]^d.$$

Proof. For $\boldsymbol{u} \in [0,1]^d$ and $n \in \mathbb{N}$

$$\begin{aligned} C_n(\boldsymbol{u}) &= F^n((F_1^n)^{\leftarrow}(u_1), \ldots, (F_d^n)^{\leftarrow}(u_d)) \\ &= F^n((F_1)^{\leftarrow}(u_1^{1/n}), \ldots, (F_d)^{\leftarrow}(u_d^{1/n})) \\ &= C^n(\boldsymbol{u}^{1/n}). \end{aligned}$$

□

Lemma 4.1 motivates to look at the limit $\lim_{n\to\infty} C_n$.

Definition 4.2 (EV copula, CDA). A copula C_0 is called an **extreme value copula (EV copula)** if there exists a copula C such that

$$\lim_{n\to\infty} C^n(\boldsymbol{u}^{1/n}) = C_0(\boldsymbol{u}), \quad \forall \boldsymbol{u} \in [0,1]^d. \tag{4.1}$$

Then C is said to be in the **copula domain of attraction** (CDA) of C_0. We use the notation $C \in \mathrm{CDA}(C_0)$.

In fact, the pointwise convergence in 4.1 holds even uniformly in $\boldsymbol{u} \in [0,1]^d$. To see this, consider a d-dimensional copula C which is in the CDA of some EV copula C_0. We analyse the sequence of copulas

$$C_n(\boldsymbol{u}) = C^n(\boldsymbol{u}^{1/n}), \quad n \in \mathbb{N}, \boldsymbol{u} \in [0,1]^d.$$

By definition $(C_n)_{n\in\mathbb{N}}$ converges pointwise to C_0. From Proposition 1.6 it follows that the C_n, $n \in \mathbb{N}$, are Lipschitz continuous with Lipschitz constant 1. Let $\epsilon > 0$. Since

$$|C_n(\boldsymbol{u}) - C_n(\boldsymbol{v})| < \epsilon, \quad \forall n \in \mathbb{N} \text{ and } \boldsymbol{u}, \boldsymbol{v} \in [0,1]^d, \text{ whenever } \sum_{i=1}^{d} |u_i - v_i| < \epsilon, \tag{4.2}$$

the family $\{C_n : n \in \mathbb{N}\}$ is uniformly equicontinuous. Note that (4.2) is the definition of uniform equicontinuity. Moreover, the family $\{C_n : n \in \mathbb{N}\}$ is uniformly bounded by 1. In view of the Arzelà–Ascoli theorem, a pointwise convergent and uniformly equicontinuous sequence of functions on the compact set $[0,1]^d$ converges uniformly.

Remark 4.3. Even if we do not assume pointwise convergence to a continuous function C_0, the Arzelà–Ascoli theorem still guarantees the existence of a subsequence $(C_{n(k)})_{k\in\mathbb{N}}$ that converges uniformly to a continuous function on $[0,1]^d$. This proves that if the limit $\lim_{n\to\infty} C^n(\boldsymbol{u}^{1/n})$ exists for all $\boldsymbol{u}$, it has to be continuous.

Next we consider the following conditions for a sequence of functions $f_n, n \in \mathbb{N}$, and a function $f : [0,1]^d \to [0,1]$.

1. The sequence $f_n : [0,1]^d \to [0,1], n \in \mathbb{N}$, converges pointwise to f.
2. f_n is continuous for all $n \in \mathbb{N}$.
3. The sequence $f_n, n \in \mathbb{N}$, is uniformly bounded.
4. f is continuous.

Under these circumstances it can be shown that the sequence (f_n) has a property called **continuous convergence**: For $\boldsymbol{x} \in [0,1]^d$ and every sequence $(\boldsymbol{x}_n)$ in $[0,1]^d$ such that $\boldsymbol{x}_n \to \boldsymbol{x}$, it follows

$$\lim_{n\to\infty} f_n(\boldsymbol{x}_n) = f(\boldsymbol{x}). \tag{4.3}$$

In particular, these four conditions hold for the sequence of copulas (C_n) and the EV copula C_0, where $C \in \text{CDA}(C_0)$. We rewrite (4.3) for this special case:

$$\lim_{n\to\infty} C^n(\boldsymbol{x}_n^{1/n}) = C_0(\boldsymbol{x}), \quad \text{if } \boldsymbol{x}_n \to \boldsymbol{x}. \tag{4.4}$$

Corollary 4.4. Let the d-dimensional df F be in the MDA of the d-dimensional EVD G. If a copula C of F is in CDA(C_0), then G also has copula C_0.

Proof. Let $\boldsymbol{a}_n > \boldsymbol{0}, \boldsymbol{b}_n \in \mathbb{R}^d, n \in \mathbb{N}$, be normalising sequences for F. Fix an $\boldsymbol{x} \in \mathbb{R}^d$. By the convergence of margins $F_i^n(a_{i,n}x_i + b_{i,n}) \to G_i(x_i)$ for $n \to \infty$ and $i = 1, \ldots, d$ the sequence

$$\boldsymbol{y}_n := (F_1^n(a_{1,n}x_1 + b_{1,n}), \ldots, F_d^n(a_{d,n}x_d + b_{d,n})), \quad n \in \mathbb{N}, \tag{4.5}$$

converges to $(G_1(x_1), \dots, G_d(x_d))$. Therefore we have

$$\begin{aligned} F^n(\boldsymbol{a}_n\boldsymbol{x} + \boldsymbol{b}_n) &= C^n(F_1(a_{1,n}x_1 + b_{1,n}), \dots, F_d(a_{d,n}x_d + b_{d,n})) \\ &= C^n(\boldsymbol{y}_n^{1/n}) \\ &\overset{n\to\infty}{\longrightarrow} C_0(G_1(x_1), \dots, G_d(x_d)) \\ &= G(\boldsymbol{x}), \quad \boldsymbol{x} \in \mathbb{R}^d, \end{aligned}$$

where the first equality followed by Lemma 4.1 and the limit from (4.4). Thus G has the copula C_0. □

Definition 4.5 (copula-max-stability). A d-dimensional copula C is **copula-max-stable** if it satisfies

$$C(\boldsymbol{u}) = C^m(\boldsymbol{u}^{1/m}), \quad \forall \boldsymbol{u} \in [0,1]^d, \tag{4.6}$$

for all $m \in \mathbb{N}$.

Remark 4.6. Copula-max-stability is completely different from the max-stability in Definition 2.11. In particular, every copula has standard uniform margins and thus can never be max-stable in the sense of Definition 2.11. However, in order to be consistent with the notation used in literature (e.g. [4, 8, 12, 13]) we call copulas with property (4.6) max-stable.

Proposition 4.7. *A copula is an EV copula if and only if it is max-stable.*

Proof. Starting with an EV copula C_0 choose $C \in \mathrm{CDA}(C_0)$ and $m \in \mathbb{N}$. Then

$$C_0(\boldsymbol{u}) = \lim_{n\to\infty} C^n(\boldsymbol{u}^{1/n}) = \lim_{n\to\infty} C^{nm}(\boldsymbol{u}^{1/nm}) = C_0^m(\boldsymbol{u}^{1/m}), \quad \boldsymbol{u} \in [0,1]^d.$$

Conversely, every max-stable copula C is an EV copula because

$$\lim_{n\to\infty} C^n(\boldsymbol{u}^{1/n}) = C(\boldsymbol{u}), \quad \boldsymbol{u} \in [0,1]^d.$$

□

This also shows that an EV copula is in its own CDA. As seen above, the convergence in (4.1) holds uniformly in $\boldsymbol{u}$. Hence, we can replace $n \in \mathbb{N}$ by a real $t > 0$ and work with the relationship

$$\lim_{t\to\infty} C^t(\boldsymbol{u}^{1/t}) = C_0(\boldsymbol{u}), \quad \boldsymbol{u} \in [0,1]^d \tag{4.7}$$

instead. As shown in [12, p. 315], alternative formulations of (4.7) are

$$\lim_{s\to 0^+} \frac{1 - C(\boldsymbol{u}^s)}{s} = -\log C_0(\boldsymbol{u}), \quad \boldsymbol{u} \in [0,1]^d \tag{4.8}$$

and

$$\lim_{s\to 0^+} \frac{1 - C(1 - sx_1, \dots, 1 - sx_d)}{s} = -\log C_0(\mathrm{e}^{-x_1}, \dots, \mathrm{e}^{-x_d}), \quad \boldsymbol{x} \geq \boldsymbol{0}. \tag{4.9}$$

Example 4.8. In Example 2.4 we calculated the limiting bivariate EVDs for the dfs F_a, F_b and F_c with standard Fréchet margins. Now we determine their EV copulas.

(a) $C_{F_a}(u_1, u_2) = u_1 u_2$ ofr $\boldsymbol{u} \in [0,1]^2$, where C_{F_a} denotes the copula of F_a, is an EV copula.

(b) C_{F_b} is the Gumbel copula, which is another EV copula.

(c) The underlying copula C_{F_c} is the Morgenstern copula C_δ^M with $\delta = 1$:

$$C_\delta^M(u_1, u_2) = u_1 u_2[1 + \delta(1-u_1)(1-u_2)], \quad \boldsymbol{u} \in [0,1]^2, \delta \in [-1,1].$$

Using the alternative formulation (4.9) we find for all $\delta \in [-1,1]$ that

$$\lim_{s \to 0^+} \frac{1 - C_\delta^M(1 - sx_1, \ldots, 1 - sx_d)}{s} = x_1 + x_2 = -\log(\mathrm{e}^{-(x_1+x_2)})$$

and therefore $C_0(u_1, u_2) = u_1 u_2$ for $\boldsymbol{u} \in [0,1]^2$ and $\delta \in [-1,1]$. This shows C_{F_c} is in the CDA of the independence copula.

○

The following theorem elegantly summarises everything one needs to know about MEVT.

Theorem 4.9 (Multivariate extension of the Fisher–Tippett theorem). [4, Theorem 3.1] Let $F(\boldsymbol{x}) = C(F_1(x_1), \ldots, F_d(x_d))$ for the marginal dfs $F_1, \ldots, F_d$ and some copula C, and let $G(\boldsymbol{x}) = C_0(G_1(x_1), \ldots, G_d(x_d))$ be a MEVD with EV copula C_0, where the marginal dfs $G_1, \ldots, G_d$ of G are univariate EVDs.
Then $F \in \mathrm{MDA}(G)$ if and only if

(i) $F_i \in \mathrm{MDA}(G_i)$ for every $i \in \{1, \ldots, d\}$ and

(ii) $\lim_{n\to\infty} C^n(\boldsymbol{u}^{1/n}) = C_0(\boldsymbol{u})$ for all $\boldsymbol{u} \in [0,1]^d$.

Proof. We provide a different and more detailed proof of this important theorem. Given $F \in \mathrm{MDA}(G)$ and normalising sequences $\boldsymbol{a}_n > \boldsymbol{0}, \boldsymbol{b}_n \in \mathbb{R}^d, n \in \mathbb{N}$, *(i)* follows from the univariate Fisher–Tippett theorem.

The difficult part is to show *(ii)*. The idea is to use the convergence of margins $F_i^n(a_{i,n}x_i + b_{i,n}) \to G_i(x_i)$ for $n \to \infty$ and $i = 1, \ldots, d$ and the Lipschitz continuity of the copula $C_n(\boldsymbol{u}) = C^n(\boldsymbol{u}^{1/n})$.

By definition $F \in \mathrm{MDA}(G)$ is equivalent to $\lim_{n\to\infty} F^n(\boldsymbol{a}_n\boldsymbol{x} + \boldsymbol{b}_n) = G(\boldsymbol{x}), \quad \boldsymbol{x} \in \mathbb{R}^d$, which is the same as

$$\lim_{n\to\infty} C_n(F_1^n(a_{1,n}x_1 + b_{1,n}), \ldots, F_d^n(a_{d,n}x_d + b_{d,n})) = C_0(G_1(x_1), \ldots, G_d(x_d)). \tag{4.10}$$

Fix $\boldsymbol{u} \in [0,1]^d$. Since the univariate EVDs $G_1, \ldots, G_d$ are continuous, we can find x_i such that $u_i = G_i(x_i),\ i = 1, \ldots, d$. By the triangle inequality we get

$$\begin{aligned}|C_n(u_1, \ldots, u_d) - C_0(u_1, \ldots, u_d)| &= |C_n(G_1(x_1), \ldots, G_d(x_d)) - C_0(G_1(x_1), \ldots, G_d(x_d))| \\ &\le |C_n(G_1(x_1), \ldots, G_d(x_d)) - C_n(F_1^n(a_{1,n}x_1 + b_{1,n}), \ldots, F_d^n(a_{d,n}x_d + b_{d,n}))| \\ &\quad + |C_n(F_1^n(a_{1,n}x_1 + b_{1,n}), \ldots, F_d^n(a_{d,n}x_d + b_{d,n})) - C_0(G_1(x_1), \ldots, G_d(x_d))|.\end{aligned}$$

Now the second term goes to zero because of (4.10), i.e.

$$\lim_{n\to\infty} |C_n(F_1^n(a_{1,n}x_1 + b_{1,n}), \dots, F_d^n(a_{d,n}x_d + b_{d,n})) - C_0(G_1(x_1), \dots, G_d(x_d))| = 0.$$

Using the Lipschitz continuity of copulas (see Proposition 1.6), the first term is bounded by a sum of terms, which converge to zero as well because of marginal convergence. In long form

$$\begin{aligned}|C_n(G_1(x_1), \dots, G_d(x_d)) - C_n(F_1^n(a_{1,n}x_1 + b_{1,n}), \dots, F_d^n(a_{d,n}x_d + b_{d,n}))|\\ \le \sum_{i=1}^{d} |G_i(x_i) - F_i^n(a_{i,n}x_i + b_{i,n})| \to 0, \quad \text{for } n \to \infty.\end{aligned}$$

Therefore, $\lim_{n\to\infty} C_n(u_1, \dots, u_d) = C_0(u_1, \dots, u_d)$ which establishes *(ii)*. Finally, given *(i)* and *(ii)*, one copies the proof of Corollary 4.4 to see $F \in \mathrm{MDA}(G)$. □

Remark 4.10. Both *(i)* and *(ii)* are relatively easy to check: *(i)* by means of univariate EVT; *(ii)* by straightforward calculation.

Example 4.11. We study the two-dimensional df

$$F(x_1, x_2) = 1 - \mathrm{e}^{-x_1} - \mathrm{e}^{-x_2} + (\mathrm{e}^{x_1} + \mathrm{e}^{x_2} - 1)^{-1}, \quad \boldsymbol{x} \ge \boldsymbol{0},$$

given in [14, p. 279]. The marginal dfs are standard exponential:

$$F_i(x) = 1 - \mathrm{e}^{-x} \quad \text{for } x > 0 \text{ and } i \in \{1, 2\}.$$

In order to find the limiting MEVD, one first analyses condition *(i)* in Theorem 4.9. By evaluating

$$\begin{aligned}F_1^n(x + \log n) &= \left(1 - \mathrm{e}^{-x-\log n}\right)^n \mathbb{1}_{(0,\infty)}(x + \log n)\\ &= \left(1 - \frac{\mathrm{e}^{-x}}{n}\right)^n \mathbb{1}_{(-\log n,\infty)}(x)\\ &\to \exp\left\{-\mathrm{e}^{-x}\right\} =: \Lambda(x), \quad x \in \mathbb{R},\end{aligned}$$

one concludes that the margins are in the MDA of the Gumbel distribution Λ. Therefore we now turn our attention to condition *(ii)* and calculate the copula C of F. We get that

$$\begin{aligned}C(u_1, u_2) &= F(F_1^{\leftarrow}(u_1), F_2^{\leftarrow}(u_2))\\ &= F(-\log(1 - u_1), -\log(1 - u_2))\\ &= 1 - (1 - u_1) - (1 - u_2) + \left(\frac{1}{1 - u_1} + \frac{1}{1 - u_2} - 1\right)^{-1}\\ &= \frac{u_1 u_2(2 - u_1 - u_2)}{1 - u_1 u_2}, \quad \boldsymbol{u} \in [0,1]^2 \backslash \{(1,1)\},\end{aligned}$$

and $C(1, 1) = 1$.

In the light of the alternative formulation (4.9) of condition *(ii)* we write

$$\frac{1-C(1-sx_1,\dots,1-sx_d)}{s}=\frac{(x_1+x_2-sx_1x_2)-(x_1+x_2)(1-sx_1-sx_2+s^2x_1x_2)}{s(x_1+x_2-sx_1x_2)} \tag{4.11}$$

for $\boldsymbol{x}\geq\mathbf{0}$ and appropriate (depending on $\boldsymbol{x}$) values of $s>0$ such that $1-s\boldsymbol{x}\in[0,1]^2$. When taking the limit in (4.11) for $s\to0^+$, it turns out that both the nominator and the denominator tend to 0. An application of l'Hôspital's rule yields

$$\begin{aligned}\lim_{s\to0^+}\frac{1-C(1-sx_1,\dots,1-sx_d)}{s}&=\lim_{s\to0^+}\left(x_1+x_2-\frac{x_1x_2}{x_1+x_2-2sx_1x_2}\right)\\&=x_1+x_2-\frac{x_1x_2}{x_1+x_2},\quad\boldsymbol{x}\geq\mathbf{0}.\end{aligned} \tag{4.12}$$

Next it needs to be checked that this last expression in (4.12) really defines an EV copula C_0 via

$$-\log C_0(\mathrm{e}^{-x_1},\mathrm{e}^{-x_2})=x_1+x_2-\frac{x_1x_2}{x_1+x_2}. \tag{4.13}$$

It turns out that

$$C_0(u_1,u_2)=\exp\left\{\log u_1u_2-\frac{\log u_1\log u_2}{\log u_1u_2}\right\}=\exp\left\{\log u_1u_2\left[1-\frac{\log u_1\log u_2}{(\log u_1u_2)^2}\right]\right\}.$$

A comparison with the Pickands representation of an EV copula in Corollary 4.13 shows that the corresponding A-function would be

$$A\left(\frac{\log u_1}{\log(u_1u_2)}\right)=1-\frac{\log u_1\log u_2}{(\log u_1u_2)^2}$$

and thus

$$A(t)=t^2-t+1,\quad t\in[0,1],$$

which is a proper Pickands dependence function since it satisfies properties (A1) and (A2). Consequently, C_0 is an EV copula. To sum up $F\in\mathrm{MDA}(G)$ where G has copula C_0 and Gumbel margins. ○

4.2 Representations of extreme value copulas

Chapter 2 covered different representations of a MEVD G:

- the stable tail dependence function ℓ,
- the spectral measure H,
- the Pickands dependence function A.

All these approaches could be used to construct MEV distributions. It is very logical to employ the same techniques to characterise the corresponding EV copulas.

Let $G(\boldsymbol{x})=C_0(G_1(x_1),\dots,G_d(x_d))$ be a MEVD with margins $G_1,\dots,G_d$ and copula C_0, and let ℓ be its stable tail dependence function. The EV copula C_0 can be written in terms of ℓ. From equation (2.10) it follows

$$C_0(\boldsymbol{u}) = \exp\left(-\ell(-\log u_1, \ldots, -\log u_d)\right), \quad \boldsymbol{u} \in [0,1]^d, \tag{4.14}$$

where by virtue of (2.13) ℓ can be expressed through the spectral measure:

$$\ell(\boldsymbol{y}) = \int_{\boldsymbol{\mathcal{S}}_d} \max_{1\leq i\leq n} \{w_i y_i\} \, \mathrm{d}H(\boldsymbol{w}), \quad \boldsymbol{y} \geq \boldsymbol{0}, \tag{4.15}$$

with the constraints

$$\int_{\boldsymbol{\mathcal{S}}_d} w_i \, \mathrm{d}H(\boldsymbol{w}) = 1, \quad i = 1, \ldots, d. \tag{4.16}$$

This representation of an EV copula also explains the somewhat mysterious constraints in (4.16). Because of

$$\begin{aligned} u_i = C_0(1, \ldots, 1, u_i, 1, \ldots, 1) &= \exp\left(-\ell(0, \ldots, 0, -\log u_i, 0, \ldots, 0)\right) \\ &= \exp\left(\log u_i \int_{\boldsymbol{\mathcal{S}}_d} w_i \, \mathrm{d}H(\boldsymbol{w})\right), \quad u_i \in [0,1] \end{aligned}$$

the constraints stem from the standard uniform margin property of copulas.

Remark 4.12. The EV copula $\widetilde{C}_0$ of the MEVD $\widetilde{G}$ (see (2.2)) is written as

$$\widetilde{C}_0(\boldsymbol{u}) = \widetilde{G}\left(\frac{1}{-\log u_1}, \ldots, \frac{1}{-\log u_d}\right), \quad \boldsymbol{u} \in [0,1]^d,$$

with $\log 0 := -\infty$.

Corollary 4.13 (Pickands representation of a two-dimensional EV copula). [8]
Every two-dimensional EV copula C_0 with Pickands dependence function A has the representation

$$C_0(u_1, u_2) = \exp\left\{\log(u_1 u_2) A\left(\frac{\log u_1}{\log(u_1 u_2)}\right)\right\}, \quad (u_1, u_2) \in (0,1)^2. \tag{4.17}$$

Once more we stress that C_0 is the independence copula if $A = 1$. In this case the spectral measure H concentrates on the points $(1,0)$ and $(0,1)$ with $H(\{(1,0)\}) = H(\{(0,1)\}) = 1$.

Remark 4.14. Now since for $\boldsymbol{u} \in (0,1)^2$

$$\exp\left\{\log(u_1 u_2) A_1\left(\frac{\log u_1}{\log(u_1 u_2)}\right)\right\} \geq \exp\left\{\log(u_1 u_2) A_2\left(\frac{\log u_1}{\log(u_1 u_2)}\right)\right\}, \tag{4.18}$$

if A_1 and A_2 are two Pickands dependence functions with $A_1(t) \leq A_2(t)$ for all $t \in [0,1]$, one proves the positive quadrant dependence of EV copulas (i.e $C_0(u_1, u_2) \geq u_1 u_2$). This follows instantly because for all Pickands dependence functions A it holds $A(t) \leq 1$ and 1 is the Pickands dependence function of the independence copula.

Setting $A(t) = \min\{t, 1-t\}$ leads to $C_0(u_1, u_2) = \min\{u_1, u_2\}$, the comonotonicity copula.

4.3 Tail dependence coefficients for extreme value copulas

Suppose C is a d-dimensional copula and $(U_1, \dots, U_d) \sim C$. Denote by C_{ij} the copula of (U_i, U_j), $1 \leq i < j \leq d$. Then C_{ij} is a two-dimensional copula with tail dependence coefficient

$$\lambda_{ij} = 2 - \lim_{q \to 1^-} \frac{\log C_{ij}(q, q)}{\log q}.$$

Moreover, if C_0 is a d-dimensional EV copula, then C_{0ij} is a two-dimensional EV copula.

The following theorem shows the importance of tail dependence coefficients for EV copulas. We use the notation λ_{0ij} for the tail dependence coefficient of C_{0ij}.

Theorem 4.15. [4, Theorem 3.4]
Let C and C_0 be d-dimensional copulas and $C \in \mathrm{CDA}(C_0)$. Then they have the same set of tail dependence coefficients, that means if λ_{ij} exists, then $\lambda_{0ij} = \lambda_{ij}$ for $1 \leq i < j \leq d$.

Proof. It is sufficient to concentrate on the bivariate case, $C \in \mathrm{CDA}(C_0)$ with tail dependecne parameters λ and λ_0, respectively. Since C is continuous we find

$$\begin{aligned}\lambda_0 &= 2 - \lim_{q \to 1^-} \frac{\log C_0(q, q)}{\log q} = 2 - \lim_{q \to 1^-} \lim_{n \to \infty} \frac{\log C^n(q^{1/n}, q^{1/n})}{\log q} \\ &= 2 - \lim_{q \to 1^-} \lim_{n \to \infty} \frac{\log C(q^{1/n}, q^{1/n})}{\log(q^{1/n})} = 2 - \lim_{v \to 1^-} \frac{\log C(v, v)}{\log v} \\ &= \lambda.\end{aligned}$$

The tail dependence coefficients of C_0 and C are the same. □

Proposition 4.16. [4, Proposition 3.5]
Let C be a d-dimensional EV copula for which $\lambda_{ij} = 0$ for all $1 \leq i < j \leq d$. Then C must be the independence copula.

Proof. As shown in Remark 3.10, from $\lambda_{ij} = 0$ it follows $C_{ij}(u_1, u_2) = u_1 u_2$ and hence pairwise independence between the components of the corresponding MEVDs. Since for a MEVD pairwise independence implies mutual independence (see for example [14, Proposition 5.27]), we conclude that $C(\boldsymbol{u}) = u_1 \cdots u_d = C^{\mathrm{ind}}(\boldsymbol{u})$. □

The last two results characterise the CDA of the independence copula.

Corollary 4.17 (CDA of the independence copula C^{ind}). A d-dimensional copula C is in the CDA of the independence copula if and only if $\lambda_{ij} = 0$ for all $1 \leq i < j \leq d$.

4.4 Copulas in no copula domain of attraction

In the light of condition *(i)* in Theorem 4.9 it is easy to give an example of a d-dimensional df F which is in no MDA, meaning that $F \notin \mathrm{MDA}(G)$ for all d-dimensional EVDs G. Simply let one of the margins $F_1 \dots, F_d$, not be in a MDA of a univariate EVD. For instance the geometric distribution and the Poisson distribution are in no univariate MDA (see [5, p. 118] for a proof).

Let C be a copula of the d-dimensional df F. Moreover, condition *(ii)* in Theorem 4.9 could fail, meaning that the limit $\lim_{n\to\infty} C^n(\boldsymbol{u}^{1/n})$ does not exist for some $\boldsymbol{u} \in [0,1]^d$. In this case $C \notin \mathrm{CDA}(C_0)$ for all EV copulas C_0. However, finding examples of such copulas C is complicated by the fact that the limit $\lim_{n\to\infty} C^n(\boldsymbol{u}^{1/n})$ highly depends on the behaviour of $C(\boldsymbol{u})$ for values of $\boldsymbol{u} \in [0,1]^d$ close to $\mathbf{1}$. This is due to the fact that $\boldsymbol{u}^{1/n} \to \mathbf{1}$ for $n \to \infty$ if $\boldsymbol{u} > \mathbf{0}$.

In this section we give two different kinds of (pathological) examples of a copula C such that $C \notin \mathrm{CDA}(C_0)$ for all EV copulas C_0. These examples are designed to fill a small gap in EVT literature and hopefully help to better understand the CDA condition (4.1).

4.4.1 An example via diagonal sections

We concentrate on the two-dimensional case, i.e. $d = 2$.

Definition 4.18 (diagonal section). The **diagonal section** δ of a two-dimensional copula C is defined by

$$\delta(u) = C(u,u), \quad u \in [0,1].$$

Lemma 4.19. Every diagonal section δ satisfies

1. $\delta(1) = 1$,
2. $0 \le \delta(u_2) - \delta(u_1) \le 2(u_2 - u_1)$ for all $0 \le u_1 \le u_2 \le 1$,
3. $0 \le \delta(u) \le u$ for all $u \in [0,1]$.

The second property follows from Proposition 1.6. The first and the third properties are obvious by definition.

Definition 4.20 (diagonal copula). Starting with a function $\delta : [0,1] \to [0,1]$ which has the three properties stated in Lemma 4.19, the **diagonal copula** C_δ is given by

$$C_\delta(u_1,u_2) = \min\left\{u_1, u_2, \frac{\delta(u_1)+\delta(u_2)}{2}\right\}, \quad (u_1,u_2) \in [0,1]^2. \tag{4.19}$$

Remark 4.21. Theorem 3.2.12 in [13] ensures that the function C_δ defined in (4.19) is a copula if δ satisfies the three properties stated in Lemma 4.19.

The diagonal copula C_δ has diagonal section δ. Before we analyse the limiting EV copula of a diagonal copula, we need an auxiliary result about asymptotic equivalence. Two functions $f, g : \mathbb{R} \to \mathbb{R}$ are said to be asymptotically equivalent for $x \to x_0 \in \overline{\mathbb{R}}$ if

$$\lim_{x\to x_0} \frac{f(x)}{g(x)} = 1.$$

We use the notation $f \sim g$ for $x \to x_0$.

Lemma 4.22. Let a, b, c and d be four functions such that $a \sim b$ and $c \sim d$ for $x \to x_0$. Then it follows that $ac \sim bd$ for $x \to x_0$.

Proof. One uses an expansion, the triangle inequality and the asymptotic equivalence of the pairs a, b and c, d to see

$$\left|\frac{a(x)c(x)}{b(x)d(x)} - 1\right| = \left|\frac{a(x)}{b(x)}\left(\frac{c(x)}{d(x)} - 1\right) + \frac{a(x)}{b(x)} - 1\right|$$
$$\leq \left|\frac{a(x)}{b(x)}\right|\left|\frac{c(x)}{d(x)} - 1\right| + \left|\frac{a(x)}{b(x)} - 1\right| \to 0 \quad \text{for } x \to x_0,$$

which shows $ac \sim bd$ for $x \to x_0$. □

The following Proposition is a reformulation of Lemma 4.6 in [11]. We give a more detailed proof.

Proposition 4.23. *Let C_δ be a diagonal copula and assume that its tail dependence coefficient λ exists. Then $C_\delta \in CDA(C_0)$, where the EV copula C_0 is given by*

$$C_0(u_1, u_2) = \min\left\{u_1, u_2, (u_1 u_2)^{\frac{2-\lambda}{2}}\right\}, \quad (u_1, u_2) \in [0, 1]^2. \tag{4.20}$$

Proof. We work with the CDA condition in (4.7). For $t \in \mathbb{R}_+$ and $(u_1, u_2) \in [0, 1]^2$

$$C_\delta^t\left(u_1^{1/t}, u_2^{1/t}\right) = \min\left\{u_1, u_2, \left[\frac{1}{2}\left(\delta(u_1^{1/t}) + \delta(u_2^{1/t})\right)\right]^t\right\}.$$

First one treats the case $(u_1, u_2) \in (0, 1)^2$. Obviously

$$\left[\frac{1}{2}\left(\delta(u_1^{1/t}) + \delta(u_2^{1/t})\right)\right]^t = \exp\left\{t \log\left(\frac{\delta(u_1^{1/t}) + \delta(u_2^{1/t})}{2}\right)\right\}.$$

Since $\log y \sim y - 1$ for $y \to 1$, and

$$\frac{\delta(u_1^{1/t}) + \delta(u_2^{1/t})}{2} \to 1 \quad \text{for } t \to \infty,$$

we get from Lemma 4.22 with $a(t) = b(t) = t$ that

$$\lim_{t\to\infty}\left[\frac{1}{2}\left(\delta(u_1^{1/t}) + \delta(u_2^{1/t})\right)\right]^t = \lim_{t\to\infty} \exp\left\{t\left(\frac{\delta(u_1^{1/t}) + \delta(u_2^{1/t})}{2} - 1\right)\right\}. \tag{4.21}$$

The argument of the exponential function in (4.21) can be rewritten as

$$t\left(\frac{\delta(u_1^{1/t}) + \delta(u_2^{1/t})}{2} - 1\right) = \frac{1}{2}\left[-t(1 - u_1^{1/t})\frac{1 - \delta(u_1^{1/t})}{1 - u_1^{1/t}} - t(1 - u_2^{1/t})\frac{1 - \delta(u_2^{1/t})}{1 - u_2^{1/t}}\right]. \tag{4.22}$$

For $u \in (0, 1)$ we have

$$\lim_{t\to\infty} -t(1 - u^{1/t}) = \log u \tag{4.23}$$

and

$$\lim_{t\to\infty} \frac{1-\delta(u^{1/t})}{1-u^{1/t}} = \lim_{q\to 1^-} \frac{1-\delta(q)}{1-q} = 2-\lambda; \tag{4.24}$$

cf. (3.4). An application of Lemma 4.22 to the limit of (4.22) combined with the two results in (4.23) and (4.24) yields that

$$\begin{aligned} \lim_{t\to\infty} t\left(\frac{\delta(u_1^{1/t})+\delta(u_2^{1/t})}{2}-1\right) &= \frac{1}{2}[\log u_1\ (2-\lambda)+\log u_2\ (2-\lambda)] \\ &= \frac{2-\lambda}{2}(\log u_1+\log u_2). \end{aligned} \tag{4.25}$$

Together with (4.21) equation (4.25) shows that

$$\lim_{t\to\infty} \left[\frac{1}{2}\left(\delta(u_1^{1/t})+\delta(u_2^{1/t})\right)\right]^t = (u_1u_2)^{\frac{2-\lambda}{2}}, \quad (u_1,u_2)\in(0,1)^2, \tag{4.26}$$

and therefore

$$\lim_{t\to\infty} C_\delta^t\left(u_1^{1/t},u_2^{1/t}\right) = \min\left\{u_1,u_2,(u_1u_2)^{\frac{2-\lambda}{2}}\right\}, \quad (u_1,u_2)\in(0,1)^2. \tag{4.27}$$

The validity of (4.27) in the boundary cases $(u_1,u_2)\in[0,1]^2\backslash(0,1)^2$ can be shown easily. If $(u_1,u_2)=(1,1)$, both sides of (4.27) are 1. In the cases $(u_1,u_2)\in\{(0,0),(1,0),(0,1)\}$ equation (4.27) obviously holds since C_δ is a copula. Both sides are zero. Since C_δ is symmetrical it remains to consider the case $u_1=1$ and $u_2\in(0,1)$. Apart from (4.24) all arguments of the above proof for the case $(u_1,u_2)\in(0,1)^2$ hold as well. Fortunately, if $u_1=1$, equation (4.22) simplifies to the extent that one does not need to apply (4.24) for $u=1$ to proceed with the proof analogously. □

Now we turn to the case where the tail dependence coefficient λ does not exist.

Corollary 4.24 (A criterion for a differentiable diagonal section δ to ensure that the corresponding diagonal copula C_δ is in no CDA). Let δ' denote the derivative of δ. If the limit $\lim_{x\to 1^-}\delta'(x)$ does not exist, then C_δ is in no CDA.

Proof. Fix $u\in(0,1)$. Because of

$$C_\delta^t\left(u^{1/t},u^{1/t}\right) = \exp\left\{t\log\delta(u^{1/t})\right\}, \quad t>0,$$

we get analogously to (4.21) that

$$\lim_{t\to\infty} C_\delta^t\left(u^{1/t},u^{1/t}\right) = \lim_{t\to\infty} \exp\left\{t(\delta(u^{1/t})-1)\right\}.$$

Thus, the convergence of $C_\delta^t(u^{1/t},u^{1/t})$ for $t\to\infty$ hinges on the existence of the limit $\lim_{t\to\infty} t(\delta(u^{1/t})-1)$. Now

$$\lim_{t\to\infty} t(\delta(u^{1/t})-1) = \lim_{q\to 0^+} \frac{\delta(u^q)-1}{q}$$

and for a differentiable δ an application of l'Hôspital's rule yields that

$$\lim_{q\to 0^+} \frac{\delta(u^q)-1}{q} = \lim_{q\to 0^+} \delta'(u^q)u^q\log u = \log u \lim_{x\to 1^-}\delta'(x).$$

If the last limit does not exist, then the corresponding diagonal copula C_δ is in no CDA. □

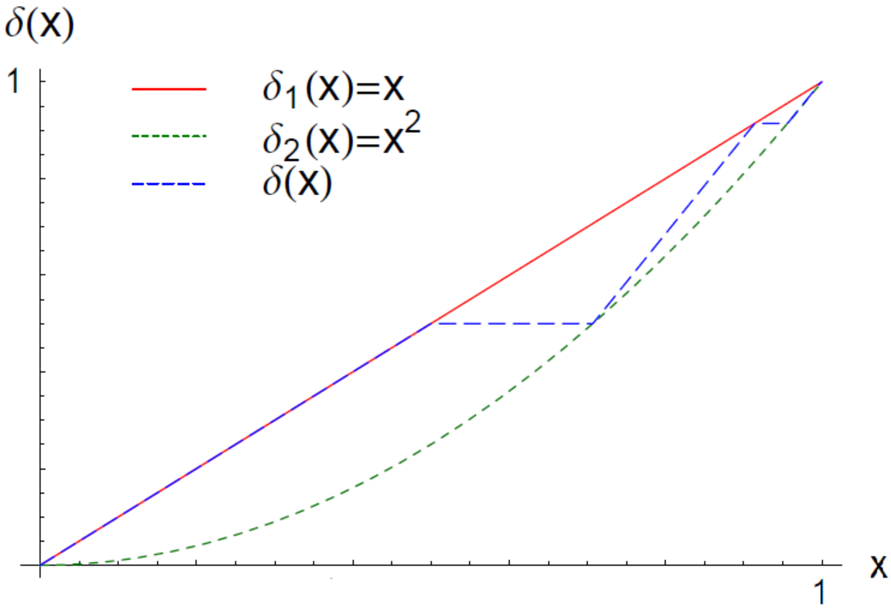

Figure 4.1: *The diagonal section δ from (4.28). δ_1 and δ_2 are the diagonal sections of the comonotonicity and the independence copula, respectively. The idea of the construction is best understood when looking at the graph of δ. First start with $\delta_1(x)$ for $x \leq 1/2$. Then continue horizontally until you hit δ_2. Then go along a line with slope 2. Once you hit δ_1 again you continue horizontally, and so forth. The secret behind this construction is that δ_1 and δ_2 are diagonal sections of EV copulas with different tail dependence coefficients, namely 1 and 0. The figure is taken from [11].*

If λ does not exist, then the limit $\lim_{t\to\infty} C_\delta^t(u_1^{1/t}, u_2^{1/t})$ does not necessarily have to exist for all $(u_1, u_2) \in [0,1]^2$.

In the following we build a diagonal section δ such that $\lim_{t\to\infty} C_\delta^t(u^{1/t}, u^{1/t})$ does not exist for all $u \in (0,1)$. For this aim we use a construction inspired by [11].
Let $x_1 = \frac{1}{2}$ and for $i \in \mathbb{N}$

$$x_{2i} = \sqrt{x_{2i-1}},$$
$$x_{2i+1} = 2x_{2i} - x_{2i}^2.$$

The sequence x_i, $i \in \mathbb{N}$, is increasing and converges to 1. Next we define an increasing function $\delta : [0,1] \to [0,1]$ by

$$\delta(x) = \begin{cases} 1 & x = 1, \\ x & x \leq \frac{1}{2}, \\ x_{2i-1} & x_{2i-1} \leq x < x_{2i}, \\ x_{2i}^2 + 2(x - x_{2i}) & x_{2i} \leq x < x_{2i+1}. \end{cases} \tag{4.28}$$

Lemma 4.25. The function δ in (4.28) satisfies the three properties in Lemma 4.19.

Proof. We discuss the three properties:

1. $\delta(1) = 1$ by definition.
2. $0 \leq \delta(u_2) - \delta(u_1) \leq 2(u_2 - u_1)$ for all $0 \leq u_1 \leq u_2 \leq 1$:
   Let $0 \leq u_1 \leq u_2 \leq 1$. First of all $\delta(u_2) - \delta(u_1) \geq 0$ holds since δ is increasing. For the other inequality consider the following cases.
   - $u_1, u_2 \in [0, 1/2]$: Then $\delta(u_2) - \delta(u_1) = u_2 - u_1 \leq 2(u_2 - u_1)$.
   - $u_1, u_2 \in [x_{2i-1}, x_{2i}]$ for some $i \in \mathbb{N}$: Then $\delta(u_2) - \delta(u_1) = 0 \leq 2(u_2 - u_1)$.
   - $u_1, u_2 \in [x_{2i}, x_{2i+1}]$ for some $i \in \mathbb{N}$: Then $\delta(u_2) - \delta(u_1) = 2(u_2 - u_1)$.

   For the general case, $u_1, u_2 \in [0, 1]$, let $D = \{j \in \mathbb{N} : x_j \in (u_1, u_2)\}$. The set $\{x_j : j \in D\}$ contains all the points of the sequence $(x_i)_{i \in \mathbb{N}}$ which lie between u_1 and u_2. Without loss of generality we can write $D = \{c, c+1, \ldots, d-1, d\}$ for integers c and d. Now

$$
\begin{aligned}
\delta(u_2) - \delta(u_1) &= (\delta(u_2) - \delta(x_d)) + \sum_{j=c+1}^{d} (\delta(x_j) - \delta(x_{j-1})) + (\delta(x_c) - \delta(u_1)) \\
&\leq 2(u_2 - x_d) + \sum_{j=c+1}^{d} 2(x_j - x_{j-1}) + 2(x_c - u_1) \\
&= 2(u_2 - u_1)
\end{aligned}
$$

   because for each of the summands one of the above three cases applies.
3. $\delta(u) \leq u$ for all $u \in [0, 1]$:
   - If $u \in [0, 1/2]$, then $\delta(u) = u \leq u$.
   - If $u \in [x_{2i-1}, x_{2i})$ for some $i \in \mathbb{N}$, then $\delta(u) = x_{2i-1} \leq u$.
   - If $u \in [x_{2i}, x_{2i+1})$ for some $i \in \mathbb{N}$, then

$$
\delta(u) = x_{2i}^2 + 2(u - x_{2i}) = u + x_{2i}^2 - x_{2i} + (u - x_{2i}).
$$

     A calculation yields that $x_{2i+1} - x_{2i} = x_{2i} - x_{2i}^2$ and consequently

$$
\delta(u) = u + x_{2i}^2 - x_{2i} + (u - x_{2i}) \leq u.
$$

□

By Lemma 4.25 and Remark 4.21 the corresponding C_δ is a copula. We study its tail dependence coefficient. In view of (3.3) we look at

$$
\lim_{x \to 1^-} \frac{1 - 2x + \delta(x)}{1 - x}.
$$

On the one hand the sequence $y_i := x_{2i-1}$, $i \in \mathbb{N}$, converges to 1 and

$$
\lim_{i \to \infty} \frac{1 - 2y_i + \delta(y_i)}{1 - y_i} = \lim_{i \to \infty} \frac{1 - 2y_i + y_i}{1 - y_i} = 1.
$$

While on the other hand for $z_i := x_{2i}$, $i \in \mathbb{N}$, we have

$$\lim_{i\to\infty} \frac{1 - 2z_i + \delta(z_i)}{1 - z_i} = \lim_{i\to\infty} \frac{1 - 2z_i + z_i^2}{1 - z_i} = \lim_{i\to\infty} (1 - z_i) = 0.$$

Therefore the tail dependence coefficient λ does not exist.

Finally we want to show that $\lim_{t\to\infty} C_\delta^t(u^{1/t}, u^{1/t})$ does not exist for all $u \in (0, 1)$, violating the CDA condition. Fix $u \in (0, 1)$. Using the sequence $y_i = x_{2i-1}$, $i \in \mathbb{N}$, from above, we define a sequence t_i, $i \in \mathbb{N}$, by choosing t_i such that

$$u^{1/t_i} = y_i$$

which is equivalent to

$$t_i = \frac{\log u}{\log y_i}.$$

The sequence t_i, $i \in \mathbb{N}$, is increasing and converges to infinity. One calculates

$$\lim_{i\to\infty} C_\delta^{t_i}(u^{1/t_i}, u^{1/t_i}) = \lim_{i\to\infty} \delta(u^{1/t_i})^{t_i} = \lim_{i\to\infty} (u^{1/t_i})^{t_i} = u. \tag{4.29}$$

Analogously, using the sequence $z_i = x_{2i}$, $i \in \mathbb{N}$, we define a sequence s_i, $i \in \mathbb{N}$, by choosing s_i such that

$$u^{1/s_i} = z_i.$$

Again the sequence s_i, $i \in \mathbb{N}$, is increasing and converges to infinity. One finds

$$\lim_{i\to\infty} C_\delta^{s_i}(u^{1/s_i}, u^{1/s_i}) = \lim_{i\to\infty} \delta(u^{1/s_i})^{s_i} = \lim_{i\to\infty} (u^{2/s_i})^{s_i} = u^2. \tag{4.30}$$

Since (4.29) and (4.30) do not lead to the same limit one concludes that the diagonal copula C_δ cannot be in a CDA.

Remark 4.26. The construction in this subsection can even be extended to get a diagonal copula with differentiable diagonal section. Note that apart from the points x_i, $i \in \mathbb{N}$, δ is differentiable on $[0, 1]$. However, one can use quadratic smoothing on appropriately small, symmetric intervals around these points to get a slightly modified $\tilde{\delta}$ which is differentiable on $[0, 1]$. Lemma 4.25 then also holds for $\tilde{\delta}$ instead of δ because $\tilde{\delta}' \leq 2$ on $[0, 1]$. We briefly sketch the last step one would need in order to use Corollary 4.24 to show that a $C_{\tilde{\delta}}$, constructed that way, is in no CDA. The sequence $s_i = \frac{x_i + x_{i+1}}{2}$, $i \in \mathbb{N}$, converges to 1. Because $\tilde{\delta}(x) = \delta(x)$ in an open neighbourhood of s_i (assuming a reasonable choice of intervals for smoothing), we see from the construction of the function δ in (4.28) that $\tilde{\delta}'(s_i) = \delta'(s_i) \in \{0, 2\}$ and $\tilde{\delta}'(s_i) \neq \tilde{\delta}'(s_{i+1})$.

4.4.2 An example via spectral measures

A multivariate extreme value distribution is characterised by its spectral measure (cf. Proposition 2.14). Let $E = [0, \infty]^d \backslash \{\mathbf{0}\}$, $\|\cdot\|$ be a norm on $\mathbb{R}^d$ and $\mathcal{C} = \{\boldsymbol{w} \in E \ : \ \|\boldsymbol{w}\| = 1\}$. The set $\boldsymbol{\mathcal{S}}_d$ in (2.12) was

$$\{\boldsymbol{w} \in [0, \infty]^d \ : \ w_1 + \cdots + w_d = 1\} = \{\boldsymbol{w} \in E \ : \ \|\boldsymbol{w}\|_1 = 1\},$$

where $\|\cdot\|_1$ denotes the sum-norm. $\boldsymbol{\mathcal{S}}_d$ is very similar to the set $\mathcal{C}$. In fact for a d-dimensional EVD G and the norm $\|\cdot\|$ there also exists a finite measure S on $\mathcal{C}$ such that

$$G(\boldsymbol{x}) = \exp\left\{\int_{\mathcal{C}} \min_{1\leq i\leq n}\{w_i \log G_i(x_i)\}\ \mathrm{d}S(\boldsymbol{w})\right\}, \quad \boldsymbol{x}\in\mathbb{R}^d. \tag{4.31}$$

This is formula (2.12) with $\boldsymbol{\mathcal{S}}_d$ replaced by $\mathcal{C}$ and H replaced by S. Note that different spectral measures S_1 and S_2 lead to different MEVDs.

Remark 4.27. Although (4.31) is more general than (2.12), we preferred (2.12) because of the simpler formulas connecting the spectral measure representation to other characterisations. For the purpose of this section a different norm than the sum-norm is more useful.

Theorem 4.28. [6, Theorem 2]
Let $\boldsymbol{X}$ be a d-dimensional random vector with df F and let $\mathcal{B}(\mathcal{C})$ denote the Borel sets of $\mathcal{C}$. Then the following statements are equivalent.

(D1) The df F is in the MDA of a MEVD G with standard Fréchet margins.

(D2)
$$\lim_{t\to\infty} t\mathbb{P}\Big(\|\boldsymbol{X}\| > t, \frac{\boldsymbol{X}}{\|\boldsymbol{X}\|}\in A\Big) = \frac{S(A)}{S(\mathcal{C})} =: \sigma(A), \quad \forall A\in\mathcal{B}(\mathcal{C}),$$
where S is the spectral measure of G with respect to $\|\cdot\|$.

$\sigma(\cdot)$ is a probability measure on $\mathcal{C}$. From now on $d = 2$. We are in a position to specify our norm of choice, the Euclidean norm:

$$\|\boldsymbol{x}\| := \sqrt{x_1^2 + x_2^2}, \quad \boldsymbol{x}\in\mathbb{R}^2.$$

Let Θ_0 and Θ_1 be two $(0, \frac{\pi}{2})$-valued random variables such that there exists a set $A_0 \in \mathcal{B}(\mathcal{C})$ with

$$\mathbb{P}\big((\cos\Theta_0, \sin\Theta_0)^\top \in A_0\big) \neq \mathbb{P}\big((\cos\Theta_1, \sin\Theta_1)^\top \in A_0\big). \tag{4.32}$$

Furthermore, let the random variable $R \geq 1$ have the tail df $\mathbb{P}(R > x) = x^{-1},\ x \geq 1$. We assume that R is independent of Θ_0 and Θ_1. Next consider the two-dimensional random vectors

$$\boldsymbol{X}_i = (R\cos\Theta_i, R\sin\Theta_i)^\top, \quad i = 0, 1$$

with dfs F_0 and F_1, respectively. For $i \in \{0, 1\}$ one finds

$$\begin{aligned}\lim_{t\to\infty} t\mathbb{P}\Big(\|\boldsymbol{X}_i\| > t, \frac{\boldsymbol{X}_i}{\|\boldsymbol{X}_i\|}\in A\Big) &= \lim_{t\to\infty} t\mathbb{P}\big(R > t, (\cos\Theta_i, \sin\Theta_i)^\top \in A\big)\\ &= \mathbb{P}\big((\cos\Theta_i, \sin\Theta_i)^\top \in A\big)\\ &=: \sigma_i(A), \quad \forall A\in\mathcal{B}(\mathcal{C}).\end{aligned}$$

Due to the construction $\sigma_0 \neq \sigma_1$. According to Theorem 4.28, $F_0 \in \mathrm{MDA}(G_0)$ and $F_1 \in \mathrm{MDA}(G_1)$, where G_0 and G_1 are two different MEVDs with standard Fréchet margins.

Now consider the densities f_0 and f_1 of F_0 and F_1. The idea for the following construction stems from [10, p. 5]. Take $\boldsymbol{y} \in \mathbb{R}^2$ with $\|\boldsymbol{y}\| > 1$. Then there exists a unique integer $j \in \mathbb{N}$ such that

$$\|\boldsymbol{y}\| \in (j!, (j+1)!].$$

Let $\boldsymbol{X}$ be an $\mathbb{R}^2_+$-valued random vector with density f given by

$$f(\boldsymbol{y}) = 0 \quad \text{for } \|\boldsymbol{y}\| \in (0,1],\ \boldsymbol{y} > \boldsymbol{0},$$

and

$$f(\boldsymbol{y}) = \begin{cases} f_0(\boldsymbol{y}) & \text{if } j \text{ is odd,} \\ f_1(\boldsymbol{y}) & \text{if } j \text{ is even,} \end{cases} \quad \text{for } \|\boldsymbol{y}\| \in (j!, (j+1)!],\ \boldsymbol{y} > \boldsymbol{0}.$$

The function f is a proper density because

$$\mathbb{P}\big(\|\boldsymbol{X}_0\| \in (j!, (j+1)!]\big) = \mathbb{P}\big(\|\boldsymbol{X}_1\| \in (j!, (j+1)!]\big), \quad \forall j \in \mathbb{N},$$

and f_0, f_1 are densities on $\mathbb{R}^2_+$. It can be shown that the marginal dfs of $\boldsymbol{X}$ are in MDA(Φ_1), (see for example [5, Proposition A 3.8]). The copula of $\boldsymbol{X}$ would be in the CDA of an EV copula if and only if (D2) in Theorem 4.28 holds for $\boldsymbol{X}$. However, one finds sequences

$$u_n = (2n)! \quad \text{and} \quad v_n = (2n+1)! \quad \text{for } n \in \mathbb{N}$$

which lead to different limits in (D2). For u_n and $A \in \mathcal{B}(\mathcal{C})$ we have

$$\begin{aligned} u_n \mathbb{P}\Big(\|\boldsymbol{X}\| > u_n, \frac{\boldsymbol{X}}{\|\boldsymbol{X}\|} \in A\Big) &= u_n \mathbb{P}\Big((2n)! < \|\boldsymbol{X}\| \le (2n+1)!, \frac{\boldsymbol{X}}{\|\boldsymbol{X}\|} \in A\Big) \\ &\quad + u_n \mathbb{P}\Big(\|\boldsymbol{X}\| > (2n+1)!, \frac{\boldsymbol{X}}{\|\boldsymbol{X}\|} \in A\Big), \end{aligned}$$

where the last term can be neglected when taking the limit since

$$u_n \mathbb{P}\Big(\|\boldsymbol{X}\| > (2n+1)!, \frac{\boldsymbol{X}}{\|\boldsymbol{X}\|} \in A\Big) \le (2n)!\, \mathbb{P}\big(R > (2n+1)!\big) = \frac{1}{2n+1}.$$

For $\|\boldsymbol{y}\| \in ((2n)!, (2n+1)!]$ the density $f(\boldsymbol{y}) = f_1(\boldsymbol{y})$ and therefore the probability $\mathbb{P}((2n)! < \|\boldsymbol{X}\| \le (2n+1)!, \frac{\boldsymbol{X}}{\|\boldsymbol{X}\|} \in A)$ can be written as

$$\mathbb{P}\Big(\|\boldsymbol{X}_1\| > (2n)!, \frac{\boldsymbol{X}_1}{\|\boldsymbol{X}_1\|} \in A\Big) - \mathbb{P}\Big(\|\boldsymbol{X}_1\| > (2n+1)!, \frac{\boldsymbol{X}_1}{\|\boldsymbol{X}_1\|} \in A\Big). \tag{4.33}$$

Using (4.33) we get

$$\begin{aligned} \lim_{n\to\infty} u_n \mathbb{P}\Big(\|\boldsymbol{X}\| > u_n, \frac{\boldsymbol{X}}{\|\boldsymbol{X}\|} \in A\Big) &= \lim_{n\to\infty} (2n)!\, \mathbb{P}\Big(\|\boldsymbol{X}_1\| > (2n)!, \frac{\boldsymbol{X}_1}{\|\boldsymbol{X}_1\|} \in A\Big) \\ &\quad - \lim_{n\to\infty} \frac{1}{2n+1} (2n+1)!\, \mathbb{P}\Big(\|\boldsymbol{X}_1\| > (2n+1)!, \frac{\boldsymbol{X}_1}{\|\boldsymbol{X}_1\|} \in A\Big) \\ &= \sigma_1(A) - \lim_{n\to\infty} \frac{1}{2n+1} \sigma_1(A) \\ &= \sigma_1(A), \quad \forall A \in \mathcal{B}(\mathcal{C}). \end{aligned}$$

Analogously, it follows that

$$\lim_{n\to\infty} v_n \mathbb{P}\Big(\|\boldsymbol{X}\| > v_n, \frac{\boldsymbol{X}}{\|\boldsymbol{X}\|} \in A\Big) = \sigma_0(A), \quad \forall A \in \mathcal{B}(\mathcal{C}).$$

Setting $A = A_0$ — in this case $\sigma_1(A_0) \neq \sigma_0(A_0)$ by construction — we see that (D2) does not hold and conclude that the copula C of $\boldsymbol{X}$ is in no CDA.

Remark 4.29. The tail dependence coefficient of the copula C might exist. Remember that with the diagonal copula approach of the previous subsection this was impossible. From Θ_0 and Θ_1 one gets spectral measures H_0 and H_1, which characterise two EV copulas via (4.14) and (4.15). By a careful choice of Θ_0 and Θ_1 one could achieve that those two EV copulas have the same tail dependence coefficient. Under these circumstances the author is optimistic that the tail dependence coefficient of the resulting copula C does exist.

Chapter 5

Multivariate Threshold Models

5.1 Multivariate threshold exceedances

In an extreme value analysis it is important to exploit as much relevant information as is available. Multivariate threshold models are concerned with all observations for which at least one component exceeds some high value. Because of their more efficient use of the often limited data they are sometimes preferred over componentwise block maxima models. In general, the use of block maxima is problematic if some blocks contain numerous high values whereas others do not contain any. This has to be kept in mind when modelling the margins of MEV distributions. To estimate the underlying copula for multivariate data, however, all information can be used.

Let $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n \in \mathbb{R}^d$ be realisations of a d-dimensional random vector with df F. The df F is not known, but it is assumed to be in the MDA of some MEVD G with EV copula C_0. Furthermore, $\mathcal{C}$ is a set of parametric EV copulas and $C_0 \in \mathcal{C}$, that means

$$\mathcal{C} = \{C_\theta : \theta \in \Theta\},$$

with $\Theta \subset \mathbb{R}^p$ for some $p \in \mathbb{N}$. Our aim is to model the upper tail of the df F above a high threshold $\boldsymbol{u} \in \mathbb{R}^d$. The threshold $\boldsymbol{u}$ satisfies $F_j(u_j) \approx 1$ for $j = 1, \ldots, d$. We are interested in an approximation for $F(\boldsymbol{x})$ for $\boldsymbol{x} > \boldsymbol{u}$. Note that only the realisations $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ are known.

First, we derive an approximation for the tails of the margins. The Pickands–Balkema–de Haan theorem (Theorem A.4) suggests to use the generalised Pareto distribution (see Definition A.3). Let H be a univariate df, which is in some MDA, and $u \in \mathbb{R}$ such that $H(u) \approx 1$. Then Theorem A.4 basically says that

$$H_u(x) \approx 1 - \left(1 + \xi \frac{x}{\beta}\right)_+^{-1/\xi}, \quad 0 < x < x_H - u, \tag{5.1}$$

where x_H denotes the right endpoint of H, that is $\sup\{x \in \mathbb{R} : H(x) < 1\}$, and the mean excess function H_u is given by

$$H_u(x) = \frac{H(x+u) - H(u)}{1 - H(u)}, \quad 0 < x < x_H - u. \tag{5.2}$$

Combining (5.2) and (5.1) yields

$$\begin{aligned} H(x+u) &= (1 - H(u))H_u(x) + H(u) \\ &\approx 1 - (1 - H(u))\left(1 + \xi \frac{x}{\beta}\right)_+^{-1/\xi}, \quad 0 < x < x_H - u. \end{aligned}$$

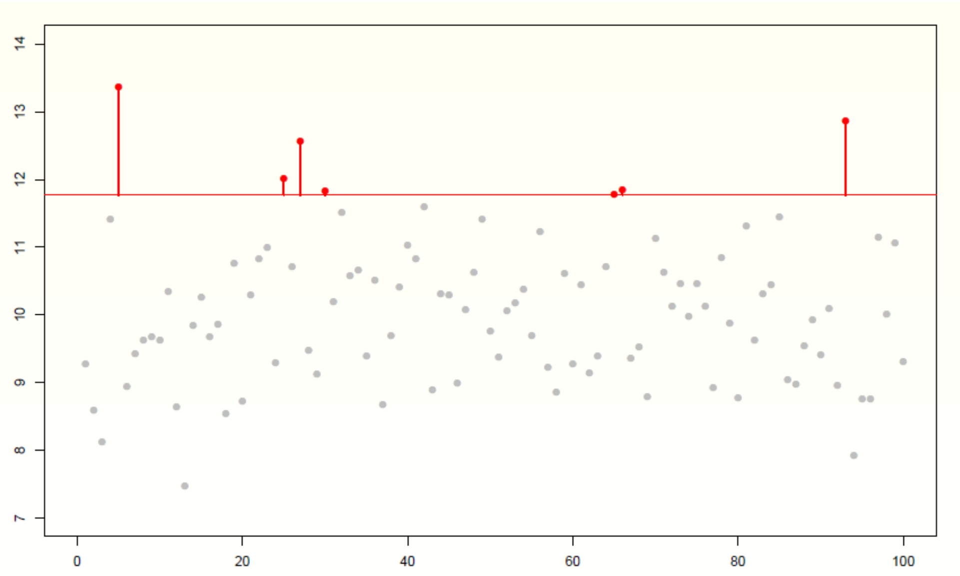

Figure 5.1: *Excesses over a threshold: the generalised Pareto distribution approach.*

Finally, with $\lambda := 1 - H(u)$, we get a GPD based parametric form for the tail:

$$H(x) \approx 1 - \lambda \left(1 + \xi \frac{x-u}{\beta}\right)_+^{-1/\xi}, \quad x > u. \tag{5.3}$$

Our next step is to approximate the copula C (cf. [12, p. 320]). From (4.7) it follows that

$$C^t(\boldsymbol{v}^{1/t}) \approx C_0(\boldsymbol{v})$$

for $\boldsymbol{v} \in [0,1]^d$ and t sufficiently large. If we now set $\boldsymbol{w} = \boldsymbol{v}^{1/t}$, this is equivalent to

$$C(\boldsymbol{w}) \approx C_0^{1/t}(\boldsymbol{w}^t) = C_0(\boldsymbol{w}), \quad \boldsymbol{w} \in [0,1]^t. \tag{5.4}$$

Since $\boldsymbol{v}^{1/t} \to \mathbf{1}$ for all $\boldsymbol{v} \in (0,1]^d$ as $t \to \infty$, this approximation should only be used for the tail of C. Fortunately, $F_i(x_i)$ is close to one for $\boldsymbol{x} > \boldsymbol{u}$ and $i \in \{1, \ldots, d\}$.

The results in (5.3) and (5.4) justify the following model:

$$F(\boldsymbol{x}) \approx C_\theta \left(1 - \lambda_1 \left(1 + \xi_1 \frac{x_1 - u_1}{\beta_1}\right)_+^{-1/\xi_1}, \ldots, 1 - \lambda_d \left(1 + \xi_d \frac{x_d - u_d}{\beta_d}\right)_+^{-1/\xi_d}\right) =: \widetilde{F}(\boldsymbol{x}), \quad \boldsymbol{x} > \boldsymbol{u}, \tag{5.5}$$

where θ is the copula parameter, $\lambda_i = 1 - F_i(u_i)$, $\beta_i > 0$ and $\xi_i \in \mathbb{R}$ for $i = 1, \ldots, d$ (see Definition A.3).

A big advantage of (5.5) is that it provides a fully parametric model for the tail. Since all EV copulas can be approximated by parametric ones, the assumption that C_0 is parametric is no restriction. Given some observations all parameters can be estimated with maximum-likelihood techniques. We briefly sketch the necessary steps.

1. Determine a suitable threshold $\boldsymbol{u}$ by looking at the empirical tail dfs and plotting the empirical tail excess functions. Since we know that the mean excess function of the GPD is linear in u, we can search for a region where the empirical mean excess functions for the given data become linear. "For such $\boldsymbol{u} \in \mathbb{R}^d$ an approximation of F_{u_i} by a GPD seems reasonable" [5, p. 167].

2. Maximise the likelihood

$$L(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n; \boldsymbol{\xi}, \boldsymbol{\beta}, \theta) = \prod_{i=1}^{n} L(\boldsymbol{x}_i),$$

   where $\boldsymbol{\xi} = (\xi_1, \dots, \xi_d)$, $\boldsymbol{\beta} = (\beta_1, \dots, \beta_d)$ and θ is the parameter of the copula. The likelihood contribution $L(\boldsymbol{x}_i)$ of an observation $\boldsymbol{x}_i$ depends on which of its coordinates exceed the corresponding threshold coordinates. This leads to the so-called censored likelihood approach.
   If in an observation $\boldsymbol{x}_i$ no component exceeds its corresponding threshold (i.e. $\boldsymbol{x}_i \leq \boldsymbol{u}$), then $L(\boldsymbol{x}_i) = \widetilde{F}(\boldsymbol{u})$ which does not influence the maximisation. In general, one works with the censored data $\max(\boldsymbol{x}_i, \boldsymbol{u})$ and only takes partial derivatives of $\widetilde{F}$ with respect to those $x_{j,i}$ which exceed their corresponding thresholds u_i. More details can be found in [1].

Example 5.1. The class of Galambos copulas

$$C^{\text{Gal}}_{\theta,\alpha,\beta}(u_1, u_2) = u_1 u_2 \exp\left\{ \left((-\alpha \log u_1)^{-\theta} + (-\beta \log u_2)^{-\theta} \right)^{-1/\theta} \right\}, \quad (u_1, u_2) \in [0,1]^2, \tag{5.6}$$

where $\alpha, \beta \in [0,1]$ and $0 < \theta < \infty$, is a class of three-parameter extreme value copulas (see [12, Chapter 7]) and a possible choice for $\mathcal{C}$ in the bivariate case. ○

5.2 Threshold copulas

In this section we condition a two-dimensional random vector $\boldsymbol{X} = (X_1, X_2)$ to exceed high thresholds and study the arising copulas. Similar considerations in the univariate case have lead to the generalised Pareto distributions, where one studied excess distributions. Now the focus is on the dependence structure. The presented material is based on [4] and [12].

Let $\boldsymbol{X} = (X_1, X_2)$ be a two-dimensional random vector with df F, which has margins F_1, F_2 and copula C. We condition on the event

$$A_v = \{X_1 > F_1^{\leftarrow}(v), X_2 > F_2^{\leftarrow}(v)\}, \quad 0 \leq v < 1.$$

This means that both components exceed their v-quantile. The conditional probability that both components exceed even higher quantiles is

$$\mathbb{P}(X_1 > F_1^{\leftarrow}(u_1), X_2 > F_2^{\leftarrow}(u_2) | A_v) = \frac{\bar{C}(u_1, u_2)}{\bar{C}(v, v)}, \quad \text{for } (u_1, u_2) \in [v, 1]^2, \tag{5.7}$$

where $\bar{C}(u_1, u_2) = \mathbb{P}(X_1 > F_1^{\leftarrow}(u_1), X_2 > F_2^{\leftarrow}(u_2)) = 1 - u_1 - u_2 + C(u_1, u_2)$ denotes the survival function of the copula C. From (1.7) one gets that

$$\bar{G}_{(v)}(u_1, u_2) = \frac{\bar{C}(u_1, u_2)}{\bar{C}(v, v)}$$

defines a bivariate survival function on $[v, 1]^2$. For the corresponding marginal survival functions on $[v, 1]$ we have

$$\bar{G}_{1,(v)}(u) = \mathbb{P}(X_1 > F_1^{\leftarrow}(u)|A_v) = \frac{\bar{C}(u, v)}{\bar{C}(v, v)}$$

and

$$\bar{G}_{2,(v)}(u) = \mathbb{P}(X_2 > F_2^{\leftarrow}(u)|A_v) = \frac{\bar{C}(v, u)}{\bar{C}(v, v)}.$$

By (1.6), a survival copula $\widehat{C}_{(v)}$ can be used to write

$$\bar{G}_{(v)}(u_1, u_2) = \widehat{C}_{(v)}(\bar{G}_{1,(v)}(u_1), \bar{G}_{2,(v)}(u_2)), \quad \boldsymbol{u} \in [v, 1]^2. \tag{5.8}$$

Here $\widehat{C}_{(v)}$ is the survival copula of $C_{(v)}$, with $C_{(v)}$ being a copula of $G_{(v)}$.

Definition 5.2 (upper threshold copula). The copula $C_{(v)}$ is called **upper threshold copula** of C at level v.

Equation (5.8) can be also written as

$$\widehat{C}_{(v)}\left(\frac{\bar{C}(u_1, v)}{\bar{C}(v, v)}, \frac{\bar{C}(v, u_2)}{\bar{C}(v, v)}\right) = \frac{\bar{C}(u_1, u_2)}{\bar{C}(v, v)}, \quad \boldsymbol{u} \in [v, 1]^2. \tag{5.9}$$

From (1.7) we get

$$C_{(v)}(u_1, u_2) = \widehat{C}_{(v)}(1 - u_1, 1 - u_2) - 1 + u_1 + u_2, \quad \boldsymbol{u} \in [0, 1]^2. \tag{5.10}$$

Example 5.3. Our aim is to calculate the corresponding upper threshold copulas at an arbitrary level $v \in [0, 1)$ for the family of copulas

$$C_\delta(u_1, u_2) = \delta \min\{u_1, u_2\} + (1 - \delta)u_1 u_2, \quad \delta \in [0, 1]. \tag{5.11}$$

The first argument of $\widehat{C}_{\delta,(v)}$ in (5.9) is

$$\frac{\bar{C}_\delta(u_1, v)}{\bar{C}_\delta(v, v)} = \frac{1 - u_1 - v + C_\delta(u_1, v)}{\bar{C}_\delta(v, v)} = \frac{1 - u_1 - v + \delta v + (1 - \delta)u_1 v}{\bar{C}_\delta(v, v)}, \quad u_1 \in [v, 1].$$

With

$$a = \frac{-1 + (1 - \delta)v}{\bar{C}_\delta(v, v)} < 0,$$

one easily sees

$$\frac{\bar{C}_\delta(u_1, v)}{\bar{C}_\delta(v, v)} = u_1 a - a, \quad u_1 \in [v, 1].$$

Since C_δ is symmetric it follows for the second argument that

$$\frac{\bar{C}_\delta(v,u_2)}{\bar{C}_\delta(v,v)} = u_2 a - a, \quad u_2 \in [v,1].$$

Next we put $x_1 = u_1 a - a$ and $x_2 = u_2 a - a$. The pair (x_1, x_2) is in $[0,1]^2$ because (u_1, u_2) is in $[v,1]^2$. Using (5.9) we get

$$\begin{aligned}\widehat{C}_{\delta,(v)}(x_1,x_2) = \frac{1}{\bar{C}_\delta(v,v)}\Big[1 &- \left(\frac{x_1+a}{a}\right) - \left(\frac{x_2+a}{a}\right) + \delta \min\left\{\left(\frac{x_1+a}{a}\right), \left(\frac{x_2+a}{a}\right)\right\} \\ &+ (1-\delta)\left(\frac{x_1+a}{a}\right)\left(\frac{x_2+a}{a}\right)\Big],\end{aligned}$$

which can be simplified to

$$\widehat{C}_{\delta,(v)}(x_1,x_2) = \frac{\delta \min\{x_1,x_2\} + (1-\delta)(1-v)x_1 x_2}{1-(1-\delta)v}. \tag{5.12}$$

Equation (5.10) finally yields

$$C_{\delta,(v)}(u_1,u_2) = \frac{\delta \min\{1-u_1, 1-u_2\} + (1-\delta)(1-v)(1-u_1)(1-u_2)}{1-(1-\delta)v} - 1 + u_1 + u_2, \tag{5.13}$$

where $\boldsymbol{u} \in [0,1]^2$. $C_{\delta,(v)}$ is the upper threshold copula of C_δ at level v. ◦

When analysing extremes large values of the quantile $v \in [0,1)$ are most interesting.

Definition 5.4 (limiting upper threshold copula). The **limiting upper threshold copula** C^{up} of a two-dimensional copula C, with upper threshold copulas $C_{(v)}$ at level $v \in [0,1)$, is given by

$$C^{\text{up}}(u_1,u_2) = \lim_{v\to 1^-} C_{(v)}(u_1,u_2), \quad \boldsymbol{u} \in [0,1]^2,$$

provided the limit exists.

Example 5.5. We continue Example 5.3 and calculate the limiting upper threshold copulas C_δ^{up} for $\delta \in [0,1]$. From (5.13) one easily sees that

$$\begin{aligned} C_\delta^{\text{up}}(u_1,u_2) &= \lim_{v\to 1^-} C_{\delta,(v)}(u_1,u_2) \\ &= \min\{1-u_1, 1-u_2\} - 1 + u_1 + u_2 \\ &= \min\{u_1,u_2\}, \quad \forall \delta \in [0,1]. \end{aligned}$$

In particular the limiting upper threshold copula $\min\{u_1,u_2\}$ is a member of the copula family in (5.11), namely C_1. Moreover, we have shown in Example 5.3 that for all levels $v \in [0,1)$ the upper threshold copulas of C_1 coincide with C_1. This is a remarkable stability property of C_1. ◦

The required calculations in Example 5.3 to establish (5.12) were rather tedious, although C_δ had a relatively simple form. Therefore, an alternative way to determine the limiting upper threshold copula C^{up} is helpful.

Theorem 5.6. [4, Theorem 3.18], [12, Theorem 7.56]
If C is a symmetric two-dimensional copula with upper tail dependence coefficient $\lambda > 0$ satisfying $C \in \mathrm{CDA}(C_0)$, then C has a limiting upper threshold copula C^{up} which is the survival copula of the df

$$G(x_1, x_2) = \frac{(x_1 + x_2)\left(1 - A\left(\frac{x_1}{x_1+x_2}\right)\right)}{\lambda}, \quad (x_1, x_2) \in [0,1]^2, \tag{5.14}$$

where A is the Pickands dependence function of the EV copula C_0.

Theorem 5.6 also shows that two copulas C_1 and C_2 belonging to the same $\mathrm{CDA}(C_0)$ necessarily have the same limiting upper threshold copula. By Theorem 4.15, C_1, C_2 and C_0 have the same tail dependence coefficient, and therefore the dfs in (5.14) are identical for C_1 and C_2.

Remark 5.7. For a symmetric copula $C \in \mathrm{CDA}(C_0)$ with tail dependence coefficient λ and limiting upper threshold copula C^{up}, with tail dependence coefficient λ^{up}, it can be shown that $\lambda^{\mathrm{up}} \in [\lambda, 1]$.

Example 5.8. We already know the copula

$$C(u_1, u_2) = \frac{u_1 u_2 (2 - u_1 - u_2)}{1 - u_1 u_2}$$

from Example 4.11, where it was shown that C is in the CDA of the extreme value copula

$$C_0(u_1, u_2) = \exp\left\{\log u_1 u_2 - \frac{\log u_1 \log u_2}{\log u_1 u_2}\right\}$$

with Pickands dependence function $A(t) = t^2 - t + 1$ for $t \in [0,1]$. Its tail dependence coefficient is

$$\lambda = 2\left(1 - A\left(\frac{1}{2}\right)\right) = \frac{1}{2} > 0.$$

We apply Theorem 5.6:

$$G(x_1, x_2) = \frac{2x_1 x_2}{x_1 + x_2}, \quad (x_1, x_2) \in [0,1]^2.$$

The margins of G are

$$G_1(x) = G_2(x) = G(1, x) = \frac{2x}{1 + x}, \quad x \in [0,1],$$

and the copula C_G of G is given by

$$\begin{aligned} C_G(u_1, u_2) &= G(G_1^{\leftarrow}(u_1), G_2^{\leftarrow}(u_2)) \\ &= G\left(\frac{u_1}{2 - u_1}, \frac{u_2}{2 - u_2}\right) \\ &= \frac{u_1 u_2}{u_1 - u_1 u_2 + u_2}. \end{aligned}$$

The limiting upper threshold copula of C and C_0 is the survival copula $\widehat{C}_G$. By (1.7)

$$\begin{aligned}\widehat{C}_G(u_1,u_2) &= C_G(1-u_1,1-u_2)+u_1+u_2-1\\ &= \frac{u_1u_2(2-u_1-u_2)}{1-u_1u_2}.\end{aligned}$$

○

The following interesting result may be used to find EV copulas.

Theorem 5.9. [4, Proposition 3.23*(ii)*]
Let C be an exchangeable, two-dimensional copula with upper tail dependence coefficient $\lambda > 0$ and let C_0 and C^{up} be its EV copula and limiting upper threshold copula, respectively. Additionally assume $C^{\mathrm{up}} \in \mathrm{CDA}(\widetilde{C}_0)$ with upper tail dependence coefficient $\tilde{\lambda}$. Then C_0 can be written in terms of the independence copula C^{ind} and $\widetilde{C}_0$:

$$C_0 = \left(C^{\mathrm{ind}}\right)^{1-c}\left(\widetilde{C}_0\right)^{c}, \quad \text{where } c = \frac{\lambda}{\tilde{\lambda}} \leq 1. \tag{5.15}$$

Remark 5.10. The inequality $c \leq 1$ follows from Remark 5.7. Note that $\tilde{\lambda}$ cannot be zero under the stated assumptions.

The next example shows how to apply Theorem 5.9 to find C_0.

Example 5.11. We determine the corresponding EV copulas for the copulas

$$C_\delta(u_1,u_2) = \delta\min\{u_1,u_2\} + (1-\delta)u_1u_2, \quad \delta \in (0,1], \tag{5.16}$$

without using one of the limit formulations in (4.7), (4.8) or (4.9). So let $C_\delta \in \mathrm{CDA}(C_{\delta,0})$. Instead of the direct calculation of $C_{\delta,0}$ we make use of Theorem 5.9. From Example 5.5 we know that

$$C^{\mathrm{up}}(u_1,u_2) = \min\{u_1,u_2\},$$

which is an EV copula and hence it is in its own CDA. With the notation of the above Theorem 5.9, $\widetilde{C}_0 = C^{\mathrm{up}}$. The copula $\widetilde{C}_0$ has tail dependence coefficient $\tilde{\lambda} = 1$. A straightforward calculation of the tail dependence coefficient λ_δ of C_δ via (3.3) yields

$$\lambda_\delta = \lim_{q\to 1^-}\frac{1-2q+\delta q+(1-\delta)q^2}{1-q} = \delta.$$

An application of Theorem 5.9 gives us

$$C_{\delta,0}(u_1,u_2) = (u_1u_2)^{1-\delta}\min\{u_1,u_2\}^{\delta}.$$

○

Example 5.11 brought together a variety of multivariate extreme value techniques:

- CDA theory,
- threshold copulas,
- EV copulas,
- and tail dependence coefficients.

Chapter 6

Practical Considerations

6.1 Block maxima approach

For block maxima models we group the d-dimensional observations in blocks of size n and concentrate on the componentwise maxima of these blocks. Blocking data makes particular sense if there are natural ways to determine the block size n.

In practical applications we are confronted with the following set-up. We have N observations from an unknown underlying d-dimensional df F, which is supposed to be in the MDA of some MEVD G. For ease of presentation we assume $N = m \cdot n$ for some $m, n \in \mathbb{N}$ and large n. Then the data can be divided into m blocks of size n. We denote the maximum of the ith block by $\boldsymbol{M}_{n,i}$. The distribution of $\boldsymbol{M}_{n,i}$ is approximated by the MEVD G. The multivariate extension of the Fisher–Tippett theorem suggests a model consisting of GEV margins coupled with an EV copula.

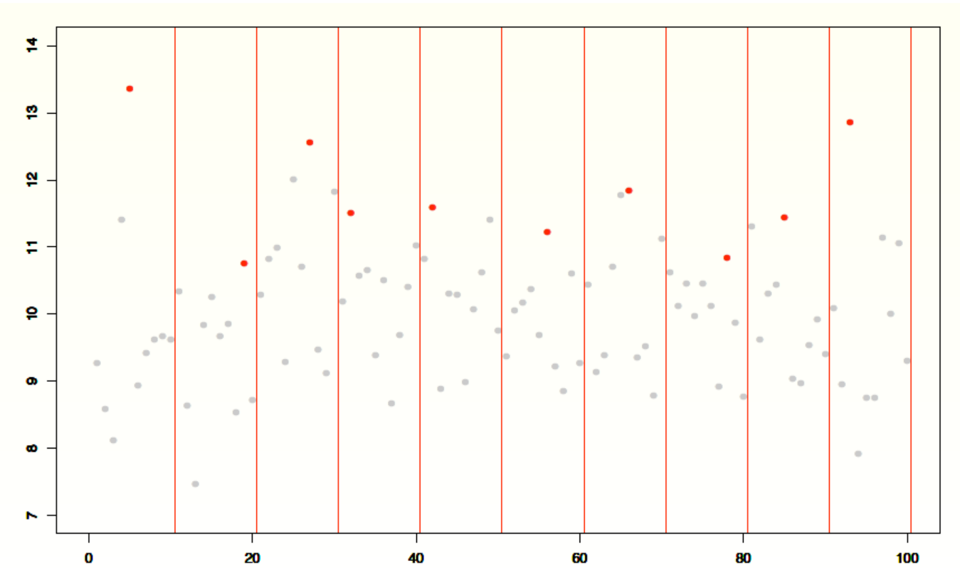

Figure 6.1: *Block maxima: the generalised extreme value distribution approach. Compare with the GPD approach in Figure* 5.1.

The choice of m and n can be a very difficult task. On the one hand n should be large so that the approximation by the generalised EVD $H_{\xi;\mu,\sigma}$ in the margin $i \in \{1, \ldots, d\}$ is acceptable, but on the other hand a large m, which is the number of block maxima we assume to come from $H_{\xi;\mu,\sigma}$, is helpful for reliable parameter estimation. A small n can lead to bias in estimation, whereas a large n leads to large estimation variance.

A general model for the tail of a d-dimensional df F could be

- a GPD based model as in (5.5) or
- a GEV based model as described above.

Both of these approaches rely on EV copulas. To successfully implement them, the availability of a wide variety of EV copulas is crucial.

6.2 Constructing new extreme value copulas

The three-parameter Galambos copulas from Example 5.1

$$C^{\mathrm{Gal}}_{\theta,\alpha,\beta}(u_1,u_2) = u_1 u_2 \exp\left\{\left((-\alpha\log u_1)^{-\theta} + (-\beta\log u_2)^{-\theta}\right)^{-1/\theta}\right\}, \tag{6.1}$$

where $\alpha,\beta\in[0,1]$ and $0<\theta<\infty$, form a reservoir of flexible EV copulas to choose from when setting up a parametric model for the tail. Another excellent choice is the asymmetric Gumbel copula given by

$$C^{\mathrm{Gu}}_{\theta,\alpha,\beta}(u_1,u_2) = u_1^{1-\alpha}u_2^{1-\beta}\exp\left\{\left((-\alpha\log u_1)^{\theta} + (-\beta\log u_2)^{\theta}\right)^{1/\theta}\right\}, \tag{6.2}$$

where $0\le\alpha,\beta\le 1$ and $1\le\theta<\infty$. These two families from [12, Chapter 7] can straightforwardly be generalised for the d-dimensional case. In the bivariate case a different way to obtain EV copulas is by using Corollary 4.13, which shows how a Pickands dependence function A defines an EV copula. It is natural to study parametric A-functions. As an example we look at a polynomial Pickands dependence function of the form

$$A(t) = a_0 + a_1 t + a_2 t^2 + a_3 t^3, \quad t\in[0,1].$$

In order to guarantee that $A(t)$ satisfies conditions (A1) and (A2) one needs some restrictions on the parameters a_0,a_1,a_2,a_3. Since $A(0)=1$ we immediately get $a_0=1$. After analysing (A1) and (A2) a bit more, it turns out that

$$A(t) = 1-(a_2+a_3)t + a_2t^2 + a_3t^3, \quad t\in[0,1], \tag{6.3}$$

with the parameter restrictions

$$a_2\ge 0, \quad a_2+3a_3\ge 0, \quad a_2+a_3\le 1, \quad a_2+2a_3\le 1$$

defines a valid Pickands dependence function. For details on the derivation see [1, p. 308].

This section provides various ways of gaining new EV copulas from known ones, such as (6.1), (6.2) or those described through (6.3) and Corollary 4.13. The presented material is based on [4, Chapter 4].

6.2.1 Product copulas

Proposition 6.1 (product EV copula). *Let $C_{01}, \dots, C_{0p}$ be d-dimensional copulas and let $A = (a_{ij})_{i=1,\dots,d}^{j=1,\dots,p} \in [0,1]^{d\times p}$ such that $\sum_{j=1}^p a_{ij} = 1$ for all $i \in \{1, \dots, d\}$. Then*

$$C_{0A}(\boldsymbol{u}) = \prod_{j=1}^{p} C_{0j}(u_1^{a_{1j}}, u_2^{a_{2j}}, \dots, u_d^{a_{dj}}), \quad \boldsymbol{u} \in [0,1]^d, \tag{6.4}$$

is a copula. If $C_{01}, \dots, C_{0p}$ are even EV copulas, then C_{0A} also is an EV copula.

Proof. First we show that C_{0A} is a copula. For this aim we find a vector $\boldsymbol{U} = (U_1, \dots, U_d)$ with standard uniform margins and df C_{0A}. Let the independent d-dimensional random vectors $\boldsymbol{Y}_j = (Y_{1,j}, \dots, Y_{d,j}) \sim C_{0j}$ for $j = 1, \dots, p$ and

$$U_i = \max_{j\in\{1,\dots,p\}} \left\{ Y_{i,j}^{1/a_{ij}} \right\}, \quad i = 1, \dots, d.$$

Then

$$\begin{aligned}
\mathbb{P}(\boldsymbol{U} \leq \boldsymbol{u}) &= \mathbb{P}\Big(Y_{1,j}^{1/a_{1j}} \leq u_1, \dots, Y_{d,j}^{1/a_{dj}} \leq u_d, \quad \forall j \in \{1, \dots, p\}\Big) \\
&= \mathbb{P}\Big(Y_{1,j} \leq u_1^{a_{1j}}, \dots, Y_{d,j} \leq u_d^{a_{dj}}, \quad \forall j \in \{1, \dots, p\}\Big) \\
&= \prod_{j=1}^{p} C_{0j}(u_1^{a_{1j}}, u_2^{a_{2j}}, \dots, u_d^{a_{dj}}), \quad \boldsymbol{u} \in [0,1]^d,
\end{aligned}$$

and therefore C_{0A} is a distribution function on $[0,1]^d$. By checking that for every $i \in \{1, \dots, d\}$ and $u_i \in [0,1]$

$$C_{0A}(1, \dots, 1, u_i, 1, \dots, 1) = \prod_{j=1}^{p} C_{0j}(1, \dots, 1, u_i^{a_{ij}}, 1, \dots, 1) = \prod_{j=1}^{p} u_i^{a_{ij}} = u_i,$$

we conclude that C_{0A} is a copula. If $C_{01}, \dots, C_{0p}$ are even EV copulas, then we have

$$\begin{aligned}
C_{0A}^t(\boldsymbol{u}^{1/t}) &= \prod_{j=1}^{p} C_{0j}^t(u_1^{a_{1j}/t}, u_2^{a_{2j}/t}, \dots, u_d^{a_{2j}/t}) \\
&= \prod_{j=1}^{p} C_{0j}(u_1^{a_{1j}}, u_2^{a_{2j}}, \dots, u_d^{a_{dj}}) = C_{0A}(\boldsymbol{u}), \quad \boldsymbol{u} \in [0,1]^d,
\end{aligned}$$

where we used property (4.6) of EV copulas for the second equality. Hence C_{0A} is an EV copula. □

Finding copulas in $\mathrm{CDA}(C_{0A})$ is very easy.

Corollary 6.2. If $C_j \in \mathrm{CDA}(C_{0j})$ for all $j \in \{1, \dots, p\}$ and $A \in [0,1]^{d\times p}$ such that $\sum_{j=1}^p a_{ij} = 1$ for all $i \in \{1, \dots, d\}$, then the copula

$$C_A(\boldsymbol{u}) = \prod_{j=1}^{p} C_j(u_1^{a_{1j}}, u_2^{a_{2j}}, \dots, u_d^{a_{dj}}), \quad \boldsymbol{u} \in [0,1]^d, \tag{6.5}$$

is in the CDA of the EV copula C_{0A} given in (6.4).

Proof. Since $C_j \in \mathrm{CDA}(C_{0j})$ for all $j \in \{1, \ldots, p\}$ we get for $\boldsymbol{u} \in [0,1]^d$ that

$$\begin{aligned}
\lim_{t\to\infty} C_A^t(\boldsymbol{u}^{1/t}) &= \lim_{t\to\infty} \prod_{j=1}^{p} C_j^t(u_1^{a_{1j}/t}, u_2^{a_{2j}/t}, \ldots, u_d^{a_{2j}/t}) \\
&= \prod_{j=1}^{p} \lim_{t\to\infty} C_j^t(u_1^{a_{1j}/t}, u_2^{a_{2j}/t}, \ldots, u_d^{a_{2j}/t}) \\
&= \prod_{j=1}^{p} C_{0j}(u_1^{a_{1j}}, u_2^{a_{2j}}, \ldots, u_d^{a_{dj}}) \\
&= C_{0A}(\boldsymbol{u}).
\end{aligned}$$

In other words, $C_A \in \mathrm{CDA}(C_{0A})$. □

6.2.2 Nested copulas

A **nested copula** can be written in terms of lower-dimensional copulas. For instance: assume C_1 and C_2 are two-dimensional copulas. If the function

$$C(u_1, u_2, u_3) := C_1(u_1, C_2(u_2, u_3)), \quad (u_1, u_2, u_3) \in [0,1]^3, \tag{6.6}$$

is a copula, then we call it a nested copula.

Remark 6.3. However, the function C in (6.6) is not necessarily a copula, as Example 3.30 in [13] shows.

Now let C_1 and C_2 be two-dimensional EV copulas such that C is a copula. Then by calculating

$$\begin{aligned}
C^t(u_1^{1/t}, u_2^{1/t}, u_3^{1/t}) &= C_1^t(u_1^{1/t}, C_2(u_2^{1/t}, u_3^{1/t})) \\
&= C_1^t(u_1^{1/t}, C_2^{1/t}(u_2, u_3)) \\
&= C_1(u_1, C_2(u_2, u_3)) \\
&= C(u_1, u_2, u_3)
\end{aligned}$$

one concludes that C is an EV copula as well. The construction of EV copulas via nested copulas can be generalised to d dimensions as the following result from [4] shows.

Proposition 6.4 (nested EV copula). *Let C_{01} be a two-dimensional EV copula and let C_{0d} be a d-dimensional EV copula such that*

$$C_{0,d+1}(\boldsymbol{u}) := C_{01}(u_1, C_{0d}(u_2, \ldots, u_{d+1})), \quad \boldsymbol{u} \in [0,1]^{d+1}, \tag{6.7}$$

defines a copula. Then $C_{0,d+1}$ is a $(d+1)$-dimensional EV copula.

6.3 Critical remarks and challenges

The implementation of EVT methodology faces many challenges. Sparse data, threshold selection, choosing the methods of parameter estimation and model checking keep EVT users busy. EVT can definitely not be regarded as an isolated science. The main challenge lies in the inhomogeneity of the data. Although there exists a range of different methods how to approach extremes or outliers, there is no universally accepted one. The method set up involves some subjectivity.

In the author's eyes, studying extreme value theory literature will inevitably lead to the following observation: The probability theory for extremes seems elegant, rigorous and extensive. The variety of different ideas is enormous. The statistical theory, however, is underdeveloped and primitive in many aspects. "Unfortunately, the current state of the art does not seem developed far enough to provide the user with a fully automatic, universally applicable methodology" [1].

Although many statistical procedures for extremes do exist, their outcomes have to be carefully checked and complemented with other methods to justify their results. EVT is no black-box machinery, which some practitioners would prefer.

MEVT allows for extrapolation outside the range of the data. We have considered extremes in all components simultaneously. The case where only some components are extreme is relatively unexplored in today's literature.

6.3.1 Copula estimation and testing for extremes

A crucial requirement when using block maxima or threshold exceedance models is copula estimation from data $\boldsymbol{x}_1, \dots, \boldsymbol{x}_n$, which are observations from an unknown d-dimensional df F with margins $F_1, \dots, F_d$ and copula C. Unfortunately one does not have observations directly from the copula C.

The margins can be estimated straightforwardly by the empirical dfs:

$$\widehat{F}_j(x) = \frac{1}{n+1} \sum_{i=1}^{n} \mathbb{1}_{[x_{i,j}\, ,\, \infty)}(x), \quad j = 1, \dots, d.$$

The empirical dfs can be used to transform the observations $\boldsymbol{x}_i$ into pseudo copula observations

$$\widehat{\boldsymbol{u}}_i = (\widehat{F}_1(x_{1,i}), \dots, \widehat{F}_d(x_{d,i})), \quad i = 1, \dots, n.$$

This construction, however, introduces dependence between the pseudo copula observations, thus violating the assumptions for a further maximum-likelihood estimation. Although the obtained maximum-likelihood estimates are still asymptotically normal, very little can be said about their finite-sample properties. That concern should be kept in mind.

Before fitting an EV copula to observations one might want to test the hypothesis

$$H_0 : \ C \text{ is an EV copula.}$$

The literature provides very few tests for H_0. In the bivariate case the best known test is based on the fact that for a two-dimensional random vector $\boldsymbol{X}$ with EV copula C_0 the random variable $W = C_0(F_1(X_1), F_2(X_2))$ satisfies

$$-1 + 8\,\mathbb{E}[W] - 9\,\mathbb{E}[W^2] = 0.$$

For details on the distribution of W consult [7]. For another class of tests we refer to [2, p. 36].

6.3.2 Violations of the maintained iid assumption

The most important assumption that we have maintained throughout this thesis is the iid assumption of the sequence of random vectors $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots$ which generate our data. However, in the majority of situations the assumption that the original data are iid is very unrealistic. A practical solution might be to form groups of data which can be assumed to be homogeneous. For example think of a data set comprising the daily maximum temperatures at a certain place. If there is a strong annual cycle in the data, they are not iid. An extraordinary hot summer day will have other characteristics than an extraordinary hot winter day. But for the daily maximum temperatures in a particular month homogeneity seems plausible.

Now what if the iid assumption does not hold? Think of a sea level estimation problem. Is there any sense in trying to predict what sea levels might occur in the future without knowledge of possible climate changes? In particular, we could be confronted with data exhibiting a trend. Detrending techniques, such as fitting a polynomial trend or differencing, might help us to filter out some iid residuals for which we could use our standard extreme value methods. Unfortunately the interpretation of results gained in such a way becomes more difficult since extremes in the detrended data do not necessarily correspond to extremes in the original data.

Roughly speaking, most environmental data sets have a more complex than an iid structure. However, that can hardly be astonishing when we think of the four seasons and their effect on our weather. Also it is widely agreed that high-frequency financial asset returns are conditionally heteroskedastic, and thus not iid.

What can happen if we turn down the iid assumption? Basically everything, as the following consideration shows. For simplicity we concentrate in the remainder of this subsection on the univariate case. Let $X_1, X_2, \ldots$ be a non-iid sequence of random variables and X be a random variable with df F. Now assume $X_i = X$ for all $i \in \mathbb{N}$. Then

$$\mathbb{P}(M_n \le x) = \mathbb{P}(X \le x) = F(x), \quad x \in \mathbb{R}.$$

Since any distribution can be a limit distribution for maxima, a general characterisation of the behaviour of extremes is impossible. What we can do is comparing the distribution of the maxima of a stationary series (stationarity implies in particular that all marginal distributions are the same) and of a series of iid random variables having the same marginal distribution. Such a series of iid random variables is called an associated iid series. If there is any difference in the limiting distributions of maxima for those two series, it can only be caused by the dependence in the original series.

The next example, which appears in many variations in extreme value literature, is a nice illustration of this argument.

Example 6.5 (dependence structure). Based on [3, Example 5.1]. Let $Y_0, Y_1, Y_2, \ldots$ be an iid sequence of random variables with df

$$F(y) = \exp\Big\{ -\frac{1}{(a+1)y} \Big\}, \quad y > 0,\ a \in [0,1].$$

Define the sequence X_i by

$$X_0 := Y_0, \qquad X_i := \max\{aY_{i-1}, Y_i\}, \quad i \in \mathbb{N}.$$

The marginal distribution of the $X_i, i \in \mathbb{N}$, is

$$\begin{aligned}\mathbb{P}(X_i \le x) &= \mathbb{P}(aY_{i-1} \le x, Y_i \le x) \\ &= \exp\Big\{-\frac{a}{(a+1)x}\Big\}\exp\Big\{-\frac{1}{(a+1)x}\Big\}\mathbb{1}_{(0,\infty)}(x) \\ &= \exp\left\{-x^{-1}\right\}\mathbb{1}_{(0,\infty)}(x) = \Phi_1(x), \quad x \in \mathbb{R}.\end{aligned}$$

Now, let $X_1^\star, X_2^\star, \ldots$ be the associated iid sequence (i.e. a sequence of independent, standard Fréchet random variables) and for $n \in \mathbb{N}$ define $M_n^\star = \max\{X_1^\star, \ldots, X_n^\star\}$. By direct calculation (see Example 2.4) it follows that

$$\mathbb{P}(M_n^\star \le nx) = \exp\left\{-x^{-1}\right\}, \quad n \in \mathbb{N}, x > 0.$$

On the other hand, for $M_n = \max\{X_1, \ldots, X_n\}$:

$$\begin{aligned}\mathbb{P}(M_n \le nx) &= \mathbb{P}(X_1 \le nx, \ldots, X_n \le nx) \\ &= \mathbb{P}(aY_0 \le nx, Y_1 \le nx, \ldots, aY_{n-1} \le nx, Y_n \le nx) \\ &= \mathbb{P}(aY_0 \le nx)\ \mathbb{P}(Y_1 \le nx, \ldots, Y_n \le nx) \\ &= \exp\Big\{-\frac{a}{(a+1)nx}\Big\}\exp\Big\{-\frac{n}{(a+1)nx}\Big\} \\ &\overset{n\to\infty}{\longrightarrow} \left[\exp\left\{-x^{-1}\right\}\right]^{\frac{1}{a+1}}, \quad x > 0.\end{aligned}$$

If $a \ne 0$, the limit distributions for M_n and $M_n^\star$ are not the same. Whereas if $a = 0$, the original sequence $X_i, i \in \mathbb{N}_0$, is iid. ∘

Remark 6.6. At the moment there is room for improvement of the theory on non-stationary extremes, especially in the multivariate case. In the near future, however, the chances for significant advances in the multivariate case are slim because there are still many unsolved issues in the iid setting.

6.3.3 Applications and outlook

Financial risk management focuses on extreme quantiles and tails. The probability of surviving big hits, i.e. extreme events, leads to credit ratings and regulatory capital requirements. Most applications of EVT are inherently multivariate. Taking into account the dependence between different asset prices is crucial for an adequate risk assessment.

Many users of statistical software are interested in confidence intervals. In practice one rarely sees confidence intervals for quantities related to extreme events. The reason being that an EV analysis is prone to a lot of uncertainty due to the limited amount of data. This uncertainty would often lead to very big confidence intervals which are useless for practical applications. A subjective explanation is that EVT is just a scientific approach to pure guesswork and cannot do magic, meaning that it can be dangerous to ask for too much.

Many applications of EVT are found in hydrology, the cradle of modern EVT. Approximately 10 percent of the human population live at the coast. Accordingly the occurrence of extreme sea levels along those low-lying and highly populated coastlines can lead to considerable loss of life and billions of Euros of damage to coastal infrastructure. There is also now a growing concern about rising mean sea levels. Over the last 150 years, global mean sea levels have on average risen by about 25 cm and it is predicted that this rise will continue over the twenty-first century at an accelerated rate. However, it is very difficult to predict future extreme sea levels. With rises in mean sea level, a particular water level will be exceeded more and more frequently as progressively less severe storm conditions are required to achieve that level. Some governments, for instance, demand dykes to be built high enough to withstand the 500-year return level of the sea level. That is a level which is expected to be reached only once in the next 500 years. If the observation period is a mere say 40 years, then the data sample is terribly small in relation to the request. Nevertheless one has to come up with a reliable estimate to minimise the risk of catastrophic structural failures due to under-design or expensive wastes due to over-design.

When it comes to extremes we are not hopeless. Although

- the central iid assumption may not be satisfied, and
- in practice it is virtually impossible to check if the copula of given data is in the CDA of some EV copula (we have presented pathological counterexamples in Chapter 4),

EVT is here to stay and it provides an excellent complement to alternatives such as graphical data analysis. Today EVT is a respected and widely used scientific discipline with its own limitations and strengths. Applications stretch from insurance and finance (cf. [5]), material sciences (cf. [3]) to the modelling of natural hazards (cf. [15]) and many more.

Appendix A

Univariate Extreme Value Theory

This section provides a brief summary of univariate extreme value theory (UEVT). One considers probabilities of the form

$$\mathbb{P}\left(\frac{M_n - b_n}{a_n} \leq x\right) = \mathbb{P}\left(M_n \leq a_n x + b_n\right) = F^n(a_n x + b_n), \quad x \in \mathbb{R},$$

where $M_n := \max\{X_1, \ldots, X_n\}$ for a sequence $X_1, X_2, \ldots$ of independent identically distributed (iid) non-degenerate one-dimensional random variables with distribution function (df)F and normalising sequences $a_n > 0$ and $b_n \in \mathbb{R}$ for $n \in \mathbb{N}$.

Suppose there exist sequences $a_n > 0$, $b_n \in \mathbb{R}$ and a non-degenerate df H such that

$$\lim_{n\to\infty} F^n(a_n x + b_n) = H(x)$$

for every continuity point $x \in \mathbb{R}$ of H, then H is called an **extreme value distribution** and F is said to be in the **maximum domain of attraction** of H. We use the notation $F \in \text{MDA}(H)$.

The two central results in EVT are the Fisher–Tippett theorem, which describes all possible limit distributions for maxima, and the Pickands–Balkema–de Haan theorem, which essentially says that for sufficiently large thresholds the exceedances follow the generalised Pareto distribution (GPD).

Theorem A.1 (Fisher–Tippett theorem). [5, Theorem 3.2.3]
Let (X_n) be a sequence of iid random variables. If there exist normalising sequences $a_n > 0$, $b_n \in \mathbb{R}$ and some non-degenerate df H such that

$$\frac{M_n - b_n}{a_n} \quad \xrightarrow{d} \quad H, \tag{A.1}$$

then H is of the type of one of the following three dfs.

Fréchet: $$\Phi_\alpha(x) = \begin{cases} 0 & \text{if } x \le 0 \\ \exp\{-x^{-\alpha}\} & \text{if } x > 0 \end{cases} \qquad \alpha > 0.$$

Weibull: $$\Psi_\alpha(x) = \begin{cases} \exp\{-(-x)^{\alpha}\} & \text{if } x \le 0 \\ 1 & \text{if } x > 0 \end{cases} \qquad \alpha > 0.$$

Gumbel: $$\Lambda(x) = \exp\{-\,\mathrm{e}^{-x}\} \qquad x \in \mathbb{R}.$$

Φ_α, Ψ_α and Λ are called (standard) extreme value distributions.

The generalised extreme value distribution is a one-parameter representation of the three extreme value distributions.

Definition A.2 (generalised extreme value (GEV) distribution). [12, Definition 7.1]
The df of the (standard) GEV distribution is given by

$$H_\xi(x) = \begin{cases} \exp\left\{-(1+\xi x)^{-1/\xi}\right\} & \text{for } \xi \in \mathbb{R}\backslash\{0\}, \\ \exp\left\{-\,\mathrm{e}^{-x}\right\} & \text{for } \xi = 0, \end{cases} \tag{A.2}$$

for all $x \in \mathbb{R}$ satisfying $1 + \xi x > 0$. A three parameter family is obtained by defining $H_{\xi;\mu,\sigma}(x) := H_\xi(\frac{x-\mu}{\sigma})$ for a location parameter $\mu \in \mathbb{R}$ and a scale parameter $\sigma > 0$.

To model the excess distribution

$$F_u(x) := \mathbb{P}(X - u \le x \,|\, X > u)$$

over a threshold u for $0 \le x < x_F - u$, one uses the generalised Pareto distribution. Here $x_F = \sup\{x \in \mathbb{R} \,:\, F(x) < 1\}$ denotes the right endpoint of F.

Definition A.3 (GPD). [12, Definition 7.16]
The df of the generalised Pareto distribution (GPD) is given by

$$G_{\xi,\beta}(x) = \begin{cases} 1 - \left(1 + \frac{\xi x}{\beta}\right)^{-1/\xi} & \text{for } \xi \ne 0\,, \\ 1 - \exp\left\{-\frac{x}{\beta}\right\} & \text{for } \xi = 0\,, \end{cases} \tag{A.3}$$

where $\beta > 0$, and $x \ge 0$ if $\xi \ge 0$ and $x \in [0, {-\beta}/{\xi}]$ if $\xi < 0$. The parameter ξ is called shape parameter.

Theorem A.4 (Pickands–Balkema–de Haan). [5, Theorem 3.4.13(b)]
For every $\xi \in \mathbb{R}$, $F \in MDA(H_\xi)$ if and only if

$$\lim_{u \nearrow x_F} \sup_{0<x<x_F-u} \left| F_u(x) - G_{\xi,\beta(u)}(x) \right| = 0 \tag{A.4}$$

for some positive function $\beta(\cdot)$.

Bibliography

[1] J. Beirlant, Y. Goegebeur, J. Teugels, and J. Segers, *Statistics of Extremes: Theory and Applications*, Wiley Series in Probability and Statistics, John Wiley & Sons Ltd., Chichester, 2004. (Cited on pages 11, 12, 43, 50, and 53.)

[2] A. Bücher, *Statistical Inference for Copulas and Extremes*, PhD thesis, Ruhr-Universität Bochum, 2011. (Cited on page 53.)

[3] S. Coles, *An Introduction to Statistical Modeling of Extreme Values*, Springer Series in Statistics, Springer, London, 2001. (Cited on pages ix, 5, 14, 19, 54, and 56.)

[4] S. Demarta, *The Copula Approach to Modelling Multivariate Extreme Values*, PhD thesis, ETH Zürich, 2007. (Cited on pages 17, 20, 25, 26, 30, 43, 46, 47, 50, and 52.)

[5] P. Embrechts, C. Klüppelberg, and T. Mikosch, *Modelling Extremal Events for Insurance and Finance*, vol. 33 of Applications of Mathematics (New York), Springer, Berlin, 1997. (Cited on pages 31, 38, 43, 56, 57, and 58.)

[6] A.-L. Fougères, *Multivariate Extremes*, In Extreme Values in Finance, Telecommunications, and the Environment, Chapman and Hall/CRC, 2004. (Cited on page 37.)

[7] K. Ghoudi, A. Khoudraji, and L.-P. Rivest, *Propriétés statistiques des copules de valeurs extrêmes bidimensionnelles*, Canad. J. Statist., 26 (1998), pp. 187–197. (Cited on page 53.)

[8] G. Gudendorf and J. Segers, *Extreme-value copulas*, arXiv:0911.1015v2 [math.ST], (2009). (Cited on pages 13, 14, 15, 21, 25, and 29.)

[9] ——, *Nonparametric estimation of multivariate extreme-value copulas*, J. Statist. Plann. Inference, 142 (2012), pp. 3073–3085. (Cited on page 14.)

[10] H. Hult and F. Lindskog, *On Kesten's counterexample to the Cramér-Wold device for regular variation*, Bernoulli, 12 (2006), pp. 133–142. (Cited on page 38.)

[11] D. Kortschak and H. Albrecher, *Asymptotic results for the sum of dependent non-identically distributed random variables*, Methodology and Computing in Applied Probability, 11 (2009), pp. 279–306. (Cited on pages 32 and 34.)

[12] A. J. McNeil, R. Frey, and P. Embrechts, *Quantitative Risk Management*, Princeton Series in Finance, Princeton University Press, Princeton, NJ, 2005. (Cited on pages 1, 2, 19, 25, 42, 43, 46, 50, and 58.)

[13] R. B. Nelsen, *An Introduction to Copulas*, Springer Series in Statistics, Springer, New York, second ed., 2006. (Cited on pages 2, 25, 31, and 52.)

[14] S. I. Resnick, *Extreme Values, Regular Variation and Point Processes*, Springer Series in Operations Research and Financial Engineering, Springer, New York, 2008. Reprint of the 1987 original. (Cited on pages x, 7, 8, 9, 10, 27, and 30.)

[15] G. Salvadori, C. de Michele, N. T. Kottegoda, and R. Rosso, *Extremes in Nature: An Approach Using Copulas*, Springer Series in Water Science and Technology, Springer, 2007. (Cited on page 56.)