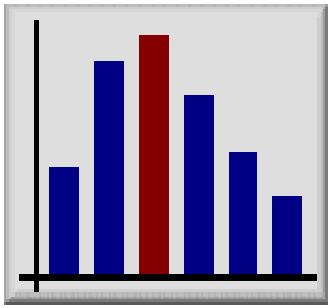
2015 CRoNoS Winter Course:



An Offspring of Multivariate Extreme Value Theory: D-Norms

Michael Falk

University of Wuerzburg, Germany

London, December 2015

'We do not want to calculate,
we want to reveal structures.'

- David Hilbert, 1930 -

Preface

Multivariate extreme value theory (MEVT) is the proper toolbox for analyzing several extremal events simultaneously. Its practical relevance in particular for risk assessment is, consequently, obvious. But on the other hand MEVT is by no means easy to access; its key results are formulated in a measure theoretic setup, a fils rouge is not visible.

Writing the 'angular measure' in MEVT in terms of a random vector, however, provides the missing fils rouge: Every result in MEVT, every relevant probability distribution, be it a max-stable one or a generalized Pareto distribution, every relevant copula, every tail dependence coefficient etc. can be formulated using a particular kind of norm on multivariate Euclidean space, called D-norm. Norms are introduced in each course on mathematics as soon as the multivariate Euclidean space is introduced. The definition of an arbitrary D-norm requires only the additional knowledge on random variables and their expectations. But D-norms do not only constitute the fils rouge through MEVT, they are of mathematical interest of their own.

In Sessions 1 to 3 we provide in the introductory chapter the theory of D-norms in

detail. The second chapter introduces multivariate generalized Pareto distributions and max-stable distributions via D-norms. The third chapter provides the extension of D-norms to functional spaces and, thus, deals with generalized Pareto processes and max-stable processes.
Session 4, in addition to a brief summary of univariate EVT and D-norms, provides a relaxed tour through the essentials of MEVT, due to the D-norms approach. Quite recent results on multivariate records complete this text.

Contents

Chapter 1

Introduction

1.1 Norms and D-Norms

General Definition of a Norm

Definition 1.1.1. A function $f : \mathbb{R}^d \to [0, \infty)$ is a *norm*, if it satisfies for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^d, \lambda \in \mathbb{R}$

$$f(\boldsymbol{x}) = 0 \iff \boldsymbol{x} = \boldsymbol{0} \in \mathbb{R}^d, \tag{1.1}$$

$$f(\lambda \boldsymbol{x}) = |\lambda| \, f(\boldsymbol{x}), \tag{1.2}$$

$$f(\boldsymbol{x} + \boldsymbol{y}) \leq f(\boldsymbol{x}) + f(\boldsymbol{y}). \tag{1.3}$$

Condition (1.2) is called homogeneity and condition (1.3) is called triangle inequality or Δ-inequality, for short.

A norm $f : \mathbb{R}^d \to [0, \infty)$ is typically denoted by

$$\|\boldsymbol{x}\| = f(\boldsymbol{x}), \qquad \boldsymbol{x} \in \mathbb{R}^d. \tag{1.4}$$

Each norm on $\mathbb{R}^d$ defines a distance, or metric on $\mathbb{R}^d$ via

$$d(x, y) = \|\boldsymbol{x} - \boldsymbol{y}\|, \qquad \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^d. \tag{1.5}$$

Well known examples of norms are the sup-norm

$$\|\boldsymbol{x}\|_\infty := \max_{1 \le i \le d} |x_i| \tag{1.6}$$

and the L_1-norm

$$\|\boldsymbol{x}\|_1 := \sum_{i=1}^{d} |x_i|, \qquad \boldsymbol{x} = (x_1, \ldots, x_d) \in \mathbb{R}^d. \tag{1.7}$$

The Logistic Norm

Not that obvious is the logistic family

$$\|x\|_p := \left(\sum_{i=1}^{d} |x_i|^p \right)^{1/p}, \qquad 1 \le p < \infty. \tag{1.8}$$

The corresponding Δ-inequality

$$f(\boldsymbol{x} + \boldsymbol{y}) \le f(\boldsymbol{x}) + f(\boldsymbol{y})$$

is known as the Minkowski-inequality[1].

[1] cf. Rudin (1976, Proposition 3.5)

Lemma 1.1.1. We have for $1 \le p \le q \le \infty$ and $\boldsymbol{x} \in \mathbb{R}^d$

(i) $\|\boldsymbol{x}\|_p \ge \|\boldsymbol{x}\|_q$,

(ii) $\lim\limits_{p\to\infty} \|\boldsymbol{x}\|_p = \|\boldsymbol{x}\|_\infty$.

Proof. (i) The inequality is obvious for $q = \infty$: $\|\boldsymbol{x}\|_\infty \le \left(\sum_{i=1}^d |x_i|^q\right)^{1/q}$.

Now consider $1 \le p \le q < \infty$ and choose $\boldsymbol{x} \ne \boldsymbol{0} \in \mathbb{R}^d$. Put $S := \|\boldsymbol{x}\|_p$. Then we have

$$\left\|\frac{\boldsymbol{x}}{S}\right\|_p = 1$$

and we have to establish

$$\left\|\frac{\boldsymbol{x}}{S}\right\|_q \le 1.$$

As

$$\frac{|x_i|}{S} \in [0, 1]$$

and thus

$$\left(\frac{|x_i|}{S}\right)^q \le \left(\frac{|x_i|}{S}\right)^p, \qquad 1 \le i \le d,$$

we obtain

$$\left\|\frac{\boldsymbol{x}}{S}\right\|_q = \left(\sum_{i=1}^{d}\left(\frac{|x_i|}{S}\right)^q\right)^{\frac{1}{q}} \leq \left(\sum_{i=1}^{d}\left(\frac{|x_i|}{S}\right)^p\right)^{\frac{1}{q}} = \left(\left\|\frac{\boldsymbol{x}}{S}\right\|_p\right)^{\frac{p}{q}} = 1^{\frac{p}{q}} = 1,$$

which is *(i)*.

(ii) We have, moreover, for $\boldsymbol{x} \neq \mathbf{0} \in \mathbb{R}^d$ and $p \in [1, \infty)$

$$\|\boldsymbol{x}\|_\infty \leq \|\boldsymbol{x}\|_p = \left(\sum_{i=1}^{d}\left(\frac{|x_i|}{\|\boldsymbol{x}\|_\infty}\right)^p\right)^{\frac{1}{p}} \|\boldsymbol{x}\|_\infty \leq d^{\frac{1}{p}} \|\boldsymbol{x}\|_\infty \xrightarrow[p\to\infty]{} \|\boldsymbol{x}\|_\infty ,$$

which implies *(ii)*. $\square$

Norms by Quadratic Forms

Let $A = (a_{ij})_{1\leq i,j\leq d}$ be a positive definite $d \times d$-matrix, i.e., the matrix A is symmetric, $A = A^\mathsf{T} = (a_{ji})_{1\leq i,j\leq d}$, and satisfies

$$\boldsymbol{x}^\mathsf{T} A\boldsymbol{x} = \sum_{1\leq i,j\leq d} x_i a_{ij} x_j > 0, \qquad \boldsymbol{x} \in \mathbb{R}^d,\ \boldsymbol{x} \neq \mathbf{0} \in \mathbb{R}^d.$$

Then

$$\|\boldsymbol{x}\|_A := \left(\boldsymbol{x}^T A\boldsymbol{x}\right)^{\frac{1}{2}}, \qquad x \in \mathbb{R}^d,$$

defines a norm on $\mathbb{R}^d$.

With $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ **we obtain, for example,** $\|\boldsymbol{x}\|_A = (x_1^2 + x_2^2)^{1/2} = \|\boldsymbol{x}\|_2$**.**

Conditions (1.1) **and** (1.2) **are obviously satisfied. The** Δ**-inequality follows by means of the Cauchy-Schwarz inequality**[2]

$$(\boldsymbol{x}^\intercal A \boldsymbol{y})^2 \le (\boldsymbol{x}^\intercal A \boldsymbol{x})\,(\boldsymbol{y}^\intercal A \boldsymbol{y})\,, \qquad \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^d,$$

as follows:

$$\begin{aligned}
\|\boldsymbol{x} + \boldsymbol{y}\|_A^2 &= (\boldsymbol{x} + \boldsymbol{y})^T A\,(\boldsymbol{x} + \boldsymbol{y}) \\
&= \boldsymbol{x}^T A \boldsymbol{x} + \boldsymbol{y}^T A \boldsymbol{x} + \boldsymbol{x}^T A \boldsymbol{y} + \boldsymbol{y}^T A \boldsymbol{y} \\
&\le \boldsymbol{x}^T A \boldsymbol{x} + 2(\boldsymbol{x}^T A \boldsymbol{x})^{\frac{1}{2}} (\boldsymbol{y}^T A \boldsymbol{y})^{\frac{1}{2}} + \boldsymbol{y}^T A \boldsymbol{y} \\
&= \left((\boldsymbol{x}^T A \boldsymbol{x})^{\frac{1}{2}} + (\boldsymbol{y}^T A \boldsymbol{y})^{\frac{1}{2}} \right)^2.
\end{aligned}$$

Definition of D-Norms

Let now $\boldsymbol{Z} = (Z_1, \dots, Z_d)$ **be a random vector (rv), whose components satisfy**

$$Z_i \ge 0, \quad E(Z_i) = 1, \qquad 1 \le i \le d.$$

Then

$$\|\boldsymbol{x}\|_D := E\left(\max_{1 \le i \le d} (|x_i|\, Z_i) \right), \qquad \boldsymbol{x} \in \mathbb{R}^d,$$

defines a norm, called D-norm and $\boldsymbol{Z}$ **is called generator of** $\|\boldsymbol{x}\|_D$**.**

[2] cf. Rudin (1976)

The homogeneity condition (1.2) **is obviously satisfied. Further, we have the bounds**

$$\begin{aligned}
\|\boldsymbol{x}\|_\infty &= \max_{1\le i\le d} |x_i| \\
&= \max_{1\le i\le d} E\left(|x_i|\, Z_i\right) \\
&\le E\left(\max_{1\le i\le d}(|x_i|\, Z_i)\right) \\
&\le E\left(\sum_{i=1}^{d} |x_i|\, Z_i\right) \\
&= \sum_{i=1}^{d} |x_i|\, E(Z_i) \\
&= \|\boldsymbol{x}\|_1\,, \qquad x\in\mathbb{R}^d,
\end{aligned}$$

i.e.,

$$\|\boldsymbol{x}\|_\infty \le \|\boldsymbol{x}\|_D \le \|\boldsymbol{x}\|_1\,, \qquad x\in\mathbb{R}^d. \tag{1.9}$$

This implies condition (1.1)**. The** $\Delta-$**inequality is easily seen by**

$$\begin{aligned}
\|\boldsymbol{x}+\boldsymbol{y}\|_D &= E\left(\max_{1\le i\le d}(|x_i+y_i|\, Z_i)\right) \\
&\le E\left(\max_{1\le i\le d}((|x_i|+|y_i|)Z_i)\right) \\
&\le E\left(\max_{1\le i\le d}(|x_i|\, Z_i)+\max_{1\le i\le d}(|y_i|\, Z_i)\right)
\end{aligned}$$

$$
\begin{aligned}
&= E\left(\max_{1\le i\le d}(|x_i|\, Z_i)\right) + E\left(\max_{1\le i\le d}(|y_i|\, Z_i)\right) \\
&= \|\boldsymbol{x}\|_D + \|\boldsymbol{y}\|_D .
\end{aligned}
$$

Basic Properties of D-Norms

Denote by $\boldsymbol{e}_j = (0,\dots,0,1,0,\dots,0) \in \mathbb{R}^d$ the j-th unit vector in $\mathbb{R}^d$, $1 \le j \le d$. Each D-norm satisfies

$$
\|\boldsymbol{e}_j\|_D = E\left(\max_{1\le i\le d}(\delta_{ij} Z_i)\right) = E(Z_j) = 1,
$$

where $\delta_{ij} = 1$ if $i = j$ and zero elsewhere, i.e., each D-norm is standardized.

Each D-norm is monotone, i.e., we have for $0 \le \boldsymbol{x} \le \boldsymbol{y}$, where this inequality is taken componentwise,

$$
\|\boldsymbol{x}\|_D = E\left(\max_{1\le i\le d}(x_i Z_i)\right) \le E\left(\max_{1\le i\le d}(y_i Z_i)\right) = \|\boldsymbol{y}\|_D .
$$

There are norms that are not monotone: Choose for example

$$
A = \begin{pmatrix} 1 & \delta \\ \delta & 1 \end{pmatrix}
$$

with $\delta \in (-1,0)$. The matrix A is positive definite, the norm $\|\boldsymbol{x}\|_A = (x^T A x)^{\frac{1}{2}} = (x_1^2 + 2\delta x_1 x_2 + x_2^2)^{\frac{1}{2}}$ is not monotone; just compare (x_1, x_2) with $(x_1, x_2 + \varepsilon)$.

Each D-norm is obviously radial symmetric, i.e., changing the sign of arbitrary components of $\boldsymbol{x} \in \mathbb{R}^d$ does not alter the value of $\|\boldsymbol{x}\|_D$. This means that the values of a

D-norm are completely determined by its values on the subset $\{\boldsymbol{x} \in \mathbb{R}^d : \boldsymbol{x} \geq \mathbf{0}\}$. The above norm $\|\cdot\|_A$ does not have this property.

1.2 Examples of D-Norms

The two Extremal D-Norms

Choose the constant generator $\boldsymbol{Z} := (1, 1, \dots, 1)$. Then

$$\begin{aligned}\|\boldsymbol{x}\|_D &= E\left(\max_{1\leq i\leq d}(|x_i|\, Z_i)\right)\\ &= E\left(\max_{1\leq i\leq d}(|x_i|)\right) = \|\boldsymbol{x}\|_\infty\,,\end{aligned}$$

i.e., the sup-norm is a D-norm.

Let $X \geq 0$ be a rv with $E(X) = 1$ and put $\boldsymbol{Z} := (X, X, \dots, X)$. Then $\boldsymbol{Z}$ is a generator of the D-norm

$$\begin{aligned}\|\boldsymbol{x}\|_D &= E(\max_{1\leq i\leq d}(|x_i|\, Z_i))\\ &= E(\max_{1\leq i\leq d}(|x_i|\, X))\\ &= \max_{1\leq i\leq d}(|x_i|)E(X)\\ &= \|\boldsymbol{x}\|_\infty\, E(X)\\ &= \|\boldsymbol{x}\|_\infty\,.\end{aligned}$$

This example shows that the generator of a D-norm is in general not uniquely determined, even its distribution is not.

Let now $\boldsymbol{Z}$ be a random permutation of $(d, 0, \dots, 0) \in \mathbb{R}^d$ with equal probability $1/d$, i.e.,

$$Z_i = \begin{cases} d, & \textbf{with probability } 1/d \\ 0, & \textbf{with probability } 1 - 1/d \end{cases}, \qquad 1 \le i \le d,$$

and $Z_1 + \cdots + Z_d = d$.

The rv $\boldsymbol{Z}$ is consequently the generator of a D-norm:

$$\begin{aligned}
\|\boldsymbol{x}\|_D &= E\left(\max_{1\le i\le d}(|x_i|\, Z_i)\right)\\
&= E\left(\max_{1\le i\le d}(|x_i|\, Z_i)\sum_{j=1}^d 1_{\{Z_j=d\}}\right)\\
&= E\left(\sum_{j=1}^d \max_{1\le i\le d}(|x_i|\, Z_i)\, 1_{\{Z_j=d\}}\right)\\
&= E\left(\sum_{j=1}^d |x_j|\; d\, 1_{\{Z_j=d\}}\right)\\
&= \sum_{j=1}^d |x_j|\; d\, E\left(1_{\{Z_j=d\}}\right)
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{j=1}^{d} |x_j| \; d\, P(Z_j = d) \\
&= \sum_{j=1}^{d} |x_j| \\
&= \|\boldsymbol{x}\|_1 ,
\end{aligned}
$$

i.e., $\|\cdot\|_1$ is a D-norm as well.

Inequality (1.9) **shows that the sup-norm $\|\cdot\|_\infty$ is the smallest D-norm and that the L_1-norm $\|\cdot\|_1$ is the largest D-norm.**

Each Logistic Norm is a D-Norm

Proposition 1.2.1. Each logistic norm $\|\boldsymbol{x}\|_p = \left(\sum_{i=1}^{d} |x_i|^p\right)^{1/p}$, $1 \leq p < \infty$, is a D-norm. For $1 < p < \infty$ a generator is given by

$$
\boldsymbol{Z} = (Z_1, \dots, Z_d) = \left(\frac{X_1}{\Gamma(1-p^{-1})}, \dots, \frac{X_d}{\Gamma(1-p^{-1})} \right),
$$

where $X_1, \dots, X_d$ are independent and identically (iid) *Fréchet-distributed* rv, i.e.,

$$
P(X_i \leq x) = \exp(-x^{-p}), \qquad x > 0,\ i = 1, \dots, d,
$$

with $E(X_i) = \Gamma(1-p^{-1})$, $1 \leq i \leq d$.

$\Gamma(s) = \int_0^\infty t^{s-1} \exp(-t)\, dt$**,** $s > 0$**, denotes the Gamma function.**

Proof. Put for notational convenience $\mu := E(X_1) = \Gamma(1 - p^{-1})$. From the fact that the expectation of a non-negative rv X is in general given by $\int_0^\infty P(X > t)\, dt$ (use Fubini's theorem), we obtain

$$\begin{aligned}
E\left(\max_{1\le i\le d} |x_i|\, Z_i\right) &= \int_0^\infty P\left(\max_{1\le i\le d} |x_i|\, Z_i > t\right)\, dt \\
&= \int_0^\infty 1 - P\left(\max_{1\le i\le d} |x_i|\, Z_i \le t\right)\, dt \\
&= \int_0^\infty 1 - P\left(Z_i \le \frac{t}{|x_i|}, 1 \le i \le d\right)\, dt \\
&= \int_0^\infty 1 - \prod_{i=1}^d P\left(Z_i \le \frac{t}{|x_i|}\right)\, dt \\
&= \int_0^\infty 1 - \prod_{i=1}^d \exp\left(-\left(\frac{|x_i|}{t\mu}\right)^p\right)\, dt \\
&= \int_0^\infty 1 - \exp\left(-\frac{\sum_{i=1}^d |x_i|^p}{(t\mu)^p}\right)\, dt
\end{aligned}$$

The substitution $t \mapsto t\left(\sum_{i=1}^d |x_i|^p\right)^{\frac{1}{p}}/\mu$ now implies that the integral above equals

$$\frac{\left(\sum_{i=1}^d |x_i|^p\right)^{\frac{1}{p}}}{\mu} \int_0^\infty 1 - \exp\left(-\frac{1}{t^p}\right)\, dt = \frac{\|\boldsymbol{x}\|_p}{E(X_1)} \int_0^\infty P(X_1 > t)\, dt$$

$$
\begin{aligned}
&= \frac{\|\boldsymbol{x}\|_p}{E(X_1)} E(X_1) \\
&= \|\boldsymbol{x}\|_p \,.
\end{aligned}
$$

$\square$

1.3 Takahashi's Characterizations

Takahashi's Characterizations of $\|\cdot\|_\infty$ and $\|\cdot\|_1$

Theorem 1.3.1 (Takahashi (1988)). Let $\|\cdot\|_D$ be an arbitrary D-norm on $\mathbb{R}^d$. Then we have the equivalences

$$
\begin{aligned}
\|\cdot\|_D = \|\cdot\|_1 &\iff \exists\, \boldsymbol{y} > \boldsymbol{0} \in \mathbb{R}^d : \; \|\boldsymbol{y}\|_D = \|\boldsymbol{y}\|_1 \,, \\
\|\cdot\|_D = \|\cdot\|_\infty &\iff \|\boldsymbol{1}\|_D = 1
\end{aligned}
$$

Corollary 1.3.1. We have for an arbitrary D-norm $\|\cdot\|_D$ on $\mathbb{R}^d$

$$
\|\cdot\|_D = \begin{cases} \|\cdot\|_\infty \\ \|\cdot\|_1 \end{cases} \iff \|\boldsymbol{1}\|_D = \begin{cases} 1 \\ d \end{cases} .
$$

Proof. To prove Theorem 1.3.1 we only have to show the implication "$\Leftarrow$". Let $(Z_1, \dots, Z_d)$ be a generator of $\|\cdot\|_D$.

(i) Suppose we have $\|\boldsymbol{y}\|_D = \|\boldsymbol{y}\|_1$ for some $\boldsymbol{y} > \mathbf{0} \in \mathbb{R}^d$, i.e.,

$$\begin{aligned} E(\max_{1\le i\le d}(y_i Z_i)) &= \sum_{i=1}^{d} y_i \\ &= \sum_{i=1}^{d} y_i E(Z_i) \\ &= E\left(\sum_{i=1}^{d} y_i Z_i\right). \end{aligned}$$

This leads to

$$E\left(\sum_{i=1}^{d} y_i Z_i\right) - E\left(\max_{1\le i\le d}(y_i Z_i)\right) = E\Bigg(\underbrace{\sum_{i=1}^{d} y_i Z_i - \max_{1\le i\le d}(y_i Z_i)}_{\ge 0}\Bigg) = 0$$

$$\begin{aligned} &\Rightarrow \quad \sum_{i=1}^{d} y_i Z_i - \max_{1\le i\le d}(y_i Z_i) = 0 \qquad a.s. \text{ (almost surely)} \\ &\Rightarrow \quad \sum_{i=1}^{d} y_i Z_i = \max_{1\le i\le d}(y_i Z_i) \qquad a.s. \end{aligned}$$

Hence $Z_i > 0$ for some $i \in \{1, \dots, d\}$ implies $Z_j = 0$ for all $j \neq i$, and we have for arbitrary

$\boldsymbol{x} \geq \mathbf{0}$

$$\begin{aligned}
&\sum_{i=1}^{d} x_i Z_i = \max_{1\le i\le d}(x_i Z_i) \qquad a.s. \\
\Rightarrow \quad & E\left(\sum_{i=1}^{d} x_i Z_i\right) = E\left(\max_{1\le i\le d}(x_i Z_i)\right) \\
\Rightarrow \quad & \|\boldsymbol{x}\|_1 = \|\boldsymbol{x}\|_D .
\end{aligned}$$

(ii) We have the following list of conclusions:

$$\begin{aligned}
&\|(1,\dots,1)\|_D = 1 \\
&\Rightarrow E\left(\max_{1\le i\le d} Z_i\right) = E(Z_j), \qquad 1 \le j \le d, \\
&\Rightarrow E\bigg(\underbrace{\max_{1\le i\le d} Z_i - Z_j}_{\geq 0}\bigg) = 0, \qquad 1 \le j \le d, \\
&\Rightarrow \max_{1\le i\le d} Z_i - Z_j = 0 \quad a.s., \qquad 1 \le j \le d, \\
&\Rightarrow Z_1 = Z_2 = \ldots = Z_d = \max_{1\le i\le d} Z_i \\
&\Rightarrow E\left(\max_{1\le i\le d}(|x_i|\, Z_i)\right) = E\left(\max_{1\le i\le d}(|x_i|\, Z_1)\right) \\
&\qquad\qquad\qquad\qquad\qquad\quad = E(\|\boldsymbol{x}\|_\infty Z_1) \\
&\qquad\qquad\qquad\qquad\qquad\quad = \|\boldsymbol{x}\|_\infty E(Z_1)
\end{aligned}$$

$$= \|\boldsymbol{x}\|_\infty, \qquad \boldsymbol{x} \in \mathbb{R}^d.$$

□

Theorem 1.3.1 can easily be generalized to sequences of D-norms.

Theorem 1.3.2. Let $\|\cdot\|_{D^n}$, $n \in \mathbb{N}$, be a sequence of D-norms on $\mathbb{R}^d$.

(i) $\forall \boldsymbol{x} \in \mathbb{R}^d : \|\boldsymbol{x}\|_{D^n} \underset{n\to\infty}{\to} \|\boldsymbol{x}\|_1 \iff \exists \boldsymbol{y} > \boldsymbol{0} : \|\boldsymbol{y}\|_{D^n} \underset{n\to\infty}{\to} \|\boldsymbol{y}\|_1$

(ii) $\forall \boldsymbol{x} \in \mathbb{R}^d : \|\boldsymbol{x}\|_{D^n} \underset{n\to\infty}{\to} \|\boldsymbol{x}\|_\infty \iff \|\boldsymbol{1}\|_{D^n} \underset{n\to\infty}{\to} 1$

Corollary 1.3.1 carries over.

Proof. Let $\left(Z_1^{(n)}, \dots, Z_d^{(n)}\right)$ be a generator of $\|\cdot\|_{D^n}$. Again we only need to show the implication "$\Leftarrow$".

(i) We suppose $\|\boldsymbol{y}\|_1 - \|\boldsymbol{y}\|_{D^n} \to_{n\to\infty} 0$ for some $\boldsymbol{y} > \boldsymbol{0} \in \mathbb{R}^d$. With the notation $M_j := \left\{y_j Z_j^{(n)} = \max_{1\le i\le d} y_i Z_i^{(n)}\right\}$ we get for every $j = 1, \dots, d$

$$\|\boldsymbol{y}\|_1 - \|\boldsymbol{y}\|_{D^n} = E\left(\underbrace{\sum_{i=1}^d y_i Z_i^{(n)} - \max_{1\le i\le d} y_i Z_i^{(n)}}_{\ge 0}\right)$$
$$\ge E\left(\left(\sum_{i=1}^d y_i Z_i^{(n)} - \max_{1\le i\le d} y_i Z_i^{(n)}\right) 1_{M_j}\right)$$

$$= E\left(\sum_{\substack{i=1\\i\neq j}}^{d} y_i Z_i^{(n)} 1_{M_j}\right)$$
$$= \sum_{\substack{i=1\\i\neq j}}^{d} y_i E\left(Z_i^{(n)} 1_{M_j}\right) \to_{n\to\infty} 0$$

as the left hand side of this equation converges to zero by assumption: $\|\boldsymbol{y}\|_1 - \|\boldsymbol{y}\|_{D^n} \to_{n\to\infty} 0$. Since $y_i > 0$ for all $i = 1, \dots, d$, we have

$$E\left(Z_i^{(n)} 1_{M_j}\right) \to_{n\to\infty} 0 \tag{1.10}$$

for all $i \neq j$. Now take an arbitrary $\boldsymbol{x} \in \mathbb{R}^d$. From (1.9) we know that known that

$$\begin{aligned}
0 &\leq \|\boldsymbol{x}\|_1 - \|\boldsymbol{x}\|_{D^n} \\
&= E\left(\underbrace{\sum_{i=1}^{d} |x_i|\, Z_i^{(n)} - \max_{1\leq i\leq d} |x_i|\, Z_i^{(n)}}_{\geq 0}\right) \\
&\leq E\left(\sum_{j=1}^{d}\left(\sum_{i=1}^{d} |x_i|\, Z_i^{(n)} - \max_{1\leq i\leq d} |x_i|\, Z_i^{(n)}\right) 1_{M_j}\right) \\
&= \sum_{j=1}^{d} E\left(\left(\sum_{i=1}^{d} |x_i|\, Z_i^{(n)} - \max_{1\leq i\leq d} |x_i|\, Z_i^{(n)}\right) 1_{M_j}\right)
\end{aligned}$$

$$= \sum_{j=1}^{d} \sum_{\substack{i=1 \\ i \neq j}}^{d} |x_i| \underbrace{E\left(Z_i^{(n)} 1_{M_j}\right)}_{\xrightarrow[n\to\infty]{\text{by (1.10)}} 0} \to_{n\to\infty} 0$$

$$\Rightarrow \forall \boldsymbol{x} \in \mathbb{R}^d : \ \|\boldsymbol{x}\|_{D^n} \to_{n\to\infty} \|\boldsymbol{x}\|_1 .$$

(ii) We use inequality (1.9) and obtain

$$\begin{aligned} 0 &\le \|\boldsymbol{x}\|_{D^n} - \|\boldsymbol{x}\|_\infty \\ &= E\left(\max_{1\le i\le d} |x_i|\, Z_i^{(n)}\right) - \max_{1\le i\le d} |x_i| \\ &\le \left(\max_{1\le i\le d} |x_i|\right) E\left(\max_{1\le i\le d} Z_i^{(n)}\right) - \max_{1\le i\le d} |x_i| \\ &= \|\boldsymbol{x}\|_\infty \left(E\left(\max_{1\le i\le d} Z_i^{(n)}\right) - 1\right) \\ &= \|\boldsymbol{x}\|_\infty \left(\|\mathbf{1}\|_{D^n} - 1\right) \xrightarrow[n\to\infty]{} 0. \end{aligned}$$

□

Theorem 1.3.3. Let $\|\cdot\|_{D^n}$, $n \in \mathbb{N}$, be a sequence of D-norms on $\mathbb{R}^d$. We have

(i) $\|\cdot\|_{D^n} \underset{n\to\infty}{\to} \|\cdot\|_1 \iff \forall\, 1 \le i < j \le d : \|\boldsymbol{e}_i + \boldsymbol{e}_j\|_{D^n} \underset{n\to\infty}{\to} 2$

(ii) $\|\cdot\|_{D^n} \underset{n\to\infty}{\to} \|\cdot\|_\infty \iff \exists\, i \in \{1,\dots,d\}\, \forall j \neq i: \|\boldsymbol{e}_i + \boldsymbol{e}_j\|_{D^n} \underset{n\to\infty}{\to} 1.$

Proof. (i) For all $1 \leq i < j \leq d$ we have

$$\begin{aligned}
&2 - \|\boldsymbol{e}_i + \boldsymbol{e}_j\|_{D^{(n)}} \\
&= E\left(Z_i^{(n)} + Z_j^{(n)}\right) - E\left(\max(Z_i^{(n)}, Z_j^{(n)})\right) \\
&= E\left(Z_i^{(n)} + Z_j^{(n)} - \max(Z_i^{(n)}, Z_j^{(n)})\right) \\
&\geq E\left(\left(Z_i^{(n)} + Z_j^{(n)} - \max(Z_i^{(n)}, Z_j^{(n)})\right) 1_{\left\{Z_j^{(n)} = \max_{1\leq k\leq d} Z_k^{(n)}\right\}}\right) \\
&= E\left(Z_i^{(n)} 1_{\left\{Z_j^{(n)} = \max_{1\leq k\leq d} Z_k^{(n)}\right\}}\right) \geq 0.
\end{aligned}$$

Therefore $E\left(Z_i^{(n)} 1_{\left\{Z_j^{(n)} = \max_{1\leq k\leq d} Z_k^{(n)}\right\}}\right) \xrightarrow[n\to\infty]{} 0$, which is (1.10) for $\boldsymbol{y} = \boldsymbol{1}$. We can repeat the remaining steps of the preceding proof and get the desired assertion.
(ii) For our given value of i we have

$$\begin{aligned}
0 &\leq \|\boldsymbol{1}\|_{D^n} - 1 \\
&= E\left(\max_{1\leq k\leq d} Z_k^{(n)} - Z_i^{(n)}\right) \\
&\leq \sum_{j=1}^d E\left(\left(\max_{1\leq k\leq d} Z_k^{(n)} - Z_i^{(n)}\right) 1_{\{Z_j^{(n)} = \max_{1\leq k\leq d} Z_k^{(n)}\}}\right) \\
&= \sum_{j=1}^d E\left(\left(\max\left(Z_i^{(n)}, Z_j^{(n)}\right) - Z_i^{(n)}\right) 1_{\{Z_j^{(n)} = \max_{1\leq k\leq d} Z_k^{(n)}\}}\right)
\end{aligned}$$

$$\leq \sum_{j=1}^{d} E\left(\max\left(Z_i^{(n)}, Z_j^{(n)}\right) - Z_i^{(n)}\right)$$

$$= \sum_{\substack{j=1 \\ j\neq i}}^{d} \left(\|\boldsymbol{e}_i + \boldsymbol{e}_j\|_{D^n} - 1\right) \to_{n\to\infty} 0$$

which proves the assertion by part *(ii)* of Theorem 1.3.2. □

Corollary 1.3.2. Let $\|\cdot\|_D$ be an arbitrary D-norm on $\mathbb{R}^d$.

(i) $\|\cdot\|_D = \|\cdot\|_1 \iff \forall\, 1 \leq i < j \leq d\colon \|\boldsymbol{e}_i + \boldsymbol{e}_j\|_D = 2 = \|\boldsymbol{e}_i + \boldsymbol{e}_j\|_1$

(ii) $\|\cdot\|_D = \|\cdot\|_\infty \iff \exists\, i \in \{1, \dots, d\}\, \forall j \neq i : \|\boldsymbol{e}_i + \boldsymbol{e}_j\|_D = 1 = \|\boldsymbol{e}_i + \boldsymbol{e}_j\|_\infty$

Proof. Put $\|\cdot\|_{D^n} = \|\cdot\|_D$ in the preceding theorem. □

Remark 1.3.1. Choose $1 \leq i < j \leq d$. Note that $\|(x, y)\|_{D_{i,j}} := \|x\boldsymbol{e}_i + y\boldsymbol{e}_j\|_D$, x, $y \in \mathbb{R}$, defines a D-norm on $\mathbb{R}^2$ with generator (Z_i, Z_j), where $(Z_1, ...Z_d)$ generates $\|\cdot\|_D$. From Takahashi's Theorem, part (i), we obtain that the condition

$$\forall 1 \leq i < j \leq d : \|\boldsymbol{e}_i + \boldsymbol{e}_j\|_D = 2$$

is equivalent with the condition

$$\forall x, y \in \mathbb{R}, 1 \le i < j \le d: \left\| x\boldsymbol{e}_i + y\boldsymbol{e}_j \right\|_D = \left\| x\boldsymbol{e}_i + y\boldsymbol{e}_j \right\|_1 = |x| + |y| \,.$$

1.4 Max-Characteristic Function

The Max-Characteristic Function of a Generator

Recall that the generator of a D-norm is not uniquely determined, even its distribution is not.

Lemma 1.4.1 (Balkema)**.** Let $\boldsymbol{X} = (X_1, X_2, \dots, X_d) \ge \boldsymbol{0}$, $\boldsymbol{Y} = (Y_1, Y_2, \dots, Y_d) \ge \boldsymbol{0}$ be rv with $E(X_i), E(Y_i) < \infty$, $1 \le i \le d$. If we have for each $\boldsymbol{x} > \boldsymbol{0} \in \mathbb{R}^d$

$$E\left(\max(1, x_1X_1, \dots, x_dX_d)\right) = E\left(\max(1, x_1Y_1, \dots, x_dY_d)\right),$$

then $\boldsymbol{X} =_D \boldsymbol{Y}$, where "$=_D$" denotes equality in distribution.

Proof. We have for $\boldsymbol{x} > \boldsymbol{0}$ and $c > 0$

$$\begin{aligned}
&E\left(\max\left(1, \frac{X_1}{cx_1}, \dots, \frac{X_d}{cx_d}\right)\right) \\
&= \int_0^\infty 1 - P\left(\max\left(1, \frac{X_1}{cx_1}, \dots, \frac{X_d}{cx_d}\right) \le t\right) dt \\
&= \int_0^\infty 1 - P(1 \le t, X_i \le tcx_i,\, 1 \le i \le d)\, dt
\end{aligned}$$

$$= 1 + \int_1^\infty 1 - P\left(X_i \le tcx_i,\, 1 \le i \le d\right)\, dt$$

The substitution $t \mapsto t/c$ yields that the right-hand side above equals

$$1 + \frac{1}{c}\int_c^\infty 1 - P(X_i \le tx_i, 1 \le i \le d)\, dt.$$

Repeating the preceding arguments with Y_i in place of X_i, we obtain from the assumption that for all $c > 0$

$$\begin{aligned} &\int_c^\infty 1 - P(X_i \le tx_i, 1 \le i \le d)\, dt \\ &= \int_c^\infty 1 - P(Y_i \le tx_i, 1 \le i \le d)\, dt. \end{aligned}$$

Taking right derivatives with respect to c we obtain for $c > 0$

$$1 - P(X_i \le cx_i,\, 1 \le i \le d) = 1 - P(Y_i \le cx_i,\, 1 \le i \le d),$$

and, thus, the assertion. $\square$

Corollary 1.4.1. If $(1, Z_1, \dots, Z_d)$ is the generator of a D-norm, then the distribution of $(Z_1, \dots, Z_d)$ is uniquely determined.

Take, for example, $\boldsymbol{Z} = (1, \dots, 1) \in \mathbb{R}^d$, which generates the sup-norm $\|\cdot\|_\infty$ on $\mathbb{R}^d$. Then $(1, \boldsymbol{Z})$ generates the sup-norm on $\mathbb{R}^{d+1}$.

Let, on the other hand, $\boldsymbol{Z}$ be a random permutation of $(d, 0, \dots, 0) \in \mathbb{R}^d$, which generates $\|\cdot\|_1$. The D-norm generated by $(1, \boldsymbol{Z})$ on $\mathbb{R}^{d+1}$ is

$$
\begin{aligned}
\|\boldsymbol{x}\|_D &= E\left(\max\left(|x_1|, |x_2| Z_2, \dots, |x_{d+1}| Z_D\right)\right)\\
&= E\left(\sum_{i=1}^{d} \left(\max\left(|x_1|, |x_2| Z_2, \dots, |x_{d+1}| Z_D\right)\right) 1(Z_i = d)\right)\\
&= \sum_{i=1}^{d} E\left(\left(\max\left(|x_1|, |x_2| Z_2, \dots, |x_{d+1}| Z_D\right)\right) 1(Z_i = d)\right)\\
&= \frac{1}{d} \sum_{i=2}^{d} \max(|x_1|, d\,|x_i|)\\
&= \sum_{i=2}^{d} \max\left(\frac{|x_1|}{d}, |x_i|\right).
\end{aligned}
$$

Let $\boldsymbol{Z} = (Z_1, \dots, Z_d)$ be the generator of a D-norm $\|\cdot\|_D$. Then we call

$$
\varphi(\boldsymbol{x}) := E(\max(1, |x_1| Z_1, \dots |x_d| Z_d)), \qquad \boldsymbol{x} \in \mathbb{R}^d,
$$

the max-characteristic function of $\boldsymbol{Z}$.

The max-characteristic function of the random permutation of $(d, 0, \dots, 0)$, for instance, is

$$
\varphi(\boldsymbol{x}) = \sum_{i=1}^{d} \max\left(\frac{1}{d}, |x_i|\right), \qquad \boldsymbol{x} \in \mathbb{R}^d.
$$

The max-characteristic function of $(1, \dots, 1)$ **is**

$$\varphi(\boldsymbol{x}) = \max\left(1, \|\boldsymbol{x}\|_\infty\right), \qquad \boldsymbol{x} \in \mathbb{R}^d.$$

We have in general

$$\varphi(\boldsymbol{x}) = E(\max(1, |x_1|\, Z_1, \dots, |x_d|\, Z_d)) \geq E(\max_{1\le i\le d}(|x_i|\, Z_i)) = \|\boldsymbol{x}\|_D$$

and, thus,

$$\begin{aligned} 0 &\le \varphi(\boldsymbol{x}) - \|\boldsymbol{x}\|_D \\ &= E(\max(1, |x_1|\, Z_1, \dots, |x_d|\, Z_d) - \max_{1\le i\le d}(|x_i|\, Z_i)) \\ &= E\left((1 - \max_{1\le i\le d}(|x_i|\, Z_i))1_{\{\max_{1\le i\le d}(|x_i|Z_i)<1\}}\right), \qquad \boldsymbol{x} \in \mathbb{R}^d, \end{aligned}$$

if Z **generates the** D**-norm** $\|\cdot\|_D$**.**

1.5 Convexity of the Set of D-Norms

The Set of D-Norms is Convex

Proposition 1.5.1. The set of D-norms on $\mathbb{R}^d$ is convex, i.e., if $\|\cdot\|_{D_1}$ and $\|\cdot\|_{D_2}$ are D-norms, then

$$\|\cdot\|_{\lambda D_1+(1-\lambda)D_2} := \lambda \|\cdot\|_{D_1} + (1-\lambda)\|\cdot\|_{D_2}$$

is for each $\lambda \in [0,1]$ a D-norm as well.

Take, for example, the convex combination of the two D-norms $\|\cdot\|_\infty$ and $\|\cdot\|_1$:

$$\lambda \|\boldsymbol{x}\|_\infty + (1-\lambda)\|\boldsymbol{x}\|_1 = \lambda \max_{1\le i\le d} |x_i| + (1-\lambda)\sum_{i=1}^d |x_i|.$$

This is the Marshall-Olkin D-norm with parameter $\lambda \in [0,1]$.

Proof of Proposition 1.5.1. Let ξ be a rv with $P(\xi = 1) = \lambda = 1 - P(\xi = 2)$ and independent of $\boldsymbol{Z}^{(1)}$ and $\boldsymbol{Z}^{(2)}$, where $\boldsymbol{Z}^{(1)}$, $\boldsymbol{Z}^{(2)}$ are generators of $\|\cdot\|_{D_1}, \|\cdot\|_{D_2}$. Then $\boldsymbol{Z} := \boldsymbol{Z}^{(\xi)}$ is a generator of $\|\cdot\|_{\lambda D_1+(1-\lambda)D_2}$, as we have for $\boldsymbol{x} \ge \boldsymbol{0}$

$$\begin{aligned}
E\left(\max_{1\le i\le d} x_i Z_i^{(\xi)}\right) &= E\left(\sum_{j=1}^2 \max_{1\le i\le d} x_i Z_i^{(\xi)} 1_{\{\xi=j\}}\right)\\
&= \sum_{j=1}^2 E\left(\left(\max_{1\le i\le d} x_i Z_i^{(\xi)}\right) 1_{\{\xi=j\}}\right)\\
&= \sum_{j=1}^2 E\left(\left(\max_{1\le i\le d} x_i Z_i^{(j)}\right) 1_{\{\xi=j\}}\right)\\
&= \sum_{j=1}^2 E\left(\max_{1\le i\le d} x_i Z_i^{(j)}\right) E\left(1_{\{\xi=j\}}\right)
\end{aligned}$$

$$= \lambda E\left(\max_{1\le i\le d} x_i Z_i^{(1)}\right) + (1-\lambda) E\left(\max_{1\le i\le d} x_i Z_i^{(2)}\right).$$

$\square$

A Bayesian Type of D-Norms

The preceding convexity of the set of D-norms can be viewed as a special case of a Bayesian type D-norm as illustrated by the following example.

Consider the logistic family $\left\{\|\cdot\|_p : p \ge 1\right\}$ of D-norms as defined in (1.8). Let f be a probability density on $[1,\infty)$, i.e., $f \ge 0$ and $\int_1^\infty f(p)\,dp = 1$. Then

$$\|\boldsymbol{x}\|_f := \int_1^\infty \|\boldsymbol{x}\|_p f(p)\,dp, \qquad \boldsymbol{x} \in \mathbb{R}^d,$$

defines a D-norm on $\mathbb{R}^d$. This can easily be seen as follows. Let X be a rv on $[1,\infty)$ with this probability density $f(\cdot)$ and suppose that X is independent from each generator $\boldsymbol{Z}_p$ of $\|\cdot\|_p$, $p \ge 1$. Then

$$\boldsymbol{Z}_f := \boldsymbol{Z}_X$$

generates the D-norm $\|\cdot\|_f$:

$$E(\boldsymbol{Z}_f) = \int_1^\infty E(\boldsymbol{Z}_X \mid X = p) f(p)\,dp = \int_1^\infty E(\boldsymbol{Z}_p) f(p)\,dp = \mathbf{1}$$

and

$$E\left(\max_{1\le i\le d}\left(|x_i|\, Z_{f,i}\right)\right)$$

$$
\begin{aligned}
&= E\left(\max_{1\le i\le d}\left(|x_i|\, Z_{X,i}\right)\right)\\
&= \int_1^\infty E\left(\max_{1\le i\le d}\left(|x_i|\, Z_{X,i}\right) \mid X = p\right) f(p)\, dp\\
&= \int_1^\infty \|\boldsymbol{x}\|_p\, f(p)\, dp.
\end{aligned}
$$

If we take, for instance, the Pareto density $f_\lambda(p) := \lambda p^{-(1+\lambda)}$**,** $p \ge 1$**, with parameter** $\lambda > 0$**, then we obtain**

$$
\|\boldsymbol{x}\|_{f_\lambda} = \int_1^\infty \left(\sum_{i=1}^p |x_i|^p\right)^{1/p} \lambda p^{-(1+\lambda)}\, dp, \qquad \boldsymbol{x} \in \mathbb{R}^d.
$$

The convex combination of two arbitrary D**-norms can obviously be embedded in this Bayesian type approach.**

1.6 D-Norms and Copulas

D-Norms and Copulas

Let the rv $\boldsymbol{U} = (U_1, \dots, U_d)$ **follow a copula, i.e., each component** U_i **is uniformly distributed on** $(0,1)$**. As** $E(U_i) = \int_0^1 u\, du = 1/2$ **, the rv** $\boldsymbol{Z} := 2\boldsymbol{U}$ **is generator of a** D**-norm.**

But not every D**-norm can be generated this way: take, for example,** $d = 2$ **and** $\|(x,y)\|_1 = |x| + |y|$**. Suppose that there exists a rv** $\boldsymbol{U} = (U_1, U_2)$ **following a copula such**

that

$$\|(x,y)\|_1 = 2E\left(\max(|x|\,U_1, |y|\,U_2)\right), \qquad x, y \in \mathbb{R}.$$

Putting $x = y = 1$ **we obtain**

$$2 = 2E\Big(\underbrace{\max(U_1, U_2)}_{\in[0,1]}\Big)$$

and, thus,

$$P(\max(U_1, U_2) = 1) = 1.$$

But

$$\begin{aligned} P(\max(U_1, U_2) = 1) &= P(\{U_1 = 1\} \cup \{U_2 = 1\}) \\ &\le P(U_1 = 1) + P(U_2 = 1) = 0. \end{aligned}$$

It is, moreover, obvious, that $\|\cdot\|_1$ **on** $\mathbb{R}^d$ **with** $d \ge 3$ **cannot be generated by** $2\boldsymbol{U}$**, as** $\|(1,\dots,1)\|_1 = d > 2E\left(\max_{1\le i\le d} U_i\right)$**.**

There are consequently strictly more D**-norms than copulas.**

1.7 Normed Generators Theorem

By $|T|$ **we denote in what follows the number of elements in a set** T**.**

The following auxiliary result can easily be proved by induction, just use the equation

$$\min(\max(a_1,\dots,a_n), a_{n+1}) = \max(\min(a_1, a_{n+1}), \dots, \min(a_n, a_{n+1}))).$$

Lemma 1.7.1. We have for arbitrary numbers $a_1, \dots, a_n \in \mathbb{R}$:

$$\max(a_1, \dots, a_n) = \sum_{\emptyset \neq T \subset \{1,\dots,n\}} (-1)^{|T|-1} \min_{i \in T} a_i,$$

$$\min(a_1, \dots, a_n) = \sum_{\emptyset \neq T \subset \{1,\dots,n\}} (-1)^{|T|-1} \max_{i \in T} a_i.$$

Corollary 1.7.1. If $\boldsymbol{Z}^{(1)}$, $\boldsymbol{Z}^{(2)}$ generate the same D-norm, then

$$E\left(\min_{1 \le i \le d}(|x_i|\, Z_i^{(1)})\right) = E\left(\min_{1 \le i \le d}(|x_i|\, Z_i^{(2)})\right), \qquad \boldsymbol{x} \in \mathbb{R}^d.$$

Proof. Corollary 1.7.1 can be seen as follows:

$$\begin{aligned}
E\left(\min_{1 \le i \le d}\left(|x_i|\, Z_i^{(1)}\right)\right) &= E\left(\sum_{\emptyset \neq T \subset \{1,\dots,d\}} (-1)^{|T|-1} \max_{j \in T}\left(|x_i|\, Z_j^{(1)}\right)\right) \\
&= \sum_{\emptyset \neq T \subset \{1,\dots,d\}} (-1)^{|T|-1} E\left(\max_{j \in T}\left(|x_i|\, Z_j^{(1)}\right)\right) \\
&= \sum_{\emptyset \neq T \subset \{1,\dots,d\}} (-1)^{|T|-1} \left\| \sum_{j \in T} |x_j|\, \boldsymbol{e}_j \right\|_D
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{\emptyset \neq T \subset \{1,\dots,d\}} (-1)^{|T|-1} E\left(\max_{j\in T}\left(|x_i|\, Z_j^{(2)}\right)\right)\\
&= E\left(\sum_{\emptyset \neq T \subset \{1,\dots,d\}} (-1)^{|T|-1} \max_{j\in T}\left(|x_i|\, Z_j^{(2)}\right)\right)\\
&= E\left(\min_{1\le i\le d}\left(|x_i|\, Z_i^{(2)}\right)\right).
\end{aligned}
$$

$\square$

Dual D-Norm Function

Let $\|\cdot\|_D$ be an arbitrary D-norm on $\mathbb{R}^d$ with arbitrary generator $\boldsymbol{Z} = (Z_1, \dots, Z_d)$. Put

$$
\wr \boldsymbol{x} \wr_D := E\left(\min_{1\le i\le \in T}(|x_i|\, Z_i)\right), \qquad \boldsymbol{x} \in \mathbb{R}^d,
$$

which we call the dual D-norm function corresponding to $\|\cdot\|_D$. It is independent of the particular generator $\boldsymbol{Z}$, but the mapping

$$
\|\cdot\|_D \to \wr \cdot \wr_D
$$

is not one-to-one. In particular we have that

$$
\wr \cdot \wr_D = 0
$$

is the least dual D-norm function, corresponding to $\|\cdot\|_D = \|\cdot\|_1$, and

$$
\wr \boldsymbol{x} \wr_D = \min_{1\le i\le d} |x_i| = \wr \boldsymbol{x} \wr_\infty, \qquad \boldsymbol{x} \in \mathbb{R}^d,
$$

is the largest dual D-norm function, corresponding to $\|\cdot\|_D = \|\cdot\|_\infty$, i.e., we have for an arbitrary dual D-norm function the bounds

$$0 = \wr \cdot \wr_1 \le \wr \cdot \wr_D \le \wr \cdot \wr_\infty .$$

While the first inequality is obvious, the second one follows from

$$|x_k| = E(|x_k|\, Z_k) \ge E\left(\min_{1\le i\le d}(|x_i|\, Z_i)\right), \qquad 1 \le k \le d.$$

The Exponent Measure Theorem

The following result is based on the characterization of a max-infinite divisible df in Balkema and Resnick (1977). We, therefore, call it Exponent Measure Theorem.

Put $\boldsymbol{E} := [\boldsymbol{0}, \boldsymbol{\infty}) \setminus \{\boldsymbol{0}\} \subset \mathbb{R}^d$ and $tB := \{t\boldsymbol{b} : \boldsymbol{b} \in B\}$ for an arbitrary set $B \subset \boldsymbol{E}$ and $t > 0$.

Theorem 1.7.1 (Exponent Measure Theorem)**.** Let $\|\cdot\|_D$ be an arbitrary D-norm on $\mathbb{R}^d$. Then

$$\nu\left([\boldsymbol{0}, \boldsymbol{x}]^{\complement} \cap \boldsymbol{E}\right) := \left\| \frac{1}{\boldsymbol{x}} \right\|_D, \qquad \boldsymbol{x} \ge \boldsymbol{0},\ \boldsymbol{x} \ne \boldsymbol{0},$$

with the convention $\|1/\boldsymbol{x}\|_D = \infty$, if some component of $\boldsymbol{x}$ is zero, defines a measure ν on

$\boldsymbol{E}$, which satisfies for each Borel subset B of $\boldsymbol{E}$

$$\nu(tB) = \frac{1}{t}\nu(B), \qquad t > 0.$$

Sketch of the proof. Let $(Z_1, \dots, Z_d)$ be a generator of the D-norm $\|\cdot\|_D$ and put for $\boldsymbol{x} \in \boldsymbol{E}$ and $\emptyset \neq T \subset \{1, \dots, d\}$

$$\nu(\pi_i > x_i,\, i \in T) := E\left(\min_{i \in T} \frac{1}{x_i} Z_i\right),$$

with the convention $0/0 = \infty$, where $\pi_i(\boldsymbol{y}) = y_i$ denotes the projection of $\boldsymbol{y} \in \boldsymbol{E}$ onto its i-th component. Note that by Corollary 1.7.1 the value of $E\left(\min_{i\in T} Z_i/x_i\right)$ does not depend on the special choice of the generator of $\|\cdot\|_D$.

The function ν is defined on a family of subsets of $\boldsymbol{E}$, which is $\cap$-stable and which generates the Borel σ-field $\mathbb{B}(\boldsymbol{E})$ in $\boldsymbol{E}$. In order to extend it to a uniquely determined measure ν on $\mathbb{B}(\boldsymbol{E})$, it has to satisfy $\nu((\boldsymbol{a}, \boldsymbol{b}]) \geq 0$ for $\boldsymbol{a}, \boldsymbol{b} \in \boldsymbol{E}$, $\boldsymbol{a} \leq \boldsymbol{b}$. This will be shown below.

From the well known inclusion exclusion principle we obtain for $\boldsymbol{0} < \boldsymbol{a} \leq \boldsymbol{b}$ the equation

$$\begin{aligned}
\nu((\boldsymbol{a}, \boldsymbol{b}]) &= \nu\left((\boldsymbol{a}, \boldsymbol{\infty}) \backslash \bigcup_{i=1}^{d} \{\pi_i > b_i\}\right) \\
&= \nu\left((\boldsymbol{a}, \boldsymbol{\infty}) \backslash \bigcup_{i=1}^{d} \{\pi_i > b_i;\, \pi_j > a_j,\, j \neq i\}\right) \\
&= \nu\left((\boldsymbol{a}, \boldsymbol{\infty})\right) - \nu\left(\bigcup_{i=1}^{d} \{\pi_i > b_i;\, \pi_j > a_j,\, j \neq i\}\right) \\
&= \nu\left((\boldsymbol{a}, \boldsymbol{\infty})\right)
\end{aligned}$$

$$
\begin{aligned}
&\qquad - \sum_{\emptyset \neq T \subset \{1,\dots,d\}} (-1)^{|T|-1} \nu(\pi_i > b_i,\ i \in T;\ \pi_j > a_j,\ j \notin T) \\
&= E\left(\min_{1\le i\le d} \frac{1}{a_i} Z_i\right) \\
&\qquad - \sum_{\emptyset \neq T \subset \{1,\dots,d\}} (-1)^{|T|-1} E\left(\min\left(\min_{i\in T} \frac{1}{b_i} Z_i,\ \min_{j\notin T} \frac{1}{a_j} Z_j\right)\right) \\
&= \sum_{T \subset \{1,\dots,d\}} (-1)^{|T|} E\left(\min\left(\min_{i\in T} \frac{1}{b_i} Z_i,\ \min_{j\notin T} \frac{1}{a_j} Z_j\right)\right) \\
&= \sum_{\boldsymbol{m} \in \{0,1\}^d} (-1)^{\sum m_i} E\left(\min_{1\le i\le d} \left(\frac{Z_i}{b_i}\right)^{m_i} \left(\frac{Z_i}{a_i}\right)^{1-m_i}\right) \\
&= E\left(\sum_{\boldsymbol{m} \in \{0,1\}^d} (-1)^{\sum m_i} \min_{1\le i\le d} \left(\frac{Z_i}{b_i}\right)^{m_i} \left(\frac{Z_i}{a_i}\right)^{1-m_i}\right).
\end{aligned}
$$

We claim that the integrand in the above expectation is nonnegative, i.e., we claim that for $\mathbb{R}^d \ni \mathbf{0} \le \boldsymbol{x} \le \boldsymbol{y}$

$$
\sum_{\boldsymbol{m} \in \{0,1\}^d} (-1)^{\sum m_i} \min_{1\le i\le d} \left(x_i^{m_i} y_i^{1-m_i}\right) \ge 0. \tag{1.11}
$$

Let U be a rv which follows the uniform distribution on $(0,1)$, and put $\boldsymbol{U} = (U, \dots, U) \in \mathbb{R}^d$. The df of $\boldsymbol{U}$ is

$$
F_{\boldsymbol{U}}(\boldsymbol{u}) := P(\boldsymbol{U} \le \boldsymbol{u}) = \min_{1\le i\le d} u_i, \qquad \boldsymbol{u} \in [0,1]^d.
$$

We, thus, obtain for $\mathbf{0} \le \boldsymbol{u} \le \boldsymbol{v} \le \mathbf{1} \in \mathbb{R}^d$ by the well known inclusion exclusion principle

$$
\begin{aligned}
P(\boldsymbol{U} \in (\boldsymbol{u}, \boldsymbol{v}]) &= \sum_{\boldsymbol{m} \in \{0,1\}^d} (-1)^{\sum m_i} F_{\boldsymbol{U}}\left(\left(u_1^{m_1} v_1^{1-m_1}, \dots, u_d^{m_d} v_d^{1-m_d}\right)\right) \\
&= \sum_{\boldsymbol{m} \in \{0,1\}^d} (-1)^{\sum m_i} \min_{1 \le i \le d} \left(u_i^{m_i} v_i^{1-m_i}\right) \\
&\ge 0.
\end{aligned}
$$

This implies inequality (1.11) by a proper scaling of $\boldsymbol{x}$, $\boldsymbol{y}$.

We have, moreover, $\nu\left(t\left(\boldsymbol{a}, \boldsymbol{b}\right]\right) = t^{-1}\nu\left(\left(\boldsymbol{a}, \boldsymbol{b}\right]\right), t > 0$. The equality $\nu_1(B) := \nu(tB) = t^{-1}\nu(B) =: \nu_2(B)$, thus, holds on a generating class closed under intersections and is, therefore, true for any Borel subset B of $\boldsymbol{E}$[3].

Finally, we have for $\boldsymbol{x} > \mathbf{0} \in \mathbb{R}^d$ by the inclusion exclusion principle and Lemma 1.7.1

$$
\begin{aligned}
\nu\left([\mathbf{0}, \boldsymbol{x}]^{\complement} \cap \boldsymbol{E}\right) &= \nu\left(\bigcup_{i=1}^{d} \{\pi_i > x_i\}\right) \\
&= \sum_{\emptyset \ne T \subset \{1,\dots,d\}} (-1)^{|T|-1} \nu(\pi_i > x_i,\, i \in T) \\
&= \sum_{\emptyset \ne T \subset \{1,\dots,d\}} (-1)^{|T|-1} E\left(\min_{i \in T} \frac{1}{x_i} Z_i\right) \\
&= E\left(\sum_{\emptyset \ne T \subset \{1,\dots,d\}} (-1)^{|T|-1} \min_{i \in T}\left(\frac{1}{x_i} Z_i\right)\right)
\end{aligned}
$$

[3]cf. Bauer (2001, p.32 Remarks)

$$= E\left(\max_{1\le i\le d}\frac{1}{x_i}Z_i\right)$$
$$= \left\|\frac{1}{\boldsymbol{x}}\right\|_D.$$

$\square$

Existence of Normed Generators

The proof of the following theorem is essentially the proof of the de Haan-Resnick representation[4] of a multivariate max-stable df with unit Fréchet margins.

Theorem 1.7.2 (Normed Generators)**.** Let $\|\cdot\|$ be an arbitrary norm on $\mathbb{R}^d$. For any D-norm $\|\cdot\|_D$ on $\mathbb{R}^d$ there exists a generator $\boldsymbol{Z}$ with the additional property $\|\boldsymbol{Z}\| = \mathrm{const}$. The distribution of this generator is uniquely determined.

Corollary 1.7.2. For any D-norm on $\mathbb{R}^d$ there exist generators $\boldsymbol{Z}^{(1)}$, $\boldsymbol{Z}^{(2)}$ with the property $\sum_{i=1}^d Z_i^{(1)} = d$ and $\max_{1\le i\le d} Z_i^{(2)} =$ const.

Proof. Choose $\|\cdot\| = \|\cdot\|_1$ in Theorem 1.7.2. Then

$$\text{const} = \left\|\boldsymbol{Z}^{(1)}\right\|_1 = \sum_{i=1}^d Z_i^{(1)}.$$

[4] de Haan and Resnick (1977)

Taking expectations on both sides yields

$$\text{const} = \sum_{i=1}^{d} E\left(Z_i^{(1)}\right) = d.$$

Choose $\|\cdot\| = \|\cdot\|_\infty$ for the second assertion. $\square$

Proof of Theorem 1.7.2. Let $\|\cdot\|_D$ be an arbitrary norm on $\mathbb{R}^d$. From the Exponent Measure Theorem 1.7.1 we know that

$$\nu\left([\mathbf{0}, \boldsymbol{x}]^\complement \cap \boldsymbol{E}\right) := \left\|\frac{1}{\boldsymbol{x}}\right\|_D, \qquad \boldsymbol{x} \geq \mathbf{0} \in \mathbb{R}^d,\ \boldsymbol{x} \neq \mathbf{0},$$

defines a measure ν on $\boldsymbol{E}$ with the property $\nu(tB) = t^{-1}\nu(B)$ for each Borel subset B of $\boldsymbol{E} = [\mathbf{0}, \boldsymbol{x}) \setminus \{\mathbf{0}\}$ and each $t > 0$.

Denote by $S_{\boldsymbol{E}} := \{\boldsymbol{z} \in \boldsymbol{E} : \|\boldsymbol{z}\| = 1\}$ the unit sphere in $\boldsymbol{E}$ with respect to the norm $\|\cdot\|$. From the equality $\nu(tB) = t^{-1}\nu(B)$ we obtain for $t > 0$ and any Borel subset A of $S_{\boldsymbol{E}}$

$$\begin{aligned}
&\nu\left(\left\{\boldsymbol{x} \in \boldsymbol{E} : \|\boldsymbol{x}\| \geq t,\ \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|} \in A\right\}\right)\\
&= \nu\left(\left\{t\boldsymbol{y} \in \boldsymbol{E} : \|\boldsymbol{y}\| \geq 1,\ \frac{\boldsymbol{y}}{\|\boldsymbol{y}\|} \in A\right\}\right)\\
&= \nu\left(t\left\{\boldsymbol{y} \in \boldsymbol{E} : \|\boldsymbol{y}\| \geq 1,\ \frac{\boldsymbol{y}}{\|\boldsymbol{y}\|} \in A\right\}\right)\\
&= \frac{1}{t}\,\nu\left(\left\{\boldsymbol{y} \in \boldsymbol{E} : \|\boldsymbol{y}\| \geq 1,\ \frac{\boldsymbol{y}}{\|\boldsymbol{y}\|} \in A\right\}\right)
\end{aligned}$$

$$=: \frac{1}{t}\Phi(A) \tag{1.12}$$

where $\Phi(\cdot)$ is the *angular measure* on $S_{\boldsymbol{E}}$ corresponding to $\|\cdot\|$.

Define the one-to-one function $T : \boldsymbol{E} \to [0, \infty) \times S_{\boldsymbol{E}}$ by

$$T(\boldsymbol{x}) = \left(\|\boldsymbol{x}\|, \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|}\right),$$

which is the transformation of a vector $\boldsymbol{x}$ on to its *polar coordinates* with respect to the norm $\|\cdot\|$.

From (1.12) we obtain that the measure $(\nu * T)(B) := \nu(T^{-1}(B))$, induced by ν and T, satisfies

$$\begin{aligned}
(\nu * T)\left((t, \infty) \times A\right) &= \nu\left(\{\boldsymbol{x} \in \boldsymbol{E} : T(\boldsymbol{x}) \in (t, \infty) \times A\}\right) \\
&= \nu\left(\left\{\boldsymbol{x} \in \boldsymbol{E} : \|x\| > t, \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|} \in A\right\}\right) \\
&= \frac{1}{t}\Phi(A) \\
&= \int_A \int_{(t,\infty)} \frac{1}{r^2}\, dr\, d\Phi \\
&= \int_{(t,\infty)\times A} \frac{1}{r^2}\, dr\, d\Phi
\end{aligned}$$

and, hence,

$$(\nu * T)(B) = \int_B r^{-2}\, dr\, d\Phi. \tag{1.13}$$

We have

$$\nu\left([\mathbf{0},\boldsymbol{x}]^{\complement}\right)=\nu\left(T^{-1}\left(T\left([\mathbf{0},\boldsymbol{x}]^{\complement}\right)\right)\right)=(\nu*T)\left(T\left([\mathbf{0},\boldsymbol{x}]^{\complement}\right)\right)$$

with

$$\begin{aligned}T\left([\mathbf{0},\boldsymbol{x}]^{\complement}\right)&=T\left(\{\boldsymbol{y}\in\boldsymbol{E}:y_i>x_i\text{ for some }i\le d\}\right)\\&=\{(r,\boldsymbol{a})\in(0,\infty)\times S_{\boldsymbol{E}}:\,ra_i>x_i\text{ for some }i\le d\}\\&=\left\{(r,\boldsymbol{a})\in(0,\infty)\times S_{\boldsymbol{E}}:\,r>\min_{1\le i\le d}\left(\frac{x_i}{a_i}\right)\right\}\end{aligned}$$

with the temporary convention $0/0=\infty$. Hence, we obtain from equation (1.13)

$$\begin{aligned}\left\|\frac{1}{\boldsymbol{x}}\right\|_D&=\nu\left([\mathbf{0},\boldsymbol{x}]^{\complement}\right)\\&=(\nu*T)\left(T\left([\mathbf{0},\boldsymbol{x}]^{\complement}\right)\right)\\&=(\nu*T)\left(\left\{(r,a)\in(0,\infty)\times S_{\boldsymbol{E}}:r>\min_{1\le i\le d}\frac{x_i}{a_i}\right\}\right)\\&=\int_{\left\{(r,a)\in(0,\infty)\times S_{\boldsymbol{E}}:r>\min_{1\le i\le d}\frac{x_i}{a_i}\right\}}s^{-2}\,ds\,d\Phi\\&=\int_{S_{\boldsymbol{E}}}\int_{\min_{1\le i\le d}\frac{x_i}{a_i}}^{\infty}s^{-2}\,ds\,\Phi(d\boldsymbol{a})\\&=\int_{S_{\boldsymbol{E}}}\frac{1}{\min_{1\le i\le d}\frac{x_i}{a_i}}\,\Phi(d\boldsymbol{a})\end{aligned}$$

$$= \int_{S_{\boldsymbol{E}}} \max_{1\le i\le d} \frac{a_i}{x_i}\, \Phi(d\boldsymbol{a})$$

now with the convention $0/0 = 0$ in the bottom line.

Note that Φ is a finite measure as can be seen as follows. Choose in the preceding equation $x_i = 1$ and let $x_j \to \infty$ for $j \neq i$. Then, by the fact that $\|\boldsymbol{e}_i\|_D = 1,\ 1 \le i \le d$, we obtain

$$1 = \int_{S_{\boldsymbol{E}}} a_i\, \Phi(d\boldsymbol{a}), \qquad 1 \le i \le d. \tag{1.14}$$

The finiteness of Φ now follows from the fact, that all norms on $\mathbb{R}^d$ are equivalent:

$$\begin{aligned} d &= \int_{S_{\boldsymbol{E}}} \sum_{i=1}^{d} a_i\, \Phi(d\boldsymbol{a}) \\ &= \int_{S_{\boldsymbol{E}}} \|\boldsymbol{a}\|_1\ \Phi(d\boldsymbol{a}) \\ &\ge \text{const} \int_{S_{\boldsymbol{E}}} \underbrace{\|\boldsymbol{a}\|}_{=1}\ \Phi(d\boldsymbol{a}) \\ &= \text{const}\ \Phi(S_{\boldsymbol{E}}), \end{aligned}$$

i.e., $\Phi(S_{\boldsymbol{E}}) < \infty$.

Put

$$m := \Phi(S_{\boldsymbol{E}}) \in (0, \infty)$$

Then

$$Q\,(\cdot) := \frac{\Phi\,(\cdot)}{m}$$

defines a probability measure on $S_{\boldsymbol{E}}$.

Let the rv $\boldsymbol{X} = (X_1, \dots X_d) \in S_{\boldsymbol{E}}$ follow this probability measure, i.e. $P\left(\boldsymbol{X} \in \cdot\right) = Q\left(\cdot\right)$. Then we have for $\boldsymbol{Z} := m\boldsymbol{X}$

$$\|\boldsymbol{Z}\| = \|m\boldsymbol{X}\| = m\,\|\boldsymbol{X}\| = m \qquad a.s.$$

as well as

$$\boldsymbol{Z} \geq \mathbf{0},$$

and

$$\begin{aligned} E(Z_i) &= E(mX_i) \\ &= mE(X_i) \\ &= m\int_{S_{\boldsymbol{E}}} a_i\left(P * \boldsymbol{X}\right)(dx) \\ &= m\int_{S_{\boldsymbol{E}}} a_i\, Q(d\boldsymbol{a}) \\ &= m\int_{S_{\boldsymbol{E}}} a_i\, \frac{\Phi(d\boldsymbol{a})}{m} \\ &= m\frac{1}{m}\int_{S_{\boldsymbol{E}}} a_i\, \Phi(d\boldsymbol{a}) \\ &= 1. \end{aligned}$$

by (1.14). Finally, we have

$$E\left(\max_{1\leq i\leq d} \frac{Z_i}{x_i}\right) = E\left(m \max_{1\leq i\leq d} \frac{X_i}{x_i}\right)$$

$$
\begin{aligned}
&= \int_{S_{\boldsymbol{E}}} m \max_{1\le i\le d} \frac{a_i}{x_i} \left(P * (X_1,\dots,X_d)\right)(d\boldsymbol{a}) \\
&= \int_{S_{\boldsymbol{E}}} m \max_{1\le i\le d} \frac{a_i}{x_i}\, Q(d\boldsymbol{a}) \\
&= m \int_{S_{\boldsymbol{E}}} \max_{1\le i\le d} \frac{a_i}{x_i} \frac{\Phi(d\boldsymbol{a})}{m} \\
&= \int_{S_{\boldsymbol{E}}} \max_{1\le i\le d} \frac{a_i}{x_i}\, \Phi(d\boldsymbol{a}) \\
&= \left\|\frac{1}{\boldsymbol{x}}\right\|_D .
\end{aligned}
$$

□

Example 1.7.1. Put $\boldsymbol{Z}^{(1)} := (1,\dots,1)$ and $\boldsymbol{Z}^{(2)} := (X,\dots,X)$, where $X \ge 0$ is a rv with $E(X) = 1$. Both generate the D-norm $\|\cdot\|_\infty$, but only $\boldsymbol{Z}^{(1)}$ satisfies $\left\|\boldsymbol{Z}^{(1)}\right\|_1 = d$.

Example 1.7.2. Let $V_1,\dots,V_d$ be independent and identically gamma distributed rv with density $\gamma_\alpha(x) := x^{\alpha-1}\exp(-x)/\Gamma(\alpha)$, $x > 0$, $\alpha > 0$. Then the rv $\tilde{\boldsymbol{Z}} \in \mathbb{R}^d$ with components

$$
\tilde{Z}_i := \frac{V_i}{V_1 + \dots + V_d}, \qquad i = 1,\dots,d,
$$

follows a symmetric *Dirichlet distribution* $\mathrm{Dir}(\alpha)$ on the closed simplex $\tilde{S}_d = \{\boldsymbol{u} \ge \boldsymbol{0} \in \mathbb{R}^d :$

$\sum_{i=1}^d u_i = 1\}$, see Ng et al. (2011, Theorem 2.1). We obviously have $E(\tilde{Z}_i) = 1/d$ and, thus,

$$\boldsymbol{Z} := d\tilde{\boldsymbol{Z}} \tag{1.15}$$

is a generator of a D-norm $\|\cdot\|_{D(\alpha)}$ on $\mathbb{R}^d$, which we call the *Dirichlet D-norm* with parameter α. We have in particular $\|\boldsymbol{Z}\|_1 = d$.

It is well-known that for a general $\alpha > 0$ the rv $\left(V_i / \sum_{j=1}^d V_j\right)_{i=1}^d$ and the sum $\sum_{j=1}^d V_j$ are independent, see, e.g., the proof of Theorem 2.1 in Ng et al. (2011). As $E(V_1+\dots+V_d) = d\alpha$, we obtain for $\boldsymbol{x} = (x_1, \dots, x_d) \in \mathbb{R}^d$

$$\begin{aligned}
\|\boldsymbol{x}\|_{D(\alpha)} &= E\left(\max_{1\le i\le d}\left(|x_i|\, Z_i\right)\right)\\
&= dE\left(\frac{\max_{1\le i\le d}\left(|x_i|\, V_i\right)}{V_1+\dots+V_d}\right)\\
&= \frac{1}{\alpha}E(V_1+\dots+V_d)E\left(\frac{\max_{1\le i\le d}\left(|x_i|\, V_i\right)}{V_1+\dots+V_d}\right)\\
&= \frac{1}{\alpha}E\left(\max_{1\le i\le d}\left(|x_i|\, V_i\right)\right).
\end{aligned}$$

A generator of $\|\cdot\|_{D(\alpha)}$ is, therefore, also given by $\alpha^{-1}(V_1, \dots, V_d)$.

Metrization of the Space of D-Norms

Denote by $\mathcal{Z}_{\|\cdot\|_D}$ the set of all generators of a given D-norm $\|\cdot\|_D$ on $\mathbb{R}^d$. Theorem 1.7.2 implies the following result.

Lemma 1.8.1. Each set $\mathcal{Z}_{\|\cdot\|_D}$ contains a generator $\boldsymbol{Z}$ with the additional property $\|\boldsymbol{Z}\|_1 = d$. The distribution of this $\boldsymbol{Z}$ is uniquely determined.

Let $\mathbb{P}$ be the set of all probability measures on $S_d := \{\boldsymbol{x} \geq \boldsymbol{0} \in \mathbb{R}^d : \|\boldsymbol{x}\|_1 = d\}$. By the preceding lemma we can identify the set $\mathbb{D}$ of D-norms on $\mathbb{R}^d$ with the subset $\mathbb{P}_D$ of those probability distributions $P \in \mathbb{P}$ which satisfy the additional condition $\int_{S_d} x_i \, P(d\boldsymbol{x}) = 1$, $i = 1, \dots, d$.

Denote by $d_W(P, Q)$ the Wasserstein metric between two probability distributions on S_d, i.e.,

$$
\begin{aligned}
&d_W(P, Q) \\
&:= \inf \left\{ E\left(\|\boldsymbol{X} - \boldsymbol{Y}\|_1\right) : \boldsymbol{X} \text{ has distribution } P, \boldsymbol{Y} \text{ has distribution } Q \right\}.
\end{aligned}
$$

As S_d, equipped with an arbitrary norm $\|\cdot\|$, is a complete separable space, the metric space $(\mathbb{P}, d_W)$ is complete and separable as well; see, e.g., Bolley (2008).

Lemma 1.8.2. The subspace $(\mathbb{P}_D, d_W)$ of $(\mathbb{P}, d_W)$ is also separable and complete.

Proof. Let P_n, $n \in \mathbb{N}$, be a sequence in $\mathbb{P}_D$, which converges with respect to d_W to $P \in \mathbb{P}$. We show that $P \in \mathbb{P}_D$. Let the rv $\boldsymbol{X}$ have distribution P and let $\boldsymbol{X}^{(n)}$ have distribution P_n, $n \in \mathbb{N}$. Then we have

$$\begin{aligned}
\sum_{i=1}^{d} \left| \int_{S_d} x_i \, P(d\boldsymbol{x}) - 1 \right| &= \sum_{i=1}^{d} \left| \int_{S_d} x_i \, P(d\boldsymbol{x}) - \int_{S_d} x_i \, P_n(d\boldsymbol{x}) \right| \\
&= \sum_{i=1}^{d} \left| E\left(X_i - X_i^{(n)} \right) \right| \\
&\le E\left(\sum_{i=1}^{d} \left| X_i - X_i^{(n)} \right| \right) \\
&= E\left(\left\| \boldsymbol{X} - \boldsymbol{X}^{(n)} \right\|_1 \right), \qquad n \in \mathbb{N}.
\end{aligned}$$

As a consequence we obtain

$$\sum_{i=1}^{d} \left| \int_{S_d} x_i \, P(d\boldsymbol{x}) - 1 \right| \le d_W(P, P_n) \to_{n \to \infty} 0,$$

and, thus, $P \in \mathbb{P}_D$. The separability of $\mathbb{P}_D$ can be seen as follows. Let $\mathcal{P}$ be a countable and dense subset of $\mathbb{P}$. Identify each distribution P in $\mathcal{P}$ with a rv $\boldsymbol{Y}$ on S_d that follows this distribution P. Put $\boldsymbol{Z} = \boldsymbol{Y}/E(\boldsymbol{Y})$, where we can assume that each component of $\boldsymbol{Y}$ has positive expectation. This yields a countable subset of $\mathbb{P}_D$, which is dense. $\square$

We can now define the distance between two D-norms $\|\cdot\|_{D_1}$, $\|\cdot\|_{D_2}$ on $\mathbb{R}^d$ by

$$\begin{aligned}
&d_W\left(\|\cdot\|_{D_1}, \|\cdot\|_{D_2}\right)\\
&:= \inf\Big\{E\left(\left\|\boldsymbol{Z}^{(1)} - \boldsymbol{Z}^{(2)}\right\|_1\right) :\\
&\qquad\qquad \boldsymbol{Z}^{(i)} \text{ generates } \|\cdot\|_{D_i}, \left\|\boldsymbol{Z}^{(i)}\right\|_1 = d,\, i = 1, 2\Big\}.
\end{aligned}$$

The space $\mathbb{D}$ of D-norms on $\mathbb{R}^d$, equipped with the distance d_W, is by Lemma 1.8.2 a complete and separable metric space.

Convergence of D-Norms and Weak Convergence of Generators

For the rest of this section we restrict ourselves to generators $\boldsymbol{Z}$ of D-norms on $\mathbb{R}^d$ that satisfy $\|\boldsymbol{Z}\|_1 = d$.

Proposition 1.8.1. Let $\|\cdot\|_{D_n}$, $n \in \mathbb{N} \cup \{0\}$, be a sequence of D-norms on $\mathbb{R}^d$ with corresponding generators $\boldsymbol{Z}^{(n)}$, $n \in \mathbb{N} \cup \{0\}$. Then we have the equivalence

$$d_W\left(\|\cdot\|_{D_n}, \|\cdot\|_{D_0}\right) \to_{n\to\infty} 0 \iff \boldsymbol{Z}^{(n)} \to_D \boldsymbol{Z}^{(0)},$$

where $\to_D$ denotes ordinary convergence in distribution.

Proof. Convergence of probability measures P_n to P_0 with respect to the Wasserstein-metric is equivalent with weak convergence together with convergence of the moments

$$\int_{S_d} \|\boldsymbol{x}\|_1 \; P_n(d\boldsymbol{x}) \to_{n\to\infty} \int_{S_d} \|\boldsymbol{x}\|_1 \; P_0(d\boldsymbol{x}),$$

see, e.g., Villani (2009). But as we have for each probability measure $P \in \mathbb{P}_D$

$$\int_{S_d} \|\boldsymbol{x}\|_1 \; P(d\boldsymbol{x}) = \int_{S_d} d\, P(d\boldsymbol{x}) = d,$$

convergence of the moments is automatically satisfied. $\square$

Lemma 1.8.3. We have for arbitrary D-norms $\|\cdot\|_{D_1}$, $\|\cdot\|_{D_2}$ on $\mathbb{R}^d$ the bound

$$\|\boldsymbol{x}\|_{D_1} \le \|\boldsymbol{x}\|_{D_2} + \|\boldsymbol{x}\|_\infty \, d_W\left(\|\cdot\|_{D_1}, \|\cdot\|_{D_2}\right)$$

and, thus,

$$\sup_{\boldsymbol{x}\in\mathbb{R}^d, \|\boldsymbol{x}\|_\infty \le r} \left| \|\boldsymbol{x}\|_{D_1} - \|\boldsymbol{x}\|_{D_2} \right| \le r\, d_W\left(\|\cdot\|_{D_1}, \|\cdot\|_{D_2}\right), \qquad r \ge 0.$$

Proof. Let $\boldsymbol{Z}^{(i)}$ be a generator of $\|\cdot\|_{D_i}$, $i = 1, 2$. We have

$$\begin{aligned}
\|\boldsymbol{x}\|_{D_1} &= E\left(\max_{1\le i\le d}\left(|x_i|\, Z_i^{(1)}\right)\right) \\
&= E\left(\max_{1\le i\le d}\left(|x_i|\left(Z_i^{(2)} + Z_i^{(1)} - Z_i^{(2)}\right)\right)\right)
\end{aligned}$$

$$
\begin{aligned}
&\le E\left(\max_{1\le i\le d}\left(|x_i|\, Z_i^{(2)}\right)\right) + \|\boldsymbol{x}\|_\infty\, E\left(\max_{1\le i\le d}\left|Z_i^{(1)} - Z_i^{(2)}\right|\right)\\
&\le \|\boldsymbol{x}\|_{D_2} + \|\boldsymbol{x}\|_\infty\, E\left(\left\|\boldsymbol{Z}^{(1)} - \boldsymbol{Z}^{(2)}\right\|_1\right),
\end{aligned}
$$

which implies the assertion. □

Chapter 2

Multivariate Generalized Pareto and Max Stable Distributions

2.1 Multivariate Simple Generalized Pareto Distributions

Multivariate Simple Generalized Pareto Distributions

Let $\boldsymbol{Z} = (Z_1, \dots, Z_d)$ be a generator of a D-norm $\|\cdot\|_D$ with the additional property

$$Z_i \le c, \qquad 1 \le i \le d, \tag{2.1}$$

for some constant $c \ge 1$, see Corollary 1.7.2. Let U be a rv that is uniformly distributed on $(0,1)$ and which is independet of $\boldsymbol{Z}$.

Put

$$\boldsymbol{V} = (V_1, \dots, V_d) := \frac{1}{U}(Z_1, \dots, Z_d) =: \frac{1}{U}\boldsymbol{Z}.$$

Note that for $x > 1$

$$P\left(\frac{1}{U} \leq x\right) = P\left(\frac{1}{x} \leq U\right) = 1 - \frac{1}{x},$$

i.e., $1/U$ follows a standard Pareto distribution (with parameter 1).

We have, moreover, for $x > c$ and $1 \leq i \leq d$ by Fubini's theorem

$$\begin{aligned}
P\left(\frac{1}{U} Z_i \leq x\right) &= P\left(\frac{Z_i}{x} \leq U\right) \\
&= E\left(1\left(\frac{Z_i}{x} \leq U\right)\right) \\
&= \int_{[0,1]\times[0,c]} 1\left(\frac{z}{x} \leq u\right) (P * (U, Z_i)) d(u, z) \\
&= \int_{[0,1]\times[0,c]} 1\left(\frac{z}{x} \leq u\right) ((P * U) \times (P * Z_i)) d(u, z) \\
&= \int_0^c \int_0^1 1\left(\frac{z}{x} \leq u\right) \, (P * U) du \, (P * Z_i) dz \\
&= \int_0^c P\left(\frac{z}{x} \leq U\right) (P * Z_i) \, dz \\
&= \int_0^c 1 - \frac{z}{x} \, (P * Z_i) \, dz \\
&= 1 - \frac{1}{x} \int_0^c z \, (P * Z_i) \, dz \\
&= 1 - \frac{1}{x} E\left(Z_i\right)
\end{aligned}$$

$$= 1 - \frac{1}{x}, \tag{2.2}$$

where $P * X$ denotes the distribution of a rv X, and $P * (X, Y) = (P * X) \times (P * Y)$ if the rv X, Y are independent.

The product Z_i/U, therefore, follows in its upper tail a standard Pareto distribution. The special case $Z_i = 1$ yields the standard Pareto distribution everywhere. We call the distribution of $\boldsymbol{V} = \boldsymbol{Z}/U$ a d-variate (simple) generalized Pareto distribution (simple GPD).

The Distribution Function of a GPD

By repeating the arguments in equation (2.2) we obtain for $\boldsymbol{x} \geq (c, \ldots, c) = \boldsymbol{c}$

$$\begin{aligned} P(\boldsymbol{V} \leq \boldsymbol{x}) &= P\left(\frac{Z_i}{U} \leq x_i, 1 \leq i \leq d\right) \\ &= P\left(\frac{Z_i}{x_i} \leq U, 1 \leq i \leq d\right) \\ &= \int_{[0,c]^d} P\left(U \geq \frac{z_i}{x_i}, 1 \leq i \leq d\right) (P * \boldsymbol{Z}) d(z_1, \ldots, z_d) \\ &= \int_{[0,c]^d} P\left(U \geq \max_{1 \leq i \leq d} \frac{z_i}{x_i}\right) (P * \boldsymbol{Z}) d(z_1, \ldots, z_d) \\ &= \int_{[0,c]^d} 1 - \max_{1 \leq i \leq d} \frac{z_i}{x_i} (P * \boldsymbol{Z}) d(z_1, \ldots, z_d) \end{aligned} \tag{2.3}$$

$$
\begin{aligned}
&= 1 - \int_{[0,c]^d} \max_{1\le i\le d} \frac{z_i}{x_i}\,(P * \boldsymbol{Z})d(z_1,\dots,z_d)\\
&= 1 - E\left(\max_{1\le i\le d} \frac{Z_i}{x_i}\right)\\
&= 1 - \left\|\frac{1}{\boldsymbol{x}}\right\|_D,
\end{aligned}
$$

i.e., the (multivariate) distribution function (df) of V is in its upper tail given by $1 - \|1/\boldsymbol{x}\|_D$.

The Survival Function of a GPD

By repeating the arguments in the derivation of equation (2.3) **again, we obtain for $\boldsymbol{x} \ge \boldsymbol{c}$**

$$
\begin{aligned}
P(\boldsymbol{V} \ge \boldsymbol{x}) &= P\left(U \le \frac{z_i}{x_i}, 1 \le i \le d\right)\\
&= \int_{[0,c]^d} P\left(U \le \frac{z_i}{x_i}, 1 \le i \le d\right)\,(P * \boldsymbol{Z})d(z_1,\dots,z_d)\\
&= \int_{[0,c]^d} P\left(U \le \min_{1\le i\le d} \frac{z_i}{x_i}\right)\,(P * \boldsymbol{Z})d(z_1,\dots,z_d)\\
&= \int_{[0,c]^d} \min_{1\le i\le d} \frac{z_i}{x_i}\,(P * \boldsymbol{Z})d(z_1,\dots,z_d)
\end{aligned}
$$

$$= E\left(\min_{1\le i\le d}\frac{Z_i}{x_i}\right)$$

$$= \wr\, 1/\boldsymbol{x}\, \wr_D. \tag{2.4}$$

An Application to Risk Assessment

Suppose that the joint random losses of a portfolio consisting of d assets are modelled by the rv $\boldsymbol{V}$.

The probability that the d losses jointly exceed the vector $\boldsymbol{x} > \boldsymbol{c}$ is by equation (2.4) given by

$$P(\boldsymbol{V} \ge \boldsymbol{x}) = E\left(\min_{1\le i\le d}\frac{Z_i}{x_i}\right).$$

If we suppose that $\|\cdot\|_D = \|\cdot\|_\infty$, then we can choose the constant function $\boldsymbol{Z} = (1,\dots 1)$ as a generator and, thus,

$$P(\boldsymbol{V} \ge \boldsymbol{x}) = \min_{1\le i\le d}\frac{1}{x_i} = \frac{1}{\max\limits_{1\le i\le d} x_i}, \qquad \boldsymbol{x} \ge (1,\dots,1).$$

If we suppose that $\|\cdot\|_D = \|\cdot\|_1$, then we can choose the random permutation of $(d,0,\dots,0)$ with equal probability $1/d$ as a generator $\boldsymbol{Z}$. in this case we have $\min\limits_{1\le i\le d} Z_i = 0$ and, thus,

$$P(\boldsymbol{V} \ge \boldsymbol{x}) = E\left(\min_{1\le i\le d}\frac{Z_i}{x_i}\right) = 0, \qquad \boldsymbol{x} \ge (d,\dots,d).$$

This example shows that assessing the risk of a portfolio is highly sensitive to the choice of the stochastic model: For $x = (d, \dots, d)$ and $\|\cdot\|_D = \|\cdot\|_\infty$, the probability for the losses jointly exceeding the value d is $1/d$, whereas for $\|\cdot\|_D = \|\cdot\|_1$ it is zero!

Risk assessment has, consequently, become a major application of extreme value analysis in recent years.

2.2 Multivariate Max-Stable Distributions

Introducing Multivariate Max-Stable Distributions

Let now $\boldsymbol{V}^{(1)} = (V_1^{(1)}, \dots, V_d^{(1)})$, $\boldsymbol{V}^{(2)} = (V_1^{(2)}, \dots, V_d^{(2)}), \dots$ be independent copies of the rv $\boldsymbol{V} = \boldsymbol{Z}/U$. Then we obtain for the vector of the componentwise maxima

$$\max_{1\le i\le n} \boldsymbol{V}^{(i)} := \left(\max_{1\le i\le n} V_1^{(i)}, \max_{1\le i\le n} V_2^{(i)}, \dots, \max_{1\le i\le n} V_d^{(i)}\right)$$

from equation (2.3) **for $\boldsymbol{x} > 0$ and n large such that $n\boldsymbol{x} > \boldsymbol{c}$**

$$\begin{aligned} & P\left(\frac{1}{n}\max_{1\le i\le n} \boldsymbol{V}^{(i)} \le \boldsymbol{x}\right) \qquad (2.5)\\ &= P\left(\max_{1\le i\le n} \boldsymbol{V}^{(i)} \le n\boldsymbol{x}\right)\\ &= P\left(\boldsymbol{V}^{(i)} \le n\boldsymbol{x},\ 1 \le i \le n\right)\\ &= \prod_{i=1}^n P\left(\boldsymbol{V}^{(i)} \le n\boldsymbol{x}\right) \end{aligned}$$

$$
\begin{aligned}
&= P\left(\boldsymbol{V} \le n\boldsymbol{x}\right)^n \\
&= \left(1 - \left\|\frac{1}{n\boldsymbol{x}}\right\|_D\right)^n \\
&= \left(1 - \frac{1}{n}\left\|\frac{1}{\boldsymbol{x}}\right\|_D\right)^n \\
&= \left(1 - \frac{\left\|\frac{1}{\boldsymbol{x}}\right\|_D}{n}\right)^n \\
&\underset{n\to\infty}{\longrightarrow} \exp\left(-\left\|\frac{1}{\boldsymbol{x}}\right\|_D\right) =: G(\boldsymbol{x}), \qquad \boldsymbol{x} > \boldsymbol{0} \in \mathbb{R}^d,
\end{aligned}
$$

where $1/x$ is meant componentwise, i.e., $1/\boldsymbol{x} = (1/x_1, \dots, 1/x_d)$.

Suppose that at least one component of x is equal to zero, say component i_0. Then

$$
\begin{aligned}
P\left(\boldsymbol{V} \le n\boldsymbol{x}\right) &\le P\left(V_{i_0} \le n x_{i_0}\right) \\
&= P\left(\frac{Z_{i_0}}{U} \le 0\right) \\
&= P\left(Z_{i_0} \le 0\right) \\
&= P\left(Z_{i_0} = 0\right) < 1
\end{aligned}
$$

by the fact that $E(Z_{i_0}) = 1$. As a consequence we obtain in this case

$$
\begin{aligned}
P\left(\frac{1}{n}\max_{1\le i\le n}\boldsymbol{V}^{(i)} \le \boldsymbol{x}\right) =& P\left(\boldsymbol{V} \le n\boldsymbol{x}\right)^n \\
\le& P\left(Z_{i_0} = 0\right)^n \underset{n\to\infty}{\longrightarrow} 0
\end{aligned}
$$

We, thus, have

$$P\left(\frac{1}{n}\max_{1\le i\le n}\boldsymbol{V}^{(i)}\le \boldsymbol{x}\right)\xrightarrow[n\to\infty]{} G(\boldsymbol{x}),\qquad \boldsymbol{x}\in\mathbb{R}^d,$$

where

$$G(\boldsymbol{x})=\begin{cases}\exp\left(-\left\|\frac{1}{\boldsymbol{x}}\right\|_D\right), & \textbf{if } \boldsymbol{x}>\boldsymbol{0},\\ 0 & \textbf{elsewhere.}\end{cases}$$

As $P(n^{-1}\max_{1\le i\le n}\boldsymbol{V}^{(i)}\le\cdot)$, $n\in\mathbb{N}$, is a sequence of df on $\mathbb{R}^d$, it is easy to check that its limit $G(\cdot)$ is a df itself[1]. It is obvious that the df G satisfies

$$G^n(n\boldsymbol{x})=\exp\left(-\left\|\frac{1}{n\boldsymbol{x}}\right\|_D\right)^n=\exp\left(-\left\|\frac{1}{\boldsymbol{x}}\right\|_D\right)=G(\boldsymbol{x}),$$

i.e.,

$$G^n(n\boldsymbol{x})=G(\boldsymbol{x}),\qquad \boldsymbol{x}\in\mathbb{R}^d,\ n\in\mathbb{N},$$

which is the so called max-stability of G:

Let the rv $\boldsymbol{\xi}\in\mathbb{R}^d$ have df G and let $\boldsymbol{\xi}^{(1)},\boldsymbol{\xi}^{(2)},\dots$ be independent copies of $\boldsymbol{\xi}$. Then we have for the vector of componentwise maxima

$$\begin{aligned}P\left(\frac{1}{n}\max_{1\le i\le n}\boldsymbol{\xi}^{(i)}\le\boldsymbol{x}\right)&=P\left(\max_{1\le i\le n}\boldsymbol{\xi}^{(i)}\le n\boldsymbol{x}\right)\\&=P\left(\boldsymbol{\xi}^{(i)}\le n\boldsymbol{x},\ 1\le i\le n\right)\\&=P\left(\boldsymbol{\xi}\le n\boldsymbol{x}\right)^n\end{aligned}$$

[1] Falk et al. (2011, Section 4.1)

$$
\begin{aligned}
&= G^n(n\boldsymbol{x}) \\
&= G(\boldsymbol{x}), \qquad \boldsymbol{x} \in \mathbb{R}^d,
\end{aligned}
$$

which explains the name max-stability.

The Simple Multivariate Max-Stable Distribution

By keeping $x_i > 0$ fixed and letting x_j tend to infinity for $j \neq i$, we obtain the marginal distribution of G:

$$
\begin{aligned}
G_i(x_i) :=& P\left(\xi_i \le x_i\right) \\
=& \lim_{\substack{x_j \to \infty \\ j \neq i}} P(\xi_i \le x_i, \xi_j \le x_j, j \neq i) \\
=& \lim_{\substack{x_j \to \infty \\ j \neq i}} G(\boldsymbol{x}) \\
=& \lim_{\substack{x_j \to \infty \\ j \neq i}} \exp\left(-\left\|\frac{1}{\boldsymbol{x}}\right\|_D\right) \\
=& \lim_{\substack{x_j \to \infty \\ j \neq i}} \exp\left(-\left\|\left(\frac{1}{x_1}, \dots, \frac{1}{x_i}, \dots, \frac{1}{x_d}\right)\right\|_D\right) \\
=& \exp\left(-\left\|\left(0, \dots, 0, \frac{1}{x_i}, 0, \dots, 0\right)\right\|_D\right) \\
=& \exp\left(-E\left(\frac{1}{x_i} Z_i\right)\right)
\end{aligned}
$$

$$= \exp\left(-\frac{1}{x_i}\right).$$

Each univariate marginal df of G is, consequently,

$$G_{F_1}(x) := \exp\left(-\frac{1}{x}\right), \qquad x > 0,$$

which is the Fréchet df with parameter 1, or unit Fréchet df for short.

We call the multivariate df G with unit Fréchet margins multivariate simple max-stable.

The Standard Multivariate Max-Stable Distribution

Let the rv $\boldsymbol{\xi} \in \mathbb{R}^d$ follow a multivariate simple max-stable df, i.e., $P(\boldsymbol{\xi} \le \boldsymbol{x}) = \exp\left(-\|1/\boldsymbol{x}\|_D\right)$, $\boldsymbol{x} > 0$. **Put**

$$\boldsymbol{\eta} = -\frac{1}{\boldsymbol{\xi}} = -\left(\frac{1}{\xi_1}, \ldots, \frac{1}{\xi_d}\right)$$

and note that $P(\xi_i = 0) = 0$, $1 \le i \le d$. **Then we obtain for $\boldsymbol{x} \le \mathbf{0} \in \mathbb{R}^d$**

$$\begin{aligned} P(\boldsymbol{\eta} \le \boldsymbol{x}) &= P\left(-\frac{1}{\xi_i} \le x_i,\ 1 \le i \le d\right) \\ &= P\left(-\frac{1}{x_i} \ge \xi_i,\ 1 \le i \le d\right) \\ &= P\left(\boldsymbol{\xi} \le -\frac{1}{\boldsymbol{x}}\right) \end{aligned}$$

$$\begin{aligned} &= \exp(-\|\boldsymbol{x}\|_D) \\ &=: G_D(\boldsymbol{x}). \end{aligned}$$

By putting for $x \in \mathbb{R}^d$

$$G_D(\boldsymbol{x}) := \exp\left(-\|(\min(x_1,0),\dots,\min(x_d,0))\|_D\right),$$

we obtain a df on $\mathbb{R}^d$, which is max-stable as well:

$$G_D^n\left(\frac{\boldsymbol{x}}{n}\right) = G_D\left(\boldsymbol{x}\right), \qquad \boldsymbol{x} \in \mathbb{R}^d,\ n \in \mathbb{N}.$$

$G_D^n(\cdot/n)$ is the df of $n\max_{1\le i\le n}\boldsymbol{\eta}^{(i)}$, where $\boldsymbol{\eta}^{(1)}, \boldsymbol{\eta}^{(2)}, \dots$ are independent copies of $\boldsymbol{\eta}$. Note that each univariate margin of G_D is the standard negative exponential df:

$$\begin{aligned} P(\eta_i \le x) &= P(\boldsymbol{\eta} \le x\boldsymbol{e}_i) \\ &= \exp\left(-\|x\boldsymbol{e}_i\|_D\right) \\ &= \exp\left(-|x|\,\|\boldsymbol{e}_i\|_D\right) \\ &= \exp(x), \qquad x \le 0. \end{aligned}$$

We call G_D multivariate standard max-stable (SMS).

Takahashi revisited

Let the rv $\boldsymbol{\eta} = (\eta_1,\dots,\eta_d)$ have in what follows the SMS df

$$P(\boldsymbol{\eta} \le \boldsymbol{x}) = G(\boldsymbol{x}) = \exp\left(-\|\boldsymbol{x}\|_D\right), \qquad \boldsymbol{x} \le \boldsymbol{0} \in \mathbb{R}^d,$$

with an arbitrary D-norm $\|\cdot\|_D$ on $\mathbb{R}^d$. Theorem 1.3.1 can now be formulated as follows.

Theorem 2.2.1. With $\boldsymbol{\eta}$ as above we have the equivalences

(i) $\eta_1, \dots, \eta_d$ are independent

$$\iff \exists\, \boldsymbol{y} < \mathbf{0} \in \mathbb{R}^d : P(\eta_i \le y_i,\, 1 \le i \le d) = \prod_{i=1}^{d} P(\eta_i \le y_i).$$

(ii) $\eta_1 = \eta_2 = \dots = \eta_d$ a.s.

$$\iff P(\eta_1 \le -1, \eta_2 \le -1, \dots, \eta_d \le -1) = 1.$$

Proof. The assumption $\eta_1, \dots, \eta_d$ are independent is equivalent with the condition $\|\cdot\|_D = \|\cdot\|_1$. The assumption $\eta_1 = \eta_2 = \dots = \eta_d$ a.s. is equivalent with the condition $\|\cdot\|_d = \|\cdot\|_\infty$. The assertion is, therefore, an immediate consequence of Theorem 1.3.1. $\square$

The following characterization is an immediate consequence of Theorem 1.3.3. Note that for arbitrary $1 \le i < j \le d$

$$\begin{aligned} P(\eta_i \le -1, \eta_j \le -1) &= P(\eta_i \le -1, \eta_j \le -1, \eta_k \le 0, k \notin \{i, j\}) \\ &= \exp\left(- \|\boldsymbol{e}_i + \boldsymbol{e}_j\|_D\right). \end{aligned}$$

Part (ii) is, obviously, trivial. We list it for the sake of completeness.

Theorem 2.2.2. With $\boldsymbol{\eta}$ as above we have the equivalences

(i) $\eta_1, \dots, \eta_d$ are independent $\iff$ $\eta_1, \dots, \eta_d$ are pairwise independent.

(ii) $\eta_1 = \eta_2 = \dots = \eta_d$ a.s. $\iff$ $\eta_1, \dots, \eta_d$ are pairwise completely dependent.

The distribution of an arbitrary d-variate max-stable rv can be obtained by means of $\boldsymbol{\eta}$ as above together with a proper non random transformation of each component η_i, $1 \le i \le d$, see, e.g., Falk et al. (2011, equation (5.47)). The preceding characterizations, therefore, carry over to an arbitrary multivariate max-stable rv (see (4.4)**).**

2.3 Standard Multivariate Generalized Pareto Distribution

Standard Multivariate Generalized Pareto Distribution

Choose $K < 0$ and put

$$\begin{aligned}\boldsymbol{W} &:= (W_1, \dots, W_d)\\ &:= \left(\max\left(-\frac{U}{Z_1}, K\right), \dots, \max\left(-\frac{U}{Z_d}, K\right)\right),\end{aligned}$$

where U is uniformly distributed on $(0,1)$ and independent of the generator $\boldsymbol{Z}$ of the D-norm $\|\cdot\|_D$, which is bounded by $c \ge 1$. The additional constant K avoids division by zero. Repeating the arguments in equation (2.3) **we obtain**

$$P(\boldsymbol{W} \le \boldsymbol{x}) = 1 - \|\boldsymbol{x}\|_D, \qquad \boldsymbol{x}_0 \le \boldsymbol{x} \le \boldsymbol{0} \in \mathbb{R}^d,$$

where $\boldsymbol{x}_0 < \boldsymbol{0} \in \mathbb{R}^d$ depends on K and c.

Repeating the arguments in equation (2.5) one obtains

$$P(n \max_{1\le i\le n} \boldsymbol{W}^{(i)} \le \boldsymbol{x}) \xrightarrow[n\to\infty]{} \exp(-\|\boldsymbol{x}\|_D), \qquad \boldsymbol{x} \le \boldsymbol{0} \in \mathbb{R}^d,$$

where $\boldsymbol{W}^{(1)}, \boldsymbol{W}^{(2)}, \dots$ are independent copies of $\boldsymbol{W}$.

We call a df H on $\mathbb{R}^d$ a standard GPD, if there exists $\boldsymbol{x}_0 < \boldsymbol{0} \in \mathbb{R}^d$ such that

$$H(\boldsymbol{x}) = 1 - \|\boldsymbol{x}\|_D, \qquad \boldsymbol{x}_0 \le \boldsymbol{x} \le \boldsymbol{0} \in \mathbb{R}^d.$$

Note that the i-th marginal df H_i of H is given by

$$H_i(x) = 1 - \|x\boldsymbol{e}_i\|_D = 1 - |x|\,\|\boldsymbol{e}_i\|_D = 1 + x, \qquad x_{0i} \le x \le 0,\ 1 \le i \le d,$$

which coincides on $[x_{0i}, 0]$ with the uniform df on $[-1, 0]$.

2.4 Max-Stable Random Vectors as Generators of D-Norms

Max-Stable Random Vectors as Generators of D-Norms

Let the rv $\boldsymbol{\eta} = (\eta_1, \dots, \eta_d)$ follow the SMS df

$$G(\boldsymbol{x}) = P(\boldsymbol{\eta} \le \boldsymbol{x}) = \exp(-\|\boldsymbol{x}\|_D), \qquad \boldsymbol{x} \le \boldsymbol{0} \in \mathbb{R}^d.$$

Choose $c \in (0,1)$. Then the rv $1/|\eta_i|^c$ has the df

$$\begin{aligned} P\left(\frac{1}{|\eta_i|^c} \le x\right) &= P\left(\frac{1}{x} \le |\eta_i|^c\right) \\ &= P\left(\frac{1}{x^{1/c}} \le -\eta_i\right) \end{aligned}$$

$$
\begin{aligned}
&= P\left(-\frac{1}{x^{1/c}} \ge \eta_i\right)\\
&= \exp\left(-\frac{1}{x^{1/c}}\right), \qquad x > 0,\ 1 \le i \le d,
\end{aligned}
$$

i.e. $1/|\eta_i|^c$ follows the Fréchet df $F_\alpha(x) = \exp(-x^{-\alpha}), x > 0$, with parameter $\alpha = 1/c$; note that $P(\eta_i = 0) = 0$.

Its expectation is

$$
\begin{aligned}
E\left(\frac{1}{|\eta_i|^c}\right) &= \int_0^\infty x\exp(-x^{-\alpha})x^{-\alpha-1}\alpha\,dx \qquad (2.6)\\
&= \alpha\int_0^\infty x^{-\alpha}\exp(-x^{-\alpha})dx\\
&= \int_0^\infty x^{-\frac{1}{\alpha}}\exp(-x)dx\\
&= \int_0^\infty x^{(1-\frac{1}{\alpha})-1}\exp(-x)dx\\
&= \Gamma\left(1-\frac{1}{\alpha}\right)\\
&= \Gamma(1-c) =: \mu_c
\end{aligned}
$$

The rv

$$
\boldsymbol{Z} = (Z_1,\dots,Z_d) := \frac{1}{\mu_c}\left(\frac{1}{|\eta_1|^c},\dots,\frac{1}{|\eta_d|^c}\right) \qquad (2.7)
$$

now satisfies $Z_i \ge 0$ and $E(Z_i) = 1,\ 1 \le i \le d$, i.e., $\boldsymbol{Z}$ is the generator of a D-norm. **Can we specify it?**

Note that the rv $(1/|\eta_1|^c, \dots, 1/|\eta_d|^c)$ follows a max-stable df with Fréchet-margins:

$$\begin{aligned} H(\boldsymbol{x}) =& P\left(\frac{1}{|\eta_i|^c} \le x_i, 1 \le i \le d\right) \\ =& P\left(\eta_i \le -\frac{1}{x_i^{1/c}}, 1 \le i \le d\right) \\ =& \exp\left(-\left\|\left(\frac{1}{x_1^{1/c}}, \dots, \frac{1}{x_d^{1/c}}\right)\right\|_D\right), \qquad \boldsymbol{x} > \boldsymbol{0} \in \mathbb{R}^d, \end{aligned}$$

and for each $n \in \mathbb{N}$:

$$\begin{aligned} H^n(n^c \boldsymbol{x}) &= \exp\left(-\left\|\frac{1}{(n^c x_1)^{1/c}}, \dots, \frac{1}{(n^c x_d)^{1/c}}\right\|_D\right)^n \\ &= \exp\left(-\frac{n}{n}\left\|\left(\frac{1}{x_1^{1/c}}, \dots, \frac{1}{x_d^{1/c}}\right)\right\|_D\right) \\ &= H(\boldsymbol{x}), \qquad \boldsymbol{x} > \boldsymbol{0} \in \mathbb{R}^d. \end{aligned}$$

Now we can specify the D-norm, which is generated by $\boldsymbol{Z} = \mu_c^{-1}(1/|\eta_1|^c, \dots, 1/|\eta_d|^c)$.

Proposition 2.4.1. The D-norm corresponding to the generator $\boldsymbol{Z}$ defined in (2.7) is given by

$$E\left(\max_{1\le i\le d}(x_iZ_i)\right) = \left\|\left(x_1^{1/c},\dots,x_d^{1/c}\right)\right\|_D^c, \qquad \boldsymbol{x}\ge\boldsymbol{0}\in\mathbb{R}^d. \tag{2.8}$$

If $\eta_1,\dots\eta_d$ in the preceding result are independent, i.e., if the corresponding D-norm is $\|\cdot\|_1$, then Proposition 2.4.1 implies that $\boldsymbol{Z}=\mu_c^{-1}\left(1/\left|\eta_1\right|^c,\dots,1/\left|\eta_d\right|^c\right)$ generates the logistic norm $\|\boldsymbol{x}\|_{1/c}=\left(\sum_{i=1}^d \left|x_i\right|^{1/c}\right)^c$. This was already observed in Proposition 1.2.1.

Proof. Recall that by Fubini's theorem

$$E(Y)=\int_0^\infty P(Y>t)\,dt,$$

if Y is an integrable rv with $Y\ge 0$ a.s. We, consequently, obtain for $\boldsymbol{x}>\boldsymbol{0}\in\mathbb{R}^d$

$$\begin{aligned}
E\left(\max_{1\le i\le d}(x_iZ_i)\right) &= \frac{1}{\mu_c}E\left(\max_{1\le i\le d}\left(\frac{x_i}{\left|\eta_i\right|^c}\right)\right)\\
&= \frac{1}{\mu_c}\int_0^\infty P\left(\max_{1\le i\le d}\left(\frac{x_i}{\left|\eta_i\right|^c}\right)\ge t\right)dt\\
&= \frac{1}{\mu_c}\int_0^\infty 1-P\left(\max_{1\le i\le d}\left(\frac{x_i}{\left|\eta_i\right|^c}\right)\le t\right)dt\\
&= \frac{1}{\mu_c}\int_0^\infty 1-P\left(\frac{x_i}{\left|\eta_i\right|^c}\le t,\,1\le i\le d\right)dt\\
&= \frac{1}{\mu_c}\int_0^\infty 1-P\left(\frac{1}{\left|\eta_i\right|^c}\le \frac{t}{x_i},\,1\le i\le d\right)dt
\end{aligned}$$

$$= \frac{1}{\mu_c} \int_0^\infty 1 - \exp\left(- \left\| \frac{1}{(t/x_1)^{1/c}}, \dots, \frac{1}{(t/x_d)^{1/c}} \right\|_D \right) dt$$
$$= \frac{1}{\mu_c} \int_0^\infty 1 - \exp\left(\frac{-1}{t^{1/c}} \left\| (x_1^{1/c}, \dots, x_d^{1/c}) \right\|_D \right) dt$$
$$= \frac{1}{\mu_c} \left\| (x_1^{1/c}, \dots, x_d^{1/c}) \right\|_D^c \int_0^\infty 1 - \exp\left(- \frac{1}{t^{1/c}} \right) dt$$

by the substitution $t \mapsto \left\| (x_1^{1/c}, \dots, x_d^{1/c}) \right\|_D^c t$.

The integral $\int_0^\infty 1 - \exp\left(-1/t^{1/c}\right)\, dt$ equals by Fubini's theorem $E(Y)$, where Y follows a Fréchet distribution with parameter $1/c$. It was shown in (2.6) that $E(Y) = \mu_c$, which completes the proof. $\square$

Iterating the Sequence of Generators

Taking this new D-norm in (2.8) as the initial D-norm and proceeding as before leads to the D-norm

$$\|(x_1, \dots, x_d)\|_{D^{(2)}} := \left\| \left(x_1^{1/c^2}, \dots, x_d^{1/c^2} \right) \right\|_D^{c^2}, \qquad \boldsymbol{x} \geq \mathbf{0} \in \mathbb{R}^d.$$

We can iterate this problem and obtain in the n-th step

$$\|(x_1, \dots, x_d)\|_{D^{(n)}} := \left\| \left(x_1^{1/c^n}, \dots, x_d^{1/c^n} \right) \right\|_D^{c^n}, \qquad \boldsymbol{x} \geq \mathbf{0} \in \mathbb{R}^d.$$

The question suggests itself: Does this sequence of D-norms converge?

Note: If we choose $\|\cdot\|_D = \|\cdot\|_\infty$, then we obtain for $\boldsymbol{x} \geq \mathbf{0} \in \mathbb{R}^d$

$$\left\| \left(x_1^{1/c}, \dots, x_d^{1/c} \right) \right\|_D^c = \left(\max_{1 \leq i \leq d} x_i^{1/c} \right)^c = \max_{1 \leq i \leq d} x_i = \|(x_1, \dots, x_d)\|_\infty .$$

The conjecture might, therefore, occur that the sequence of D-norms converges to the sup-norm $\|\cdot\|_\infty$, if it converges.

Recall that $\|\cdot\|_\infty \leq \|\cdot\|_D \leq \|\cdot\|_1$ for an arbitrary D-norm and that $c \in (0,1)$. Consequently, we obtain

$$\begin{aligned} \|(x_1, \dots, x_d)\|_{D^{(n)}} &= \left\| \left(x_1^{1/c^n}, \dots, x_d^{1/c^n} \right) \right\|_D^{c^n} \\ &\leq \left\| (x_1^{1/c^n}, \dots, x_d^{1/c^n}) \right\|_1^{c^n} \\ &= \left(\sum_{i=1}^d x_i^{1/c^n} \right)^{c^n} \\ &\xrightarrow[n \to \infty]{} \|(x_1, \dots x_d)\|_\infty , \qquad \boldsymbol{x} \geq \mathbf{0} \in \mathbb{R}^d, \end{aligned}$$

by Lemma 1.1.1 and, hence,

$$\|(x_1, \dots, x_d)\|_{D^{(n)}} \xrightarrow[n \to \infty]{} \|(x_1, \dots x_d)\|_\infty .$$

Chapter 3

The Functional D-Norm

3.1 Introduction

Some Basic Definitions

By $C[0,1] := \{g : [0,1] \to \mathbb{R},\ g \text{ is continuous}\}$ we denote the set of continuous functions from the interval $[0,1]$ to the real line. By $E[0,1]$ we denote the set of those bounded functions $f : [0,1] \to \mathbb{R}$ with only a finite number of discontinuities. Note that $E[0,1]$ is a linear space: If $f_1, f_2 \in E[0,1]$ and $x_1, x_2 \in \mathbb{R}$, then $x_1 f_1 + x_2 f_2 \in E[0,1]$ as well.

Let now $\boldsymbol{Z} = (Z_t)_{t\in[0,1]}$ be a stochastic process on $[0,1]$, i.e., Z_t is a rv for each $t \in [0,1]$. We require that each sample path of $(Z_t)_{t\in[0,1]}$ is a continuous function on $[0,1]$, $\boldsymbol{Z} \in C[0,1]$, for short. We also require that

$$Z_t \geq 0, \quad E(Z_t) = 1, \qquad t \in [0,1],$$

and

$$E\left(\sup_{0\le t\le 1} Z_t\right) < \infty.$$

Then

$$\|f\|_D := E\left(\sup_{0\le t\le 1}\left(|f(t)|\, Z_t\right)\right), \qquad f \in E\left[0,1\right],$$

defines a norm on $E\left[0,1\right]$**: We, obviously, have** $\|f\|_D \ge 0$ **and**

$$\begin{aligned}\|f\|_D = E\left(\sup_{0\le t\le 1}\left(|f(t)|\, Z_t\right)\right) &\le E\left(\left(\sup_{t\in[0,1]} |f(t)|\right)\left(\sup_{t\in[0,1]} Z_t\right)\right)\\ &= \left(\sup_{t\in[0,1]} |f(t)|\right) E\left(\sup_{t\in[0,1]} Z_t\right) < \infty.\end{aligned}$$

Let $\|f\|_D = 0$**. We want to show that** $f = 0$**. Suppose that there exists** $t_0 \in [0,1]$ **with** $f(t_0) \ne 0$**, then**

$$\begin{aligned}0 &= \|f\|_D\\ &= E\left(\sup_{t\in[0,1]}\left(|f(t)|\, Z_t\right)\right)\\ &\ge E(|f(t_0)|\, Z_{t_0})\\ &= |f(t_0)|\, E(Z_{t_0})\\ &= |f(t_0)| > 0,\end{aligned}$$

which is a clear contradiction. We, thus, have established the implication

$$\|f\|_D = 0 \implies f = 0.$$

The reverse implication is obvious. Homogeneity is obvious as well: We have for $f \in E[0,1]$ and $\lambda \in \mathbb{R}$

$$\begin{aligned}
\|\lambda f\|_D &= E\left(\sup_{0\le t\le 1}\left(|\lambda f(t)|\, Z_t\right)\right) \\
&= E\left(|\lambda| \sup_{0\le t\le 1}\left(|f(t)|\, Z_t\right)\right) \\
&= |\lambda|\, E\left(\sup_{0\le t\le 1}\left(|f(t)|\, Z_t\right)\right) \\
&= |\lambda|\, \|f\|_D \,.
\end{aligned}$$

The triangle inequality for $\|\cdot\|_D$ follows from the triangle inequality for real numbers $|x+y| \le |x| + |y|$, $x, y \in \mathbb{R}$:

$$\begin{aligned}
\|f_1 + f_2\|_D &= E\left(\sup_{0\le t\le 1}\left(|f_1 + f_2|\, Z_t\right)\right) \\
&\le E\left(\sup_{0\le t\le 1}\left(|f_1|\, Z_t + |f_2|\, Z_t\right)\right) \\
&\le E\left(\sup_{0\le t\le 1}\left(|f_1|\, Z_t\right) + \sup_{0\le t\le 1}\left(|f_2|\, Z_t\right)\right) \\
&= E\left(\sup_{0\le t\le 1}\left(|f_1|\, Z_t\right)\right) + E\left(\sup_{0\le t\le 1}\left(|f_2|\, Z_t\right)\right)
\end{aligned}$$

$$= \|f_1\|_D + \|f_2\|_D\,, \qquad f_1, f_2 \in E\,[0,1]\,.$$

Measurability of Integrand

Note that $(f(t)Z_t)_{t\in[0,1]}$ is for each $f \in E\,[0,1]$ a stochastic process whose sample paths have only a finite number of discontinuities, namely those of the function f. We, therefore, can find a sequence of increasing index sets $T_n = \{t_1, \dots t_n\}$, $n \in \mathbb{N}$, such that

$$\sup_{t\in[0,1]} (|f(t)|\, Z_t) = \lim_{n\to\infty} \left(\max_{1\le i\le n} (|f(t_i)|\, Z_{t_i}) \right).$$

As $\max_{1\le i\le n} (|f(t_i)|\, Z_{t_i})$ is for each $n \in \mathbb{N}$ a rv, the limit of this sequence, i.e., $\sup_{t\in[0,1]} (|f(t)|\, Z_t)$, is a rv as well. We, therefore, can compute its expectation, which is finite by the bound

$$\begin{aligned}\sup_{t\in[0,1]} (|f(t)|\, Z_t) &=: \|f\boldsymbol{Z}\|_\infty \\ &\le \sup_{t\in[0,1]} (|f(t)|) \sup_{t\in[0,1]} Z_t \\ &= \|f\|_\infty \|\boldsymbol{Z}\|_\infty\end{aligned}$$

and taking expectations. Recall that each function $f \in E[0,1]$ is by the definition of $E[0,1]$ bounded. The process $\boldsymbol{Z} = (Z_t)_{t\in[0,1]}$ is again called generator of the D-norm $\|\cdot\|_D$.

Example of a Generator: The Brown-Resnick Process

A nice example of a generator process is the Brown-Resnick process (Brown and Resnick (1977))

$$Z_t := \exp\left(B_t - \frac{t}{2}\right), \qquad t \in [0,1],$$

where $\boldsymbol{B} := (B_t)_{t\in[0,1]}$ is a standard Brownian motion on $[0,1]$. That is, $\boldsymbol{B} \in C[0,1]$, $B_0 = 0$ and the increments $B_t - B_s$ are independent and normal $N(0, t-s)$ distributed rv with mean zero and variance $t-s$. As a consequence, each B_t with $t > 0$ is $N(0,t)$-distributed. We, therefore, have

$$Z_t > 0, \qquad t \in [0,1],$$

and, for $t > 0$,

$$\begin{aligned}
E(Z_t) &= \exp\left(-\frac{t}{2}\right) E(\exp(B_t)) \\
&= \exp\left(-\frac{t}{2}\right) E\left(\exp\left(t^{1/2}\frac{B_t}{t^{1/2}}\right)\right) \\
&= \exp\left(-\frac{t}{2}\right) \int_{-\infty}^{\infty} \exp(t^{1/2}x)\frac{1}{(2\pi)^{1/2}}\exp\left(-\frac{x^2}{2}\right)\,dx \\
&= \int_{-\infty}^{\infty} \frac{1}{(2\pi)^{1/2}}\exp\left(-\frac{(x-t^{1/2})^2}{2}\right)\,dx \\
&= 1,
\end{aligned}$$

as $\exp\left(-(x-t^{1/2})^2/2\right)/(2\pi)^{1/2}$ **is the density of the normal** $N(t^{1/2},1)$**-distribution.**

It is well known[1] **that for** $x \ge 0$

$$P\left(\sup_{t\in[0,1]} B_t > x\right) = 2P(B_1 > x)$$

and, thus,

$$\begin{aligned}
E\left(\sup_{t\in[0,1]} Z_t\right) &\le E\left(\sup_{t\in[0,1]} \exp\left(B_t\right)\right)\\
&= E\left(\exp\left(\sup_{t\in[0,1]} B_t\right)\right)\\
&= \int_0^\infty P\left(\exp\left(\sup_{t\in[0,1]} B_t\right) > x\right)\,dx\\
&\le 1 + \int_1^\infty P\left(\sup_{t\in[0,1]} B_t > \log(x)\right)\,dx\\
&= 1 + 2\int_1^\infty P(B_1 > \log(x))\,dx\\
&= 1 + 2E(\exp(B_1))\\
&< \infty
\end{aligned}$$

as $\exp(B_1)$ **is standard lognormal distributed, with expectation** $\exp(1/2)$**. The computation of the corresponding** D**-norm is, however, not obvious.**

[1] http://ocw.mit.edu/courses/sloan-school-of-management/15-070j-advanced-stochastic-processes-fall-2013/lecture-notes/MIT15_070JF13_Lec7.pdf

BOUNDS FOR THE FUNCTIONAL D-NORM

Lemma 3.1.1. Each functional D-norm is equivalent with the sup-norm $\|\cdot\|_\infty$, precisely,

$$\|f\|_\infty \leq \|f\|_D \leq \|f\|_\infty \|1\|_D \,, \qquad f \in E[0,1] \,.$$

Proof. Let $\boldsymbol{Z} = (Z_t)_{t\in[0,1]}$ be a generator of $\|\cdot\|_D$. We have for each $t_0 \in [0,1]$ and $f \in E[0,1]$

$$\begin{aligned} |f(t_0)| &= E\left(|f(t_0)|\, Z_{t_0}\right) \\ &\leq E\left(\sup_{t\in[0,1]} \left(|f(t)|\, Z_t\right)\right) \\ &= \|f\|_D \\ &\leq E\left(\|f\|_\infty \|\boldsymbol{Z}\|_\infty\right) \\ &= \|f\|_\infty \|1\|_D \,, \end{aligned}$$

which proves the lemma. □

THE FUNCTIONAL L_p-NORM IS NOT A D-NORM

Different to the multivariate case, the functional logistic norm is not a functional D-norm.

Corollary 3.1.1. Each p-norm $\|f\|_p := \left(\int_0^1 |f(t)|^p \, dt\right)^{1/p}$ with $p \in [1, \infty)$ is not a D-norm.

Proof. Choose $\varepsilon \in (0, 1)$ and put $f_\varepsilon(\cdot) := 1_{[0,\varepsilon]}(\cdot) \in E\,[0, 1]$. Then $\|f_\varepsilon\|_\infty = 1 > \varepsilon^{1/p} = \|f_\varepsilon\|_p$. The p-norm, therefore, does not satisfy the first inequality in the preceding result. □

A Functional Version of Takahashi's Theorem

The next consequence of Lemma 3.1.1 is obvious. This is a functional version of Takahashi's Theorem 1.3.1 for $\|\cdot\|_\infty$. Note that there cannot exist an extension to the functional case with $\|\cdot\|_1$, as this is not a functional D-norm by the preceding result.

Corollary 3.1.2. A functional D-norm $\|\cdot\|_D$ is the sup-norm $\|\cdot\|_\infty$ iff $\|1\|_D = 1$.

3.2 Generalized Pareto Processes

Defining a Simple Generalized Pareto Process

Let $\boldsymbol{Z} = (Z_t)_{t\in[0,1]}$ be the generator of a functional D-norm $\|\cdot\|_D$ with the additional property

$$Z_t \le c, \qquad t \in [0, 1], \tag{3.1}$$

for some constant $c \geq 1$. For each functional D-norm there exists a generator with this additional property, see de Haan and Ferreira (2006, equation (9.4.9)). This might be viewed as a funcitonal analogue of the Normed Generators Theorem 1.7.2. Let U be a rv that is uniformly distributed on $(0, 1)$ and which is independent of Z. Put

$$\boldsymbol{V} := (V_t)_{t\in[0,1]} := \frac{1}{U}(Z_t)_{t\in[0,1]} =: \frac{1}{U}\boldsymbol{Z}. \tag{3.2}$$

Repeating the arguments in equation (2.3) **we obtain for $g \in E[0, 1]$ with $g(t) \geq c$, $t \in [0, 1]$,**

$$\begin{aligned}
&P(\boldsymbol{V} \leq g) \qquad (3.3)\\
&= P\left(\frac{1}{U}\boldsymbol{Z} \leq g\right)\\
&= P\left(U \geq \frac{Z_t}{g(t)},\, t \in [0,1]\right)\\
&= \int_{[0,c]^{[0,1]}} P\left(U \geq \frac{z_t}{g(t)},\, t \in [0,1]\right) (P * \boldsymbol{Z})\left(d(z_t)_{t\in[0,1]}\right)\\
&= \int_{[0,c]^{[0,1]}} P\left(U \geq \sup_{t\in[0,1]} \frac{z_t}{g(t)}\right) (P * \boldsymbol{Z})\left(d(z_t)_{t\in[0,1]}\right)\\
&= \int_{[0,c]^{[0,1]}} 1 - P\left(U \leq \sup_{t\in[0,1]} \frac{z_t}{g(t)}\right) (P * \boldsymbol{Z})\left(d(z_t)_{t\in[0,1]}\right)\\
&= 1 - \int_{[0,c]^{[0,1]}} \sup_{t\in[0,1]} \frac{z_t}{g(t)} \, (P * \boldsymbol{Z})\left(d(z_t)_{t\in[0,1]}\right)
\end{aligned}$$

$$
\begin{aligned}
&= 1 - E\left(\sup_{t\in[0,1]} \frac{Z_t}{g(t)}\right)\\
&= 1 - \left\|\frac{1}{g}\right\|_D,
\end{aligned}
$$

i.e., the functional df of the process V is in its upper tail given by $1 - \|1/g\|_D$. We have, moreover,

$$
P(V_t \le x) = 1 - \frac{1}{x}, \qquad x \ge c,\ t \in [0,1],
$$

i.e. each marginal df of the process V is in its upper tail equal to the standard Pareto distribution. We, therefore, call the process V simple generalized Pareto process; see Ferreira and de Haan (2014) and Dombry and Ribatet (2015).

Survival Function of a Simple Generalized Pareto Process

The following result extends the survival function of a multivariate GPD as in equation (2.4) **to simple generalized Pareto processes.**

Proposition 3.2.1. Let $\boldsymbol{Z} = (Z_t)_{t\in[0,1]}$ be the generator of a functional D-norm $\|\cdot\|_D$ with the additional property $\|\boldsymbol{Z}\|_\infty \le c$ for some constant $c \ge 1$. Then we obtain for $g \in E[0,1]$

with $g(t) \geq c$, $t \in [0,1]$,

$$P(\boldsymbol{V} \geq g) = P(\boldsymbol{V} > g) = E\left(\inf_{t\in[0,1]}\left(\frac{Z_t}{g(t)}\right)\right).$$

Proof. Repeating the arguments in equation (3.3), we obtain

$$\begin{aligned}
&P(\boldsymbol{V} > g)\\
&= \int_{[0,c]^{[0,1]}} P\left(U < \frac{z_t}{g(t)},\ t \in [0,1]\right) (P * \boldsymbol{Z})\left(d(z_t)_{t\in[0,1]}\right)\\
&= \int_{[0,c]^{[0,1]}} P\left(U \leq \inf_{t\in[0,1]} \frac{z_t}{g(t)}\right) (P * \boldsymbol{Z})\left(d(z_t)_{t\in[0,1]}\right)\\
&= \int_{[0,c]^{[0,1]}} \inf_{t\in[0,1]} \frac{z_t}{g(t)}\ (P * \boldsymbol{Z})\left(d(z_t)_{t\in[0,1]}\right)\\
&= E\left(\inf_{t\in[0,1]} \frac{Z_t}{g(t)}\right).
\end{aligned}$$

□

Excursion stability of a Generalized Pareto Process

Corollary 3.2.1. We obtain under the conditions of Lemma 3.2.1 and the additional condition

$E\left(\inf_{t\in[0,1]} Z_t\right) > 0$

$$P(\boldsymbol{V} \ge xg \,|\, \boldsymbol{V} \ge g) = \frac{1}{x}, \qquad x \ge 1.$$

Proof. We have

$$\begin{aligned}
&P(\boldsymbol{V} \ge xg \,|\, \boldsymbol{V} \ge g) \\
&= \frac{P(\boldsymbol{V} \ge xg, \boldsymbol{V} \ge g)}{P(\boldsymbol{V} \ge g)} \\
&= \frac{P(\boldsymbol{V} \ge xg)}{P(\boldsymbol{V} \ge g)} \\
&= \frac{E\left(\inf_{t\in[0,1]} \frac{Z_t}{xg(t)}\right)}{E\left(\inf\limits_{t\in[0,1]} \frac{Z_t}{g(t)}\right)} = \frac{1}{x}.
\end{aligned}$$

$\square$

The conditional excursion probability $P(\boldsymbol{V} \ge xg|\boldsymbol{V} \ge g) = 1/x$, $x \ge 1$, does not depend on g. We, therefore, call the process $\boldsymbol{V}$ excursion stable.

Sojourn Time of a Stochastic Process

The time, which the process $\boldsymbol{V} = (V_t)_{t\in[0,1]}$ spends above the function $g \in E[0,1]$ $g \geq c \geq 1$, is called its sojourn time above g, denoted by

$$ST(g) = \int_0^1 1_{(g(t),\infty)}(V_t)\ dt.$$

From Fubini's theorem we obtain

$$\begin{aligned} E\left(ST(g)\right) &= E\left(\int_0^1 1_{(g(t),\infty)}(V_t)\ dt\right) \\ &= \int_0^1 E\left(1_{(g(t),\infty)}(V_t)\right)\ dt \\ &= \int_0^1 P\left(V_t > g(t)\right) dt \\ &= \int_0^1 \frac{1}{g(t)} dt. \end{aligned}$$

Recall that $P(V_t \leq x) = 1 - 1/x$, $x \geq c$, $t \in [0,1]$.

By choosing the constant function $g(t) := s \geq c$, we obtain for the expected sojourn time of the process $\boldsymbol{V}$ above the constant s

$$E(ST(s)) = E\left(\int_0^1 1_{(s,\infty)}(V_t)\, dt\right) = \frac{1}{s}.$$

This implies

$$E(ST(s) \mid ST(s) > 0) = \frac{E(ST(s))}{1 - P(ST(s) = 0)}$$
$$= \frac{1/s}{1 - P(V_t \le s, t \in [0,1])}$$
$$= \frac{1}{\|1\|_D},$$

independent of $s \ge c$**.**

3.3 Max-Stable Processes

Introducing Max-Stable Processes

Let $\boldsymbol{V}^{(1)}, \boldsymbol{V}^{(2)}, \ldots$ **be a sequence of independent copies of** $\boldsymbol{V} = \boldsymbol{Z}/U$**, where the generator** $\boldsymbol{Z}$ **satisfies the additional boundedness condition** (3.1)**. We obtain for** $g \in E[0,1]$**,** $g > 0$**,**

$$P\left(\frac{1}{n} \max_{1 \le i \le n} \boldsymbol{V}^{(i)} \le g\right)$$
$$= P\left(\boldsymbol{V}^{(i)} \le ng,\, 1 \le i \le n\right)$$
$$= \prod_{i=1}^{n} P\left(\boldsymbol{V}^{(i)} \le ng\right)$$

$$
\begin{aligned}
&= P\left(\boldsymbol{V} \le ng\right)^n \\
&= \left(1 - \left\|\frac{1}{ng}\right\|_D\right)^n \\
&\xrightarrow[n\to\infty]{} \exp\left(-\left\|\frac{1}{g}\right\|_D\right),
\end{aligned}
$$

where the mathematical operations $\max_{1\le i\le n} \boldsymbol{V}_i^{(n)}$, etc. are taken componentwise.

The question now occurs: Is there a stochastic process $\boldsymbol{\xi} = (\xi_t)_{t\in[0,1]}$ on $[0,1]$ with

$$
P(\boldsymbol{\xi} \le g) = \exp\left(-\left\|\frac{1}{g}\right\|_D\right), \qquad g \in E\left[0,1\right],\, g > 0?
$$

If ξ actually exists: Does it have continuous sample paths?

If such ξ exists, it is a max-stable process: Let $\boldsymbol{\xi}^{(1)}, \boldsymbol{\xi}^{(2)}, \dots$ be a sequence of independent copies of the process $\boldsymbol{\xi}$. Then we obtain for $g \in E[0,1]$, $g > 0$, and $n \in \mathbb{N}$

$$
\begin{aligned}
P\left(\frac{1}{n}\max_{1\le i\le n} \boldsymbol{\xi}^{(i)} \le g\right) &= P\left(\max_{1\le i\le n} \boldsymbol{\xi}^{(i)} \le ng\right) \\
&= P\left(\boldsymbol{\xi}^{(i)} \le ng,\, 1 \le i \le n\right) \\
&= \prod_{i=1}^n P\left(\boldsymbol{\xi}^{(i)} \le ng\right) \\
&= P\left(\boldsymbol{\xi} \le ng\right)^n \\
&= \exp\left(-\left\|\frac{1}{ng}\right\|_D\right)^n
\end{aligned}
$$

$$
\begin{aligned}
&= \exp\left(-n\left\|\frac{1}{ng}\right\|_D\right)^n \\
&= P(\boldsymbol{\xi} \leq g).
\end{aligned}
$$

For the existence of such processes see Theorem 4.7.1.

Chapter 4

Tutorial: D-Norms & Multivariate Extremes

4.1 Univariate Extreme Value Theory

Let X be $\mathbb{R}$-valued random variable (rv) and suppose that we are only interested in large values of X, where we call a realization of X large, if it exceeds a given high threshold $t \in \mathbb{R}$. In this case we choose the data window $A = (t, \infty)$ or, better adapted to our purposes, we put $t \in \mathbb{R}$ on a linear scale and define

$$A_n = (a_n t + b_n, \infty)$$

for some norming constants $a_n > 0,\ b_n \in \mathbb{R}$. We are, therefore, only interested in values of $X \in A_n$.

Denote by F the distribution function (df) of X. We obtain for $s \geq 0$

$$P\{X \leq a_n(t+s) + b_n \mid X > a_n t + b_n\}$$
$$= 1 - \frac{1 - F(a_n(t+s) + b_n)}{1 - F(a_n t + b_n)},$$

thus facing the problem:

What is the limiting behavior of

$$\frac{1 - F(a_n(t+s) + b_n)}{1 - F(a_n t + b_n)} \longrightarrow_{n\to\infty} ? \tag{4.1}$$

Extreme Value Distributions

Let $X_1, X_2, \ldots$ **be independent copies of** X**. Suppose that there exist constants** $a_n > 0,\ b_n \in \mathbb{R}$ **such that for** $x \in \mathbb{R}$

$$P\left(\frac{\max_{1\le i\le n} X_i - b_n}{a_n} \le x\right) = F^n(a_n x + b_n) \ \to_{n\to\infty} G(x)$$

for some (non degenerate) limiting df G**. Then we say that** F **belongs to the domain of attraction of** G**, denoted by** $F \in \mathcal{D}(G)$**. In this case we deduce from the Taylor expansion** $\log(1+\varepsilon) = \varepsilon + O(\varepsilon^2)$ **for** $\varepsilon \to 0$ **the equivalence**

$$\begin{aligned}
&F^n(a_n x + b_n) \ \longrightarrow_{n\to\infty} \ G(x)\\
&\quad\Leftrightarrow \quad n\log(1 - (1 - F(a_n x + b_n))) \ \longrightarrow_{n\to\infty} \ \log(G(x))\\
&\quad\Leftrightarrow \quad n(1 - F(a_n x + b_n)) \ \longrightarrow_{n\to\infty} \ -\log(G(x))
\end{aligned}$$

if $0 < G(x) \le 1$**, and hence,**

$$\frac{1 - F(a_n(t+s) + b_n)}{1 - F(a_n t + b_n)} \ \longrightarrow_{n\to\infty} \ \frac{\log(G(t+s))}{\log(G(t))} \tag{4.2}$$

if $0 < G(t) < 1$**.**

By the meanwhile classical article by Gnedenko (1943) (see also de Haan (1975) and Galambos (1987)) we know that $F \in \mathcal{D}(G)$ only if $G \in \{G_\alpha : \alpha \in \mathbb{R}\}$, with

$$G_\alpha(x) = \begin{cases} \exp\big(-(-x)^\alpha\big), & x \le 0, \\ 1, & x > 0, \end{cases} \qquad \textbf{for } \alpha > 0,$$

$$G_\alpha(x) = \begin{cases} 0, & x \le 0, \\ \exp(-x^\alpha), & x > 0, \end{cases} \qquad \textbf{for } \alpha < 0$$

and

$$G_0(x) := \exp(-e^{-x}), \qquad x \in \mathbb{R},$$

being the family of (reverse) Weibull, Fréchet and the Gumbel distribution. Note that $G_{-1}(x) = \exp(x)$, $x \le 0$, is the standard inverse exponential df.

Max-Stability of Extreme Value Distributions

The characteristic property of the class of the extreme value distributions (EVD) $\{G_\alpha : \alpha \in \mathbb{R}\}$ is their max-stability, i.e., for each $\alpha \in \mathbb{R}$ and each $n \in \mathbb{N}$ there exist constants $a_n > 0$, $b_n \in \mathbb{R}$, depending on α, such that

$$G^n(a_n x + b_n) = G(x), \qquad x \in \mathbb{R}. \tag{4.3}$$

For $G(x) = \exp(x)$, $x \le 0$, for example, we have $a_n = 1/n$, $b_n = 0$, $n \in \mathbb{N}$:

$$G^n\left(\frac{x}{n}\right) = \exp\left(\frac{x}{n}\right)^n = \exp(x) = G(x).$$

Let $\eta^{(1)}, \eta^{(2)}, \dots$ be independent copies of a rv η that follows the df G_α. In terms of rv, equation (4.3) **means**

$$P\left(\frac{\max_{1\le i\le n}\eta^{(i)} - b_n}{a_n} \le x\right) = P(\eta \le x), \qquad x \in \mathbb{R}.$$

This is the reason, why G_α is called a max-stable df, and the set $\{G_\alpha : \alpha \in \mathbb{R}\}$ collects all univariate max-stable distributions which are non degenerate, i.e., they are not concentrated in one point in $\mathbb{R}$ see, e.g., Galambos (1987, Theorem 2.4.1).

Generalized Pareto Distributions

If we assume that $F \in \mathcal{D}(G_\alpha)$, we obtain from (4.2) **that**

$$\begin{aligned}
&P\left(\frac{X-b_n}{a_n} \le t+s \mid \frac{X-b_n}{a_n} > t\right)\\
&= 1 - \frac{n(1-F(a_n(t+s)+b_n))}{n(1-F(a_nt+b_n))}\\
&\longrightarrow_{n\to\infty} 1 - \frac{\log(G_\alpha(t+s))}{\log(G_\alpha(t))}\\
&= \begin{cases} H_\alpha\left(1+\frac{s}{t}\right), & \textbf{if } \alpha \ne 0,\\ H_0(s), & \textbf{if } \alpha = 0. \end{cases} \qquad s \ge 0,
\end{aligned}$$

provided $0 < G_\alpha(t) < 1$. The family

$$H_\alpha(s) := 1 + \log(G_\alpha(s)), \qquad 0 < G_\alpha(s) < 1,$$

$$= \begin{cases} 1-(-s)^\alpha,\ -1\le s\le 0, & \text{if } \alpha>0,\\ 1-s^\alpha,\ s\ge 1, & \text{if } \alpha<0,\\ 1-\exp(-s),\ s\ge 0, & \text{if } \alpha=0, \end{cases}$$

of df parameterized by $\alpha \in \mathbb{R}$ is the class of (univariate) generalized Pareto df (GPD) coming along with the family of EVD. Notice that H_α with $\alpha < 0$ is a Pareto distribution, H_1 is the uniform distribution on $(-1, 0)$, and H_0 is the standard exponential distribution.

The preceding considerations are the reason, why random exceedances above a high threshold are typically modelled as iid observations coming from a (univariate) GPD.

It was, for example, first observed by van Dantzig (1960)[1] that floods, which exceed some high threshold, follow approximately an exponential df.

Consequence: Suppose that your data are realizations from iid observations, whose common df is in the domain of attraction of an extreme value df. Almost every textbook df satisfies this condition. Then the approximation of exceedances above high thresholds by a GPD is, consequently, a straightforward option and typically used in risk assessment.

4.2 Multivariate Extreme Value Distributions

In complete accordance with the univariate case we call a df G on $\mathbb{R}^d$ max-stable, if for every $n \in \mathbb{N}$ there exists vectors $\boldsymbol{a}_n > 0$, $\boldsymbol{b}_n \in \mathbb{R}^d$ such that

$$G^n\left(\boldsymbol{a}_n\boldsymbol{x} + \boldsymbol{b}_n\right) = G(\boldsymbol{x}), \qquad \boldsymbol{x} \in \mathbb{R}^d. \tag{4.4}$$

[1] http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.189.3302&rep=rep1&type=pdf

All operations on vectors such as addition, multiplication etc. are always meant componentwise. The preceding equation can again be formulated in terms of componentwise maxima of independent copies $\boldsymbol{\eta}^{(1)}, \boldsymbol{\eta}^{(2)}, \dots$ of a rv $\boldsymbol{\eta} = (\eta_1, \dots, \eta_d)$ that realizes in $\mathbb{R}^d$, and which follows the df G:

$$P\left(\frac{\max_{1\le i\le n} \boldsymbol{\eta}^{(i)} - \boldsymbol{b}_n}{\boldsymbol{a}_n} \le \boldsymbol{x}\right) = P\left(\boldsymbol{\eta} \le \boldsymbol{x}\right), \qquad \boldsymbol{x} \in \mathbb{R}^d.$$

Note that also $\max$ is taken componentwise as well as division.

Different to the univariate case, the class of multivariate max-stable distributions or multivariate extreme value distributions (EVD) is no longer a parametric one, indexed by some α. This is obviously necessary for the univariate margins of G. Instead, a nonparametric part occurs, which can be best described in terms of D-norms.

What is a D-Norm?

Definition 4.2.1. A norm $\|\cdot\|_D$ on $\mathbb{R}^d$ is a *D-norm*, if there exists a rv $\boldsymbol{Z} = (Z_1, \dots, Z_d)$ with $Z_i \ge 0$, $E(Z_i) = 1$, $1 \le i \le d$, such that

$$\|\boldsymbol{x}\|_D = E\left(\max_{1\le i\le d} \left(|x_i|\, Z_i\right)\right),$$

$\boldsymbol{x} = (x_1, \dots, x_d) \in \mathbb{R}^d$.

In this case the rv $\boldsymbol{Z}$ is called generator of $\|\cdot\|_D$.

Example 4.1. Here is a list of D-norms and their generators:

- $\|\boldsymbol{x}\|_\infty = \max_{1\le i\le d} |x_i|$, generated by $\boldsymbol{Z} = (1,\dots,1)$.
- $\|\boldsymbol{x}\|_1 = \sum_{i=1}^d |x_i|$, generated by $\boldsymbol{Z}$ = random permutation of $(d,0,\dots,0) \in \mathbb{R}^d$ with equal probability $1/d$.
- $\|\boldsymbol{x}\|_\lambda = \left(\sum_{i=1}^d |x_i|^\lambda\right)^{1/\lambda}$, $1<\lambda<\infty$. Let $X_1,\dots,X_d$ be independent and identically Fréchet-distributed rv, i.e., $P(X_i \le x) = \exp(-x^{-\lambda})$, $x>0$, $\lambda>1$. Then $\boldsymbol{Z} = (Z_1,\dots,Z_d)$ with

$$Z_i := \frac{X_i}{\Gamma(1-\frac{1}{\lambda})}, \quad i=1,\dots,d,$$

generates $\|\cdot\|_\lambda$.

Characterization of a Standard Max-Stable Distribution

A df G on $\mathbb{R}^d$ is a standard max-stable or standard extreme value df iff it is max-stable in the sense of equation (4.4), and if it has standard negative exponential margins:

$$G(0,\dots,0,x_i,0,\dots,0) = \exp(x_i), \qquad x_i \le 0,\ 1\le i\le d.$$

The theory of D-norms now allows a mathematically elegant characterization of a standard max-stable df.

Theorem 4.2.1 (Pickands (1981), de Haan and Resnick (1977), Vatan (1985)). A df G on $\mathbb{R}^d$ is a standard max-stable df $\iff$ there exists a D-norm $\|\cdot\|_D$ on $\mathbb{R}^d$ such that

$$G(\boldsymbol{x}) = \exp\left(-\|\boldsymbol{x}\|_D\right), \qquad \boldsymbol{x} \le \boldsymbol{0} \in \mathbb{R}^d.$$

Characterization of an Arbitrary Max-Stable Distribution

Any multivariate max-stable df $G_{\alpha_1,\dots,\alpha_d}$ with univariate margins $G_{\alpha_1},\dots,G_{\alpha_d}$ can be represented as

$$\begin{aligned} G_{\alpha_1,\dots,\alpha_d}(\boldsymbol{x}) &= G\left(\psi_{\alpha_1}(x_1),\dots,\psi_{\alpha_d}(x_d)\right) \\ &= \exp\left(-\left\|\left(\psi_{\alpha_1}(x_1),\dots,\psi_{\alpha_d}(x_d)\right)\right\|_D\right), \qquad \boldsymbol{x} \in \mathbb{R}^d, \end{aligned} \tag{4.5}$$

where $G(\boldsymbol{x}) = \exp\left(-\|\boldsymbol{x}\|_D\right)$, $\boldsymbol{x} \le \boldsymbol{0} \in \mathbb{R}^d$, is a standard EVD and

$$\psi_{\alpha_i}(x) = \log\left(G_{\alpha_i}(x)\right), \qquad 0 < G_{\alpha_i}(x),\ 1 \le i \le d,$$

see, e.g., Falk et al. (2011, equation (5.47)).

Pickands Dependence Function

Take an arbitrary D-norm on $\mathbb{R}^d$. We, obviously, can write for $\boldsymbol{x} \neq \boldsymbol{0} \in \mathbb{R}^d$

$$\|\boldsymbol{x}\|_D = \|\boldsymbol{x}\|_1 \left\| \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|_1} \right\|_D =: \|\boldsymbol{x}\|_1 A\left(\frac{\boldsymbol{x}}{\|\boldsymbol{x}\|_1}\right),$$

where $A(\cdot)$ is a function on the unit sphere $\boldsymbol{S} = \left\{\boldsymbol{y} \in \mathbb{R}^d : \|\boldsymbol{y}\|_1 = 1\right\}$ with respect to the norm $\|\cdot\|_1$. It is evident that it suffices to define the function $A(\cdot)$ on $\boldsymbol{S}_+ := \left\{\boldsymbol{u} \geq \boldsymbol{0} \in \mathbb{R}^{d-1} : \sum_{i=1}^{d-1} u_i \leq 1\right\}$ by putting

$$A(\boldsymbol{u}) := \left\| \left(u_1, \ldots, u_{d-1}, 1 - \sum_{i=1}^{d-1} u_i\right) \right\|_D .$$

The function $A(\cdot)$ is known as Pickands dependence function and we can represent any SMS df G as

$$\begin{aligned} G(\boldsymbol{x}) &= \exp\left(-\|\boldsymbol{x}\|_D\right) \\ &= \exp\left(\left(\sum_{i=1}^{d} x_i\right) A\left(\frac{x_1}{\sum_{i=1}^d x_i}, \ldots, \frac{x_{d-1}}{\sum_{i=1}^d x_i}\right)\right) \end{aligned}$$

and an arbitrary max-stable df correspondingly.

In particular in case $d = 2$ we obtain

$$A(u) = \|(u, 1-u)\|_D = E\left(\max(uZ_1, (1-u)Z_2)\right), \qquad 0 \leq u \leq 1,$$

with $A(0) = A(1) = 1$, $\max(u, 1-u) \leq A(u) \leq u + (1-u) = 1$. For a further analysis of the function $A(\cdot)$ we refer to Falk et al. (2011, Chapter 6).

For an appealing approach to the estimation of Pickands dependence function $A(\cdot)$ in the general case $d \geq 2$ using Bernstein polynomials we refer to Marcon et al. (2014).

Characterization of Multivariate Domain of Attraction

In complete analogy to the univariate case we say that a multivariate df F on $\mathbb{R}^d$ is in the domain of attraction of an arbitrary multivariate EVD G, again denoted by $F \in \mathcal{D}(G)$, if there are vectors $\boldsymbol{a}_n > \mathbf{0}$, $\boldsymbol{b}_n \in \mathbb{R}^d$, $n \in \mathbb{N}$, such that

$$F^n(\boldsymbol{a}_n \boldsymbol{x} + \boldsymbol{b}_n) \longrightarrow_{n\to\infty} G(\boldsymbol{x}), \qquad \boldsymbol{x} \in \mathbb{R}^d.$$

Recall: A copula on $\mathbb{R}^d$ is the df of a rv $\boldsymbol{U} = (U_1, \dots, U_d)$ with the property that each U_i follows the uniform distribution on $(0,1)$. Sklar's theorem plays a major role.

Theorem 4.2.2 (Sklar (1959, 1996)). For every df F on $\mathbb{R}^d$ there exists a copula C such that

$$F(\boldsymbol{x}) = C(F_1(x_1), \dots, F_d(x_d)), \qquad \boldsymbol{x} = (x_1, \dots, x_d) \in \mathbb{R}^d,$$

where $F_1, \dots, F_d$ are the univariate margins of F.

If F is continuous, then C is uniquely determined and given by $C(\boldsymbol{u}) = F(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d))$, $\boldsymbol{u} = (u_1, \dots, u_d) \in (0,1)^d$, where $F_i^{-1}(u) = \inf\{t \in \mathbb{R} : F_i(t) \geq u\}$, $u \in (0,1)$, is the generalized inverse of F_i.

Proposition 4.2.1 (Deheuvels (1984), Galambos (1987)). A d-variate df $F \in \mathcal{D}(G) \iff$ this is true for the univariate margins of F together with the condition that the copula C_F of F satisfies the expansion

$$C_F(\boldsymbol{u}) = 1 - \|\mathbf{1} - \boldsymbol{u}\|_D + o\left(\|\mathbf{1} - \boldsymbol{u}\|\right)$$

as $\boldsymbol{u} \to \mathbf{1}$, uniformly for $\boldsymbol{u} \in [0,1]^d$, where $\|\cdot\|_D$ is the D-norm on $\mathbb{R}^d$ that corresponds to G in the sense of equation (4.5).

Idea: Skip the $o(\|\mathbf{1} - \boldsymbol{u}\|)$-term.
Problem:

Is $C(\boldsymbol{u}) := 1 - \|\mathbf{1} - \boldsymbol{u}\|_D$, $\boldsymbol{u} \in [0,1]^d$, a copula?

Answer:

Only in dimension $d \in \{1, 2\}$[2].

Multivariate Generalized Pareto Distributions

A d-dimensional df W is called a multivariate GPD iff there exists a d-dimensional EVD G and $\boldsymbol{x}_0 \in \mathbb{R}^d$ with $G(\boldsymbol{x}_0) < 1$ such that

$$W(\boldsymbol{x}) = 1 + \log(G(\boldsymbol{x})), \qquad \boldsymbol{x} \geq \boldsymbol{x}_0. \tag{4.6}$$

[2] Michel (2008, Theorem 6)

Note: $1+\log(G(\boldsymbol{x}))$, $G(\boldsymbol{x}) \geq 1/e$, **does not define a df in general unless** $d \in \{1,2\}$, **see above.**

For a standard max-stable df $G(\boldsymbol{x}) = \exp\left(-\|\boldsymbol{x}\|_D\right)$, $\boldsymbol{x} \leq \mathbf{0} \in \mathbb{R}^d$, **we obtain**

$$W(\boldsymbol{x}) = 1+\log(G(\boldsymbol{x})) = 1-\|\boldsymbol{x}\|_D\,, \qquad \boldsymbol{x} \in [x_0, 0]^d.$$

Note: Each univariate margin $W_i(x) = 1+x$, $x_0 \leq x \leq 0$, **is the df of a uniform distribution on** $[x_0, 1]$.

Domain of Attraction for Copulas

Each univariate margin of an arbitrary copula is the uniform distribution on $(0,1)$. **Its df is** $F_U(u) = u$, $u \in [0,1]$. **We, therefore, obtain with** $a_n = 1/n$, $b_n = 1$, $n \in \mathbb{N}$,

$$\begin{aligned} F_U^n\left(a_n x + b_n\right) &= F_U^n\left(\frac{x}{n}+1\right) \\ &= \left(1+\frac{x}{n}\right)^n \qquad \textbf{if } n \textbf{ is large} \\ &\longrightarrow_{n\to\infty} \exp(x), \qquad x \leq 0, \end{aligned}$$

i.e., each univariate margin of an arbitrary copula is automatically in the domain of attraction of the EVD $G(x) = \exp(x)$, $x \leq 0$.

Replacing in the preceding Proposition 4.2.1 the df F **by a copula** C **immediately yields the following characterization.**

Corollary 4.2.1. A copula $C \in \mathcal{D}(G) \iff$

$$C(\boldsymbol{u}) = 1 - \|\boldsymbol{1} - \boldsymbol{u}\|_D + o(\|\boldsymbol{1} - \boldsymbol{u}\|)$$

as $\boldsymbol{u} \to \boldsymbol{1}$, uniformly for $\boldsymbol{u} \in [0,1]^d$.

Message: A copula $C(u)$ can reasonably be approximated for u close to 1 only by a shifted GPD $W(\boldsymbol{u} - \boldsymbol{1}) = 1 - \|\boldsymbol{1} - \boldsymbol{u}\|_D$.

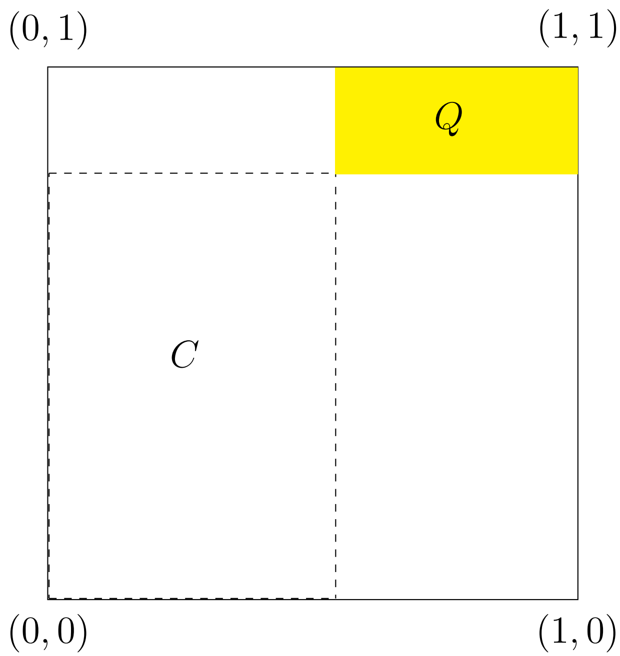
This message has the following implication for risk assessment: If you want to model the copula underlying multivariate data above some high threshold u_0, you should try a GPD copula, which is given in its upper tail by

$$Q(\boldsymbol{u}) = 1 - \|\boldsymbol{1} - \boldsymbol{u}\|_D, \qquad \boldsymbol{u}_0 \le \boldsymbol{u} \le \boldsymbol{1},$$

where $\|\cdot\|_D$ is a D-norm.

Multivariate Piecing-Together

It is possible to cut off the upper tail of an arbitrary copula C and to substitute it by a GPD copula as above such that the result is again a copula, see Aulbach et al. (2012a,b):



Multivariate Piecing Together

4.3 Extreme Value Copulas et al.

Extreme Value Copulas

An extreme value copula on $\mathbb{R}^d$ is the copula of an arbitrary d-variate max-stable df G^*. It has by equation (4.5) the representation

$$C_{G^*}(\boldsymbol{u}) = \exp\left(-\left\|(\log(u_1), \dots, \log(u_d))\right\|_D\right), \qquad \boldsymbol{u} \in (0,1]^d,$$

and, thus, we have by elementary arguments the following equivalences:

A copula C_F is in the max-domain of attraction of a standard max-stable df G

$\iff$ $C_F(\boldsymbol{u}) = 1 - \|\boldsymbol{1} - \boldsymbol{u}\|_D + o(\|\boldsymbol{1} - \boldsymbol{u}\|)$, $\qquad \boldsymbol{u} \in [0,1]^d$,

$\iff$ $\lim_{t\downarrow 0} \frac{1 - C_F(\boldsymbol{1} + t\boldsymbol{x})}{t} = \ell_{G^*}(\boldsymbol{x})$, $\qquad \boldsymbol{x} \le \boldsymbol{0} \in \mathbb{R}^d$,

with $\ell_{G^*}(\boldsymbol{x}) := -\log(C_{G^*}(\exp(\boldsymbol{x}))) = \|\boldsymbol{x}\|_D$, $\boldsymbol{x} \le \boldsymbol{0}$, known as the stable tail dependence function (Huang (1992)) of G^*. This opens the way to estimate an underlying D-norm by using estimators of the stable tail dependence function[3].

Example 4.3.1. Take an arbitrary Archimedean copula

$$C_\varphi(\boldsymbol{u}) = \varphi^{-1}(\varphi(u_1) + \cdots + \varphi(u_m)),$$

(McNeil and Nešlehová (2009, Theorem 2.2). If φ is differentiable from the left in $x = 1$ with left derivative $\varphi'(1-) \neq 0$, then

$$\lim_{t\downarrow 0} \frac{1 - C_\varphi(\boldsymbol{1} + t\boldsymbol{x})}{t} = \sum_{i \le m} |x_i| = \|\boldsymbol{x}\|_1, \qquad \boldsymbol{x} \le \boldsymbol{0} \in \mathbb{R}^m,$$

$\implies$ $C_\varphi \in \mathcal{D}(G)$ with $G(\boldsymbol{x}) = \exp(-\|\boldsymbol{x}\|_1)$, $\boldsymbol{x} \le \boldsymbol{0}$, having independent margins $\implies$ margins of C_φ are tail independent.

[3]http://www.actuaries.org/ASTIN/Colloquia/Hague/Presentations/Krajina.pdf

This concerns Clayton, Frank copula, but not the Gumbel copula with generator $\varphi_G(t) = (-\log(t))^\lambda$, $0 < t \leq 1$, $\lambda > 1$.

4.4 A Measure of Tail Dependence

The Extremal Coefficient

To measure the dependence among the univariate margins by just one number, Smith (1990) introduced the extremal coefficient as that constant $\varepsilon > 0$ which satisfies

$$G^*(x, \dots, x) = F^\varepsilon(x), \qquad x \in \mathbb{R},$$

where G^* is an arbitrary d-dimensional max-stable df with identical margins F.

If $\varepsilon = d$ we have independence of the margins, if $\varepsilon = 1$ we have complete dependence.

Question: Can we characterize this ε? Does it exist at all?

Without loss of generality we can transform as in equation (4.5) the margins of G^* to the standard negative exponential distribution $\exp(x)$, $x \leq 0$, thus obtaining a standard max-stable df G and, threfore,

$$G(x, \dots, x) = \exp(-\|(x, \dots, x)\|_D) = \exp(x\|\mathbf{1}\|_D) = \exp(x)^{\|\mathbf{1}\|_D},$$

$x \leq 0$**, yielding**

$$\varepsilon = \|\mathbf{1}\|_D .$$

The extremal coefficient is, therefore, the D-norm of the vector 1.

If a df F is in the domain of attraction of an arbitrary multivariate EVD G^* with corresponding D-norm as in equation (4.5), then $\varepsilon = \|1\|_D$ is a measure of the tail

dependence of F. **This is a crucial measure for assessing the risk inherent in a portfolio etc.**

4.5 Takahashi's Theorem

Takahashi's Theorem for D-Norms

The following result can easily be established by elementary arguments (see Theorem 1.3.1).

Theorem 4.5.1 (Takahashi (1988)). We have for an arbitrary D-norm $\|\cdot\|_D$ on $\mathbb{R}^d$:

(i) $\|\cdot\|_D = \|\cdot\|_1 \iff \exists\, \boldsymbol{y} > \boldsymbol{0}: \ \|\boldsymbol{y}\|_D = \|\boldsymbol{y}\|_1$,

(ii) $\|\cdot\|_D = \|\cdot\|_\infty \iff \|\boldsymbol{1}\|_D = 1$.

Consequence: The margins of a multivariate EVD are independent iff this is true for at least one point. They are completely dependent if they are dependent at one point.

The next result can easily be established as well (see Theorem 1.3.3).

Theorem 4.5.2. Let $\|\cdot\|_D$ be an arbitrary D-norm on $\mathbb{R}^d$ and denote by $\boldsymbol{e}_i := (0,\dots,0,1,0,\dots,0)$ the i-th unit vector in $\mathbb{R}^d$. We have

$$\|\cdot\|_D = \|\cdot\|_1 \iff \|\boldsymbol{e}_i + \boldsymbol{e}_j\|_D = 2,\ 1 \le i \ne j \le d.$$

Speaking in terms of multivariate EVD, the preceding result states: The margins of an arbitrary multivariate EVD are independent iff they are pairwise independent.

4.6 Some General Remarks on D-Norms

- The generator $\boldsymbol{Z}$ of a D-norm $\|\cdot\|_D$ is in general not uniquely determined, even its distribution is not.
- We have the bounds
$$\|\cdot\|_\infty \le \|\cdot\|_D \le \|\cdot\|_1$$
for an arbitrary D-norm; $\|\cdot\|_\infty$, $\|\cdot\|_1$ are D-norms themselves.
- The index D means dependence:
$$G(\boldsymbol{x}) = \exp\left(-\|\boldsymbol{x}\|_\infty\right) = \text{complete dependence of the margins of } G$$
$$G(\boldsymbol{x}) = \exp\left(-\|\boldsymbol{x}\|_1\right) = \text{independence of the margins of } G.$$

Copulas as Generators of a D-Norm

By the way, talking about dependence: Let the rv $\boldsymbol{U} = (U_1, \dots, U_d)$ follow an arbitrary copula on $\mathbb{R}^d$, i.e., each U_i is on $(0,1)$ uniformly distributed. Then
$$\boldsymbol{Z} := 2\boldsymbol{U}$$
is obviously the generator of a D-norm.

Not each D-norm can be generated this way: The bivariate D-norm $\|\cdot\|_1$ cannot. There are, consequently, strictly more D-norms than copulas.

4.7 Functional $\mathbb{D}$-Norm

Denote by $E[0,1]$ the set of functions $f:[0,1]\to\mathbb{R}$ that are bounded and have only a finite number of discontinuities. This is obviously a linear space. By $C[0,1]$ we denote the subset of continuous functions.

Generator of a Functional D-Norm

Let $\boldsymbol{Z}=(Z_t)_{t\in[0,1]}$ be a stochastic process with continuous sample paths, i.e., $\boldsymbol{Z}\in C[0,1]$, with the additional properties

$$Z_t\ge 0,\quad E(Z_t)=1,\qquad t\in[0,1],$$

and

$$E\left(\sup_{t\in[0,1]} Z_t\right)<\infty.$$

Then

$$\|f\|_D := E\left(\sup_{t\in[0,1]} \left(|f(t)|\, Z_t\right)\right),\qquad f\in E[0,1],$$

defines a norm on $E[0,1]$, called D-norm, with generator $\boldsymbol{Z}$.

Max-Stable Processes

Let $\boldsymbol{\eta} = (\eta_t)_{t\in[0,1]}$ be a stochastic process in $C[0,1]$, with the additional property that each component η_t follows the standard negative exponential distribution $\exp(x)$, $x \le 0$. The following result, which goes back to Giné et al. (1990), can now be formulated in terms of the functional D-norm:

Theorem 4.7.1. A process $\boldsymbol{\eta}$ as above is max-stable $\iff$ there exists a D-norm $\|\cdot\|_D$ on $E[0,1]$ such that

$$P(\boldsymbol{\eta} \le f) = \exp\left(-\|f\|_D\right), \qquad f \in E^-[0,1].$$

We call a max-stable process η as above standard max-stable (SMS). It, obviously, satisfies with 1 denoting the constant function 1 on $[0,1]$

$$\begin{aligned} & P\left(\sup_{t\in[0,1]} \eta_t \le x\right) \\ &= P\left(\eta_t \le x,\, t \in [0,1]\right) \\ &= P(\boldsymbol{\eta} \le x\mathbf{1}) \\ &= \exp\left(-\|x\mathbf{1}\|_D\right) \\ &= \exp\left(x\|\mathbf{1}\|_D\right), \qquad x \le 0, \end{aligned}$$

i.e., the rv $X := \sup_{t\in[0,1]} \eta_t$ is negative exponential distributed

$$P(X \le x) = \exp(x/\vartheta), \qquad x \le 0,$$

with parameter $\vartheta = 1/\left\|\mathbf{1}\right\|_D$. As a consequence we obtain in particular

$$\begin{aligned}
&P(\eta_t = 0 \text{ for some } t \in [0,1]) \\
&= P\left(\sup_{t\in[0,1]} \eta_t = 0\right) \\
&= 1 - P(X < 0) \\
&= 1 - P(X \le 0) \\
&= 0.
\end{aligned}$$

We can now put

$$\boldsymbol{\xi} := \frac{1}{\boldsymbol{\eta}}.$$

The process $\boldsymbol{\xi} = (\xi_t)_{t\in[0,1]}$ has continuous sample paths, each margin ξ_t is standard Fréchet distributed

$$P(\xi_t \le y) = P\left(\eta_t \le -\frac{1}{y}\right) = \exp\left(-\frac{1}{y}\right), \qquad y > 0,$$

and we have for $g \in E[0,1]$, $g > 0$,

$$P(\boldsymbol{\xi} \le g) = P\left(\boldsymbol{\eta} \le -\frac{1}{g}\right) = \exp\left(-\left\|\frac{1}{g}\right\|_D\right).$$

The process $\boldsymbol{\xi}$ is, consequently, max-stable as well. It is called simple max-stable in the literature.

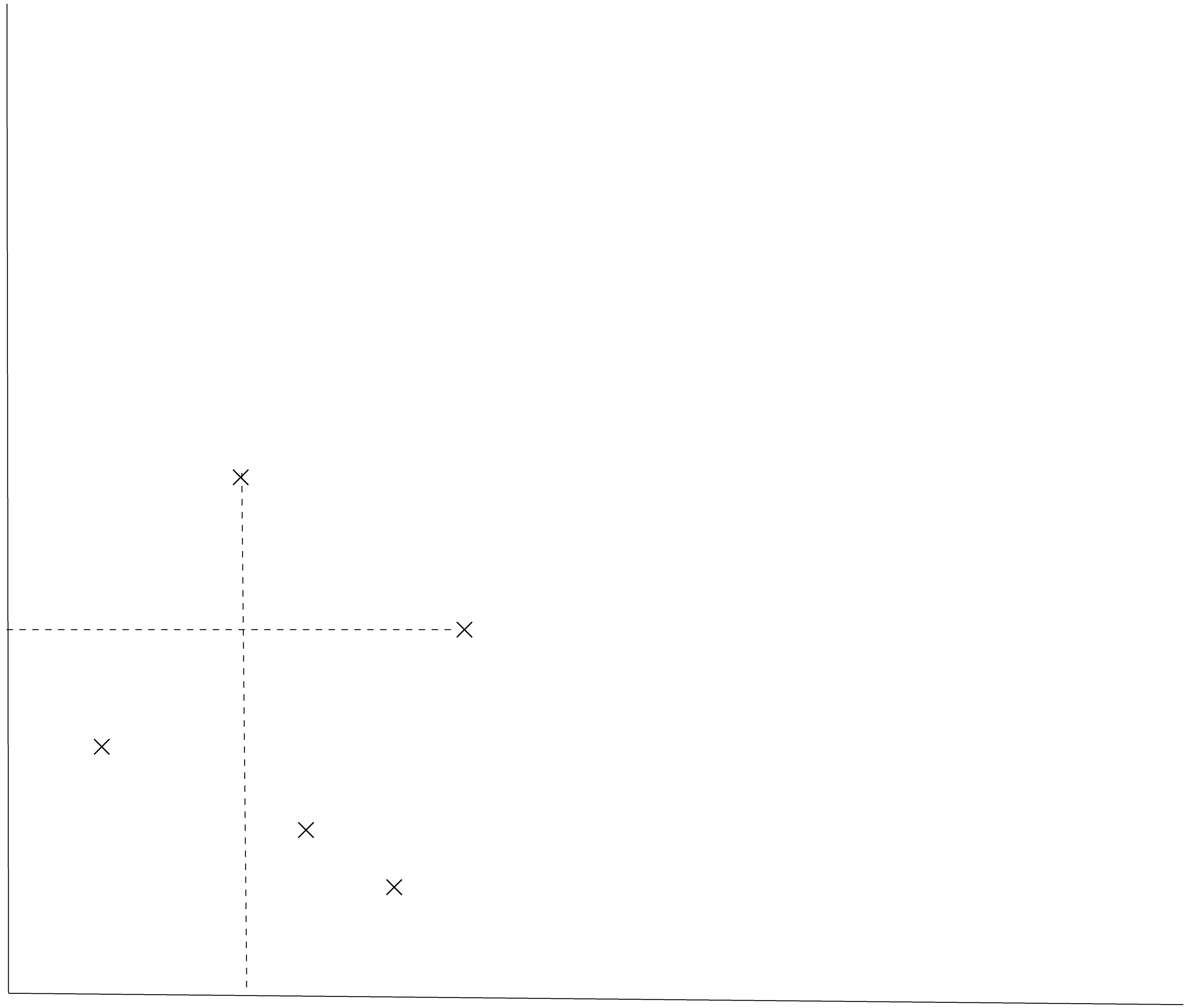
4.8 Some Quite Recent Results on Multivariate Records

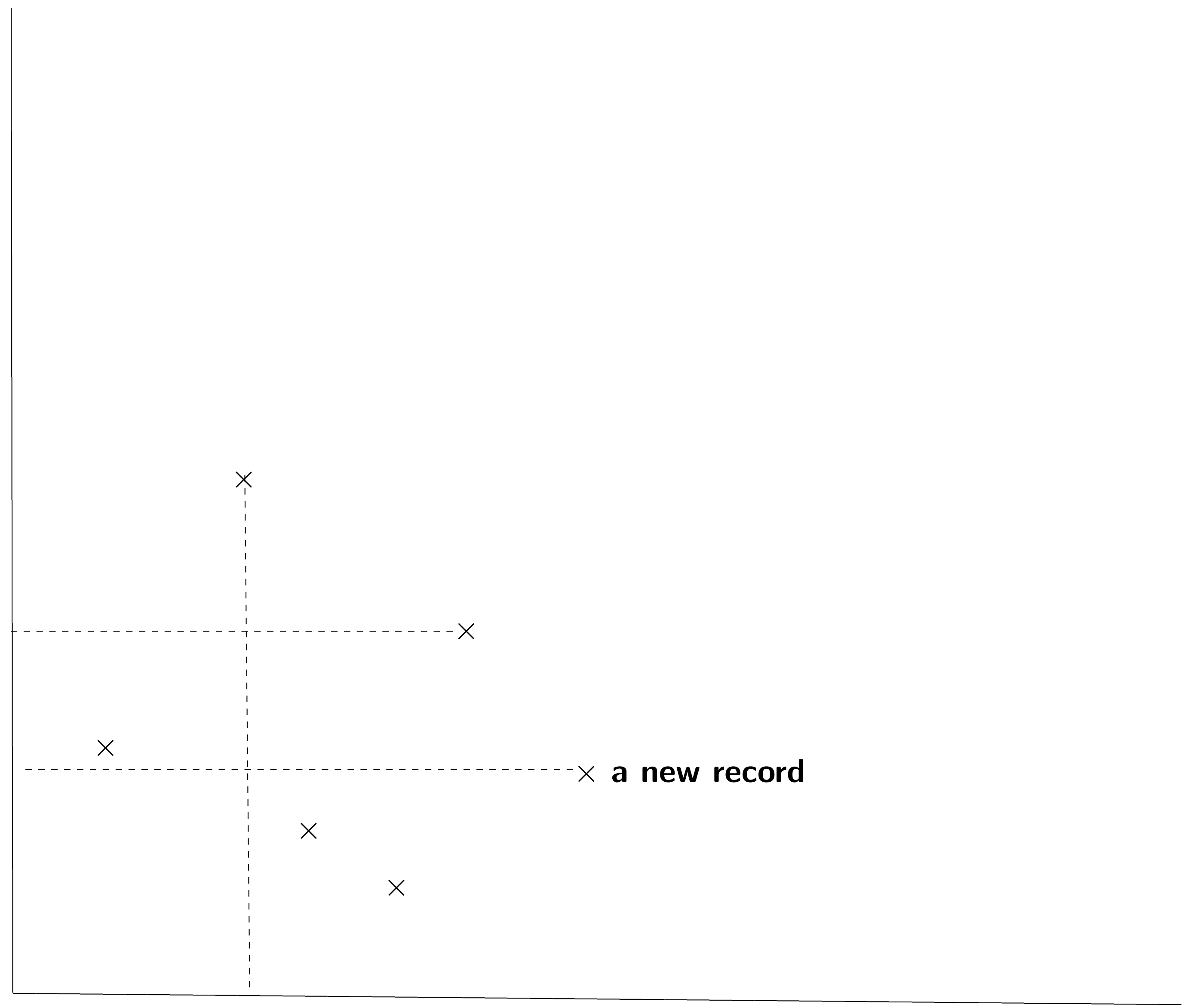
Multivariate Records

The subsequent results are joint work with Clément Dombry and Maximilian Zott (Dombry et al. (2015)). Let $\boldsymbol{X}_1, \boldsymbol{X}_2, \dots$ be independent copies of a rv $\boldsymbol{X} \in \mathbb{R}^d$. We say that $\boldsymbol{X}_k$ is a (multivariate simple) record, if

$$\boldsymbol{X}_k \not\leq \max_{1 \leq i \leq k-1} \boldsymbol{X}_i,$$

i.e., if at least one component of $\boldsymbol{X}_k$ is strictly larger than the corresponding components of $\boldsymbol{X}_1, \dots, \boldsymbol{X}_{k-1}$.




a new record

Record Times

We denote by $N(n)$, $n \geq 1$, the record times, i.e., those subsequent random indices at which a record occurs. Precisely, $N(1) = 1$, as $\boldsymbol{X}_1$ is, clearly, a record, and, for $n \geq 2$,

$$N(n) := \min\left\{j : \; j > N(n-1), \; \boldsymbol{X}_j \not\leq \max_{1 \leq i \leq N(n-1)} \boldsymbol{X}_i\right\}.$$

As the df F is continuous, the distribution of $N(n)$ does not depend on F and, therefore, we assume in what follows without loss of generality that F is a copula C on $\mathbb{R}^d$, i.e., each component of $\boldsymbol{X}_i$ is on $(0,1)$ uniformly distributed.

Expectation of Record Time

We have for $j \geq 2$

$$P(N(2) = j) = \int_{[0,1]^d} C(\boldsymbol{u})^{j-2}(1 - C(\boldsymbol{u}))\, C(d\boldsymbol{u})$$

and, thus,

$$E(N(2)) = \int_{[0,1]^d} \frac{C(\boldsymbol{u})}{1 - C(\boldsymbol{u})}\, C(d\boldsymbol{u}) + 2.$$

Suppose now that $d = 1$. Then we have $\boldsymbol{u} = u \in [0,1]$, $C(u) = u$ and

$$E(N(2)) = \int_0^1 \frac{u}{1-u}\, du + 2 = \infty,$$

which is well-known (Galambos (1987, Theorem 6.2.1)). Because $N(n) \geq N(2)$, $n \geq 2$, we have $E(N(n)) = \infty$ for $n \geq 2$ as well.

Suppose next that $d \geq 2$ and that the margins of C are independent, i.e.,

$$C(\boldsymbol{u}) = \prod_{i=1}^{d} u_i, \qquad \boldsymbol{u} = (u_1, \dots, u_d) \in [0,1]^d.$$

Then we obtain

$$\int_{[0,1]^d} \frac{C(\boldsymbol{u})}{1 - C(\boldsymbol{u})} C(d\boldsymbol{u}) = \int_0^1 \cdots \int_0^1 \frac{\prod_{i=1}^d u_i}{1 - \prod_{i=1}^d u_i} du_1 \dots du_d < \infty$$

by elementary arguments and, thus, $E(N(2)) < \infty$. This observation gives rise to the problem to characterize those copulas C on $[0,1]^d$ with $d \geq 2$, such that $E(N(2))$ is finite. Note that $E(N(2)) = \infty$ if the components of C are completely dependent.

Characterization of Finite Expectation

Lemma 4.8.1. We have $E(N(2)) < \infty$ iff

$$\int_1^\infty P\left(X_i \geq \frac{1}{t}, 1 \leq i \leq d\right) dt < \infty. \tag{4.7}$$

Dual D-Norm Function

Let $\|\cdot\|_D$ be an arbitrary D-norm on $\mathbb{R}^d$ with arbitrary generator $\boldsymbol{Z} = (Z_1, \dots, Z_d)$. Put

$$\wr\, \boldsymbol{x} \,\wr_D := E\left(\min_{1\le i\le \in T}(|x_i|\, Z_i)\right), \qquad \boldsymbol{x} \in \mathbb{R}^d,$$

which we call the dual D-norm function corresponding to $\|\cdot\|_D$. It is independent of the particular generator $\boldsymbol{Z}$, but the mapping

$$\|\cdot\|_D \to \wr \cdot \wr_D$$

is not one-to-one. In particular we have that

$$\wr \cdot \wr_D = 0$$

is the least dual D-norm function, corresponding to $\|\cdot\|_D = \|\cdot\|_1$, and

$$\wr\, \boldsymbol{x} \,\wr_D = \min_{1\le i\le d} |x_i| = \wr\, \boldsymbol{x} \,\wr_\infty, \qquad \boldsymbol{x} \in \mathbb{R}^d,$$

is the largest dual D-norm function, corresponding to $\|\cdot\|_D = \|\cdot\|_\infty$, i.e., we have for an arbitrary dual D-norm function the bounds

$$0 = \wr \cdot \wr_1 \le \wr \cdot \wr_D \le \wr \cdot \wr_\infty.$$

While the first inequality is obvious, the second one follows from

$$|x_k| = E(|x_k|\, Z_k) \ge E\left(\min_{1\le i\le d}(|x_i|\, Z_i)\right), \qquad 1 \le k \le d.$$

Expansion of Survival Function of $C \in \mathcal{D}(G)$ via Dual D-Norm Function

We obtain the following consequence.

Lemma 4.8.2. Let G be a sms df with corresponding D-norm $\|\cdot\|_D$. Then we have for an arbitrary copula C the implication

$$C \in \mathcal{D}(G) \implies P(\boldsymbol{X} \geq \boldsymbol{u}) = \wr\, \boldsymbol{1} - \boldsymbol{u} \,\wr_D + o(\|\boldsymbol{1} - \boldsymbol{u}\|) \tag{4.8}$$

as $\boldsymbol{u} \to \boldsymbol{1}$, uniformly for $\boldsymbol{u} \in [0,1]^d$, where $\boldsymbol{X}$ is a rv whose df is C.

Note that the reverse implication in the preceding result does not hold, as the mapping $\|\cdot\|_D \to \wr \cdot \wr_D$ is not one to one.

Infinite Expectation of Record Time

Proposition 4.8.1. Suppose that $C \in \mathcal{D}(G)$, where the D-norm corresponding to G satisfies $\wr\, \boldsymbol{1} \,\wr_D > 0$. Then $E(N(2)) = \infty$.

Another Tail Dependence Coefficient

Within the class of (bivariate) copula that are tail independent,

$$\bar{\chi} := \lim_{u \uparrow 1} \frac{2\log(1-u)}{\log(P(X_1 > u, X_2 > u))} - 1$$

is a popular measure of tail comparison, provided this limit exists (Coles et al. (1999); Heffernan (2000)). In this case we have $\bar{\chi} \in [-1, 1]$, cf. Beirlant et al. (2004, (9.83)). For a bivariate normal copula with coefficient of correlation $\rho \in (-1, 1)$ it is, for instance, well known that $\bar{\chi} = \rho$.

Proposition 4.8.2. Let $\boldsymbol{X} = (X_1, \ldots, X_d)$ follow a copula C in $\mathbb{R}^d$ with $C \in \mathcal{D}(G)$ and G having independent margins. Suppose that there exist indices $k \neq j$ such that

$$\bar{\chi}_{k,j} = \lim_{u \uparrow 1} \frac{2\log(1-u)}{\log(P(X_k > u, X_i > u))} - 1 \in (-1, 1).$$

Then we have $E(N(2)) < \infty$.

Corollary 4.8.1. We have $E(N(2)) < \infty$ for multivariate normal rv unless all components are completely dependent.

Index

Bibliography

Aulbach, S., Bayer, V., and Falk, M. (2012a). A multivariate piecing-together approach with an application to operational loss data. Bernoulli 18, 455–475. doi:10.3150/10-BEJ343.

Aulbach, S., Falk, M., and Hofmann, M. (2012b). The multivariate piecing-together approach revisited. J. Multivariate Anal. 110, 161–170. doi:10.1016/j.jmva.2012.02.002.

Balkema, A. A., and Resnick, S. I. (1977). Max-infinite divisibility. J. Appl. Probab. 14, 309–319. doi:10.2307/3213001.

Bauer, H. (2001). Measure and Integration Theory, De Gruyter Studies in Mathematics, vol. 26. De Gruyter, Berlin.

Beirlant, J., Goegebeur, Y., Segers, J., and Teugels, J. (2004). Statistics of Extremes: Theory and Applications. Wiley Series in Probability and Statistics. Wiley, Chichester, UK. doi:10.1002/0470012382.

Bolley, F. (2008). Separability and completeness for the Wasserstein distance. In Séminaire de Probabilités XLI (C. Donati-Martin, M. Émery, A. Rouault, and

C. Stricker, eds.), Lecture Notes in Mathematics, vol. 1934, 371–377. Springer, Berlin. doi:10.1007/978-3-540-77913-1_17.

Brown, B. M., and Resnick, S. I. (1977). Extreme values of independent stochastic processes. J. Appl. Probab. 14, 732–739. doi:10.2307/3213346.

Coles, S. G., Heffernan, J. E., and Tawn, J. A. (1999). Dependence measure for extreme value analyses. Extremes 2, 339–365.

Deheuvels, P. (1984). Probabilistic aspects of multivariate extremes. In Statistical Extremes and Applications (J. Tiago de Oliveira, ed.), 117–130. D. Reidel, Dordrecht.

Dombry, C., Falk, M., and Zott, M. (2015). On functional records and champions. Tech. Rep., Université de Franche-Comté and University of Würzburg. arXiv:1510.04529 [math.PR].

Dombry, C., and Ribatet, M. (2015). Functional regular variations, pareto processes and peaks over threshold. In Special Issue on Extreme Theory and Application (Part II) (Y. Wang and Z. Zhang, eds.), Statistics and Its Interface, vol. 8, 9–17. doi:10.4310/SII.2015.v8.n1.a2.

Falk, M., Hüsler, J., and Reiss, R.-D. (2011). Laws of Small Numbers: Extremes and Rare Events. 3rd ed. Springer, Basel. doi:10.1007/978-3-0348-0009-9.

Ferreira, A., and de Haan, L. (2014). The generalized Pareto process; with a view towards application and simulation. Bernoulli 20, 1717–1737. doi:10.3150/13-BEJ538.

Galambos, J. (1987). The Asymptotic Theory of Extreme Order Statistics. 2nd ed. Krieger, Malabar.

Giné, E., Hahn, M., and Vatan, P. (1990). Max-infinitely divisible and max-stable sample continuous processes. Probab. Theory Related Fields 87, 139–165. doi:10.1007/BF01198427.

Gnedenko, B. (1943). Sur la distribution limite du terme maximum d'une série aléatoire. Ann. of Math. (2) 44, 423–453. doi:10.2307/1968974.

de Haan, L. (1975). On regular variation and its application to the weak convergence of sample extremes, MC Tracts, vol. 32. 3rd ed. Centrum Voor Wiskunde en Informatica, Amsterdam, http://persistent-identifier.org/?identifier=urn:nbn:nl:ui:18-18567.

de Haan, L., and Ferreira, A. (2006). Extreme Value Theory: An Introduction. Springer Series in Operations Research and Financial Engineering. Springer, New York. doi:10.1007/0-387-34471-3. See http://people.few.eur.nl/ldehaan/EVTbook.correction.pdf and http://home.isa.utl.pt/~anafh/corrections.pdf for corrections and extensions.

de Haan, L., and Resnick, S. (1977). Limit theory for multivariate sample extremes. Probab. Theory Related Fields 40, 317–337. doi:10.1007/BF00533086.

Heffernan, J. E. (2000). A directory of coefficients of tail dependence. Extremes 3, 279–290.

Huang, X. (1992). Statistics of Bivariate Extreme Values. Ph.D. thesis, Tinbergen Institute Research Series.

Marcon, G., Padoan, S. A., Naveau, P., and Muliere, P. (2014). Multivariate nonparametric estimation of the pickands dependence function using bernstein polynomials. Tech. Rep. arXiv:1405.5228 [math.PR].

McNeil, A. J., and Nešlehová, J. (2009). Multivariate archimedean copulas, d-monotone functions and ℓ_1-norm symmetric distributions. Ann. Statist. 37, 3059–3097. doi:10.1214/07-AOS556.

Michel, R. (2008). Some notes on multivariate generalized Pareto distributions. J. Multivariate Anal. 99, 1288–1301. doi:10.1016/j.jmva.2007.08.007.

Ng, K. W., Tian, G.-L., and Tang, M.-L. (2011). Dirichlet and Related Distributions. Theory, Methods and Applications. Wiley Series in Probability and Statistics. Wiley, Chichester, UK. doi:10.1002/9781119995784.

Pickands, J., III (1981). Multivariate extreme value distributions. Proc. 43th Session ISI (Buenos Aires) 859–878.

Rudin, W. (1976). Principles of Mathematical Analysis. International Student Edition, 3rd ed. McGraw-Hill International, Tokyo.

Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. Pub. Inst. Stat. Univ. Paris 8, 229–231.

Sklar, A. (1996). Random variables, distribution functions, and copulas – a personal look backward and forward. In Distributions with fixed marginals and related topics (L. Rüschendorf, B. Schweizer, and M. D. Taylor, eds.), Lecture Notes – Monograph Series, vol. 28, 1–14. Institute of Mathematical Statistics, Hayward, CA. doi:10.1214/lnms/1215452606.

Smith, R. L. (1990). Max-stable processes and spatial extremes. Preprint, Univ. North Carolina, http://www.stat.unc.edu/faculty/rs/papers/RLS_Papers.html.

Takahashi, R. (1988). Characterizations of a multivariate extreme value distribution. Adv. in Appl. Probab. 20, 235–236. doi:10.2307/1427279.

Vatan, P. (1985). Max-infinite divisibility and max-stability in infinite dimensions. In Probability in Banach Spaces V: Proceedings of the International Conference held in Medford, USA, July 1627, 1984 (A. Beck, R. Dudley, M. Hahn, J. Kuelbs, and M. Marcus, eds.), Lecture Notes in Mathematics, vol. 1153, 400–425. Springer, Berlin. doi:10.1007/BFb0074963.

Villani, C. (2009). Optimal Transport: Old and New, Grundlehren der mathematischen Wissenschaften, vol. 338. Springer, Berlin. doi:10.1007/978-3-540-71050-9.