UNSW SYDNEY

MATH5805 – A Modern Introduction to Extreme Value Theory

Week 7 - Multivariate Extremes: the componentwise maxima (I)

Dr Boris Béranger
School of Mathematics and Statistics
UNSW Sydney

CRICOS Provider Code 00098G

Lecture outline

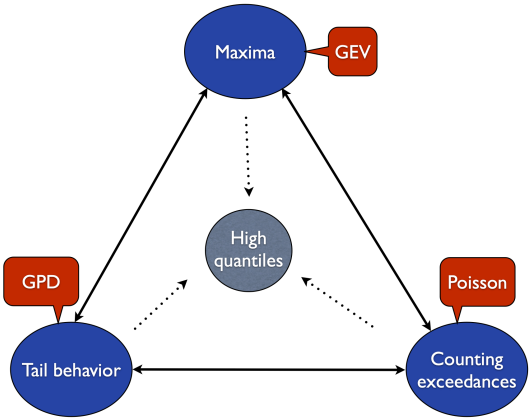
1. Introduction
2. Limit laws
3. Quantifying and representing extremal dependence
4. Parametric models for extremal dependence

1. Introduction



Goals

In **Week 2-5** we have seen several examples of univariate analyses using the block-maxima, peaks-over-threshold and Poisson point process approaches.



Goals

We have seen several ways to define an extreme value in the univariate setting, we will see that it is even more complex in the multivariate context!



This lecture is about introducing **Multivariate Extremes**. In particular, we

- Extend the Block Maxima approach to more than one random variables
- Characterise the dependence structure between the multiple extremes

Note that throughout this lecture we will make the assumption that $\{\mathbf{X}_i = (X_{i,1}, \ldots, X_{i,d}), i = 1, \ldots, n\}$ is a sequence of i.i.d. d-dimensional random variables with **common joint distribution function** F and **marginal distributions** $F_1, \ldots, F_d$.

Motivation

Many practical problems in finance, the environment, etc., are **high-dimensional by nature**. E.g., when analysing the air quality in an area, the amount of pollution depends on the levels of different pollutants and the interaction between them.

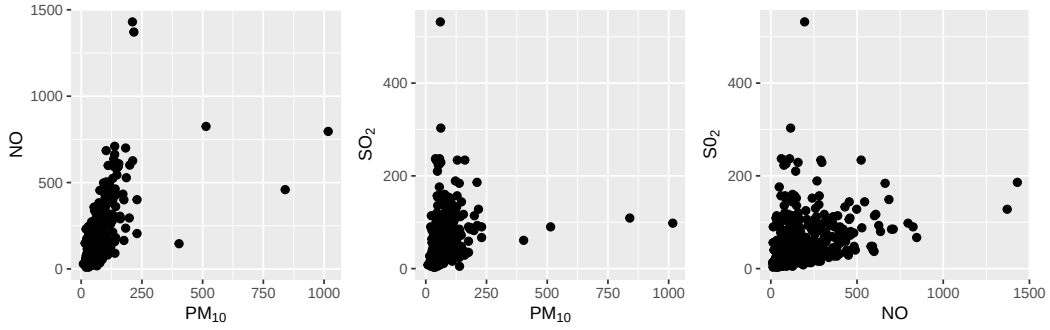
Example - The 2003 European heatwave One of the hottest summers on record in Europe, especially in France. Temperatures exceeded 40° in many places in northern Europe, and reach almost 50° in Portugal!

- It led to health crises in several countries and combined with drought to create a crop shortfall in Southern Europe.
- It has been linked to forest fires in Portugal, to dried-up rivers, to buckling of train track in the UK, and to flash floods in the Alps.
- Over 35,000 more people than usual are thought to have died in Europe that summer.

Motivation

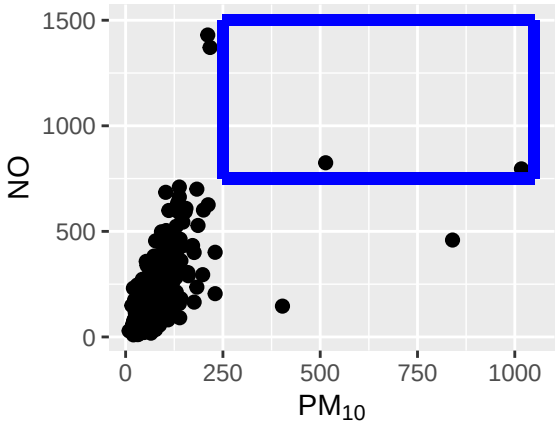
Example - Pollution in Leeds, UK

Air pollutants (PM_{10}, NO and SO_2) levels in Leeds, UK. Data recorded over the winter period (Nov-Feb) between 1994 and 1998.



Example - Pollution in Leeds, UK

What is the probability of observing data in the **blue** box?



Defining an extremes 'per block'

Example - $d = 2$

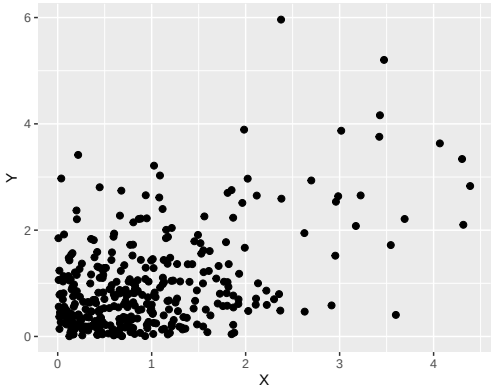
Consider the daily concentration level of **two** air pollutants X and Y in 2024 ($n \times m = 365$).

We lose the temporal element in the scatterplot!



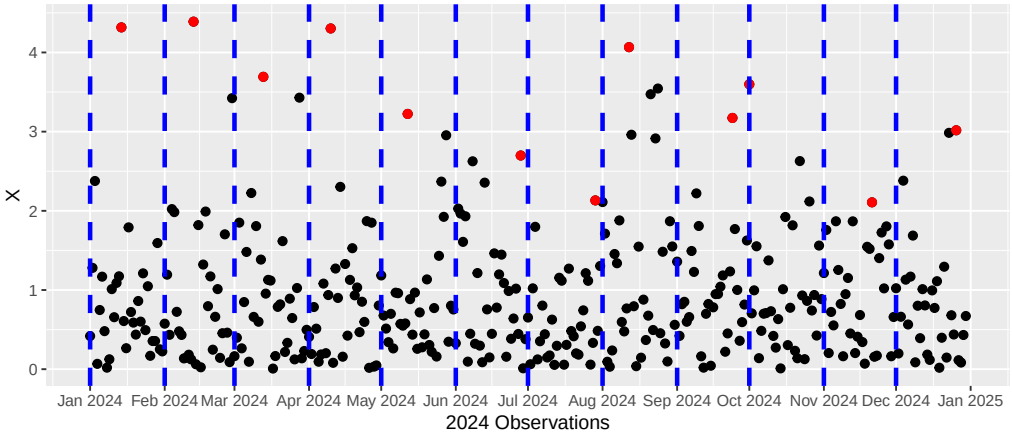
How to define bivariate blocks?

Short answer: We can't!



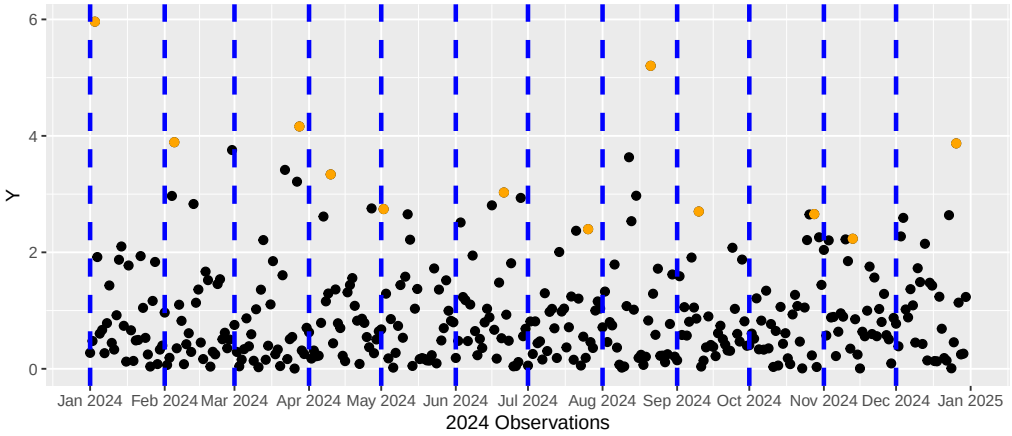
Defining an extreme ‘per block’

Simulated daily levels of air pollutant X in 2024 with monthly maxima.



Defining an extreme ‘per block’

Simulated daily levels of air pollutant Y in 2024 with monthly maxima.



Defining an extreme 'per block'

We define **marginal blocks**, $M_{n,j} = \max_{i=1,\ldots,n} X_{i,j}$ and the associated **vector of componentwise maxima**

$$\boldsymbol{M_n} = (M_{n,1}, \ldots, M_{n,d}).$$



This may not correspond to an actual observation (or be physically possible)!

Example - $d = 2$

In our simulated example we have $\boldsymbol{M_n} = (M_{n,x}, M_{n,y})$ and the observed data matrix is

$$\begin{pmatrix} \max_{i \in \text{Jan 2024}} X_i & \max_{i \in \text{Jan 2024}} Y_i \\ \vdots & \vdots \\ \max_{i \in \text{Dec 2024}} X_i & \max_{i \in \text{Dec 2024}} Y_i \end{pmatrix}.$$

Note e.g., $(\max_{i \in \text{Jan 2024}} X_i, \max_{i \in \text{Jan 2024}} Y_i)$ does not correspond to a (X_i, Y_i).

Defining an extreme 'per block'

Assuming n is finite, if the joint distribution F of $\mathbf{X}$ is known then

$$\begin{aligned} F_{\boldsymbol{M_n}}(\mathbf{x}) = \mathbb{P}(\boldsymbol{M_n} \leq \mathbf{x}) &= \mathbb{P}\left(\max_{i=1,\ldots,n} X_{i,1} \leq x_1, \ldots, \max_{i=1,\ldots,n} X_{i,d} \leq x_d\right) \\ &= \mathbb{P}(\mathbf{X}_1 \leq \mathbf{x}, \ldots, \mathbf{X}_n \leq \mathbf{x}) \\ &= F^n(\mathbf{x}). \end{aligned}$$

Since F is often unknown, **the interest is in the limiting distribution of M_n as $n \to \infty$.**



2. Limit laws

The asymptotic distribution of $\boldsymbol{M_n}$ is derived in a similar approach to the univariate case.

Asymptotic distribution of M_n

Assume there exists sequences of normalising constants $\boldsymbol{a_n} = (a_{n,1}, \ldots, a_{n,d}) > \mathbf{0} = (0, \ldots, 0)$ and $\boldsymbol{b_n} = (b_{n,1}, \ldots, b_{n,d}) \in \mathbb{R}^d$ such that

$$\mathbb{P}\left(\frac{\boldsymbol{M_n} - \boldsymbol{b_n}}{\boldsymbol{a_n}} \leq \mathbf{y}\right) = F^n(\boldsymbol{a_n}\mathbf{y} + \boldsymbol{b_n}) \xrightarrow{n\to\infty} G(\mathbf{x}), \tag{1}$$

for all continuity points $\mathbf{x}$ of a non-degenerate distribution G. The class of limiting distribution G is called **multivariate extreme value distributions (MEVDs)**.

Limit Laws

MEVDS have several attractive **properties**:

• **Max-stability (Resnick, 1987, Proposition 5.9)** A distribution G on $\mathbb{R}^d$ is max-stable if for every $n \in \mathbb{N}$, there exists sequences $\boldsymbol{a_n} > \mathbf{0}$ and $\boldsymbol{b_n} \in \mathbb{R}^d$ such that

$$G(\boldsymbol{a_n}\mathbf{x} + \boldsymbol{b_n}) = G^{1/n}(\mathbf{x}).$$

• As a consequence, G is such that G^a is a distribution for every $a > 0$. A class of distributions that satisfies this property is named *max-infinitely divisible* (max-id).

• **Max-infinite divisibility**: a distribution G on $\mathbb{R}^d$ is max-id if for any $n \in \mathbb{N}$ there exists a distribution F_n such that $G = F_n^n$.

• The above implies that G can always be defined through the distribution of the sample maxima of n i.i.d. random vectors.

Marginal standardisation

Considered separately for $j = 1, \ldots, d$, the $\mathbf{X}_{i,j}$'s are sequences of independent random variables. **Univariate EVT can be applied marginally.**

By construction we have that by the **extremal type theorem**, there exists some constants $(a_{n,1}, \ldots, a_{n,d}) > \mathbf{0}$ and $(b_{n,1}, \ldots, b_{n,d}) \in \mathbb{R}^d$ such that

$$M_{n,j}^* \xrightarrow{d} Y_j \sim G_{\xi_j}, \quad j = 1, \ldots, d.$$

The margins are GEV distributed!

For simplicity, we can **assume the marginal distributions known**.

Marginal standardisation

The precise choice of marginal distribution itself is not so important. Still, a particularly useful choice is that of **unit Fréchet**.

We know that marginally for $j = 1, \ldots, d$,

$$\mathbb{P}(Y_j \leq y) = \exp\left\{ -\left[(1 + \xi_j y)_+ \right]^{-1/\xi_j} \right\},$$

then $Y_j' = (1 + \xi_j Y_j)_+^{1/\xi_j}$ **has unit Fréchet distribution,** $F_{Y_j'}(z) = \exp(-1/z), z > 0$.



Prove the above result.

Marginal standardisation

If we let $Z_{i,j} = (1 + \xi_j X_{i,j})_+^{1/\xi_j}$, then we can show that, when appropriately normalised, the limiting distribution of $\widetilde{M}_{n,j} = \max_{i=1,\ldots,n} Z_{i,j}$ is unit Fréchet.



Prove the above result using the limiting distribution of $M_{n,j} = \max_{i=1,\ldots,n} X_{i,j}$.

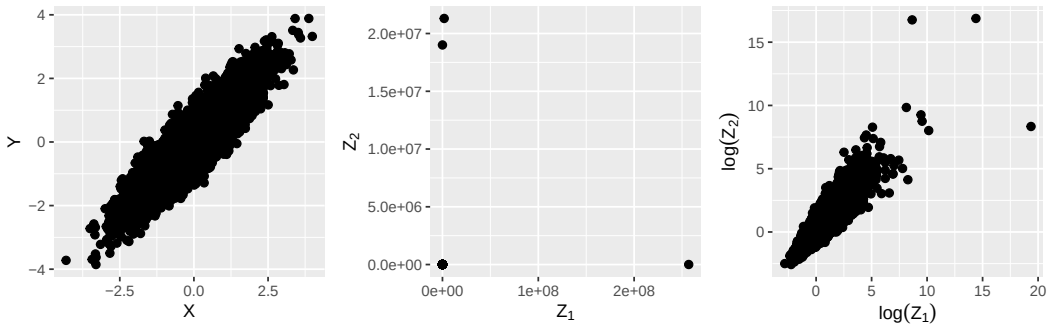
Note that since the limiting distribution is unit Fréchet, then from the **Week 3** material, a convenient choice of normalising constant for $\widetilde{M}_{n,j}$ is $b_{n,j} = 0$ and $a_{n,j} = n$.

Find an example of transformation to unit Fréchet scale in the following slides.

Marginal standardisation

Example - Normal distribution

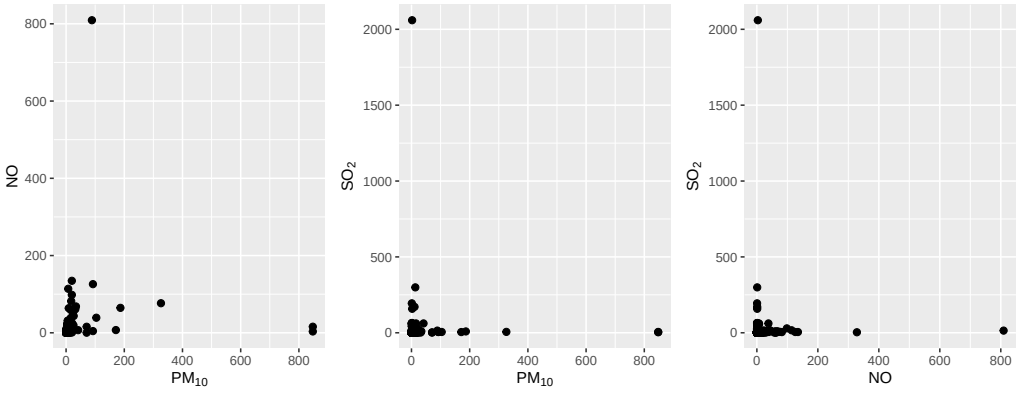
We first generate $n = 10,000$ observations from the bivariate normal distribution with correlation $\rho = 0.9$ (left) and then transform them to unit Fréchet scale (centre). For better visualisation the axes are finally put on the log scale (right).



Marginal standardisation

Example - Pollution in Leeds, UK

Standardising the Leeds pollution dataset



Theorem - Multivariate Extreme Value Distributions (MEVDs)
Let $\widetilde{M}_{n,j} = \max_{i=1,\dots,n} Z_{i,j}/n, j = 1, \dots, d$ be the linearly rescaled componentwise maxima of independent vectors $(X_1, \dots, X_d)$, transformed to have limiting unit Fréchet distributions, then for $z_1, \dots, z_d > 0$

$$\left(\widetilde{M}_{n,1} \leq z_1, \dots, \widetilde{M}_{n,d} \leq z_d\right) \xrightarrow{d} \mathbf{Y} \sim G(z_1, \dots, z_d) = \exp\left\{-\mathbf{V(z_1, \dots, z_d)}\right\},$$

where G is a non-degenerate distribution function and V is the **exponent (dependence) function**.

Comments:

- The proposed transformations can also be applied to the point processes $\{a_{n,j}^{-1}(X_{i,j} - b_{n,j}) : i = 1, \ldots, n\}$ for the X_j's.
- Note that for all $j = 1, \ldots, d$, $G(\infty, \ldots, z_j, \ldots, \infty) = \exp(-1/z_j)$ where $z_j > 0$, meaning the marginal distirbutions are indeed unit Fréchet.
- The exponent function represents the dependence structure of multiple extremes (extremal dependence). It depends on the unknown F.

Theorem - Characterisation of the exponent function

The exponent function defined in the previous Theorem can be written as

$$V(z_1, \ldots, z_d) = d \int_{\mathbb{W}} \max\left(\frac{\omega_1}{z_1}, \ldots, \frac{\omega_d}{z_d}\right) \mathrm{d}H(\boldsymbol{\omega}),$$

where $\mathbb{W} = \{\boldsymbol{\omega} \in \mathbb{R}_+^d \backslash \{\mathbf{0}\} : \omega_1 + \cdots + \omega_d = 1\}$, i.e. the d-dimensional unit simplex and $H(\cdot)$ is the **spectral measure** satisfying

$$\int_{\mathbb{W}} \omega_j \mathrm{d}H(\boldsymbol{\omega}) = \frac{1}{d}, \quad \forall j = 1, \ldots, d.$$

If H is differentiable with density h known as the *angular density*, we can write $\mathrm{d}H(\boldsymbol{\omega}) = h(\boldsymbol{\omega})\mathrm{d}\boldsymbol{\omega}$.

Details about this results will unravel throughout the coming lectures.

Example - Independence

If H is not differentiable but places mass $1/d$ on the d-dimensional vectors $\boldsymbol{e}_j = (0, \ldots, 0, 1, 0, \ldots, 0)$, then

$$V(z_1, \ldots, z_d) = \frac{1}{z_1} + \ldots + \frac{1}{z_d} = \sum_{j=1}^{d} \frac{1}{z_j},$$

and thus $$G(z_1, \ldots, z_d) = \prod_{j=1}^{d} \exp\left\{-\frac{1}{z_j}\right\} = \prod_{j=1}^{d} G(\infty, \ldots, z_j, \ldots, \infty).$$

Since the joint distribution is the product of the marginal distributions we are in presence of independent random variables.



Write down the detailed steps when $d = 3$.

Example - Perfect dependence

If H is not differentiable but places mass 1 on $\boldsymbol{\omega} = \left(\frac{1}{d}, \ldots, \frac{1}{d}\right)$, then

$$V(z_1, \ldots, z_d) = d \times \max\left(\frac{1}{dz_1}, \ldots, \frac{1}{dz_d}\right),$$

which implies that

$$G(z_1, \ldots, z_d) = \exp\left\{-\max\left(\frac{1}{z_1}, \ldots, \frac{1}{z_d}\right)\right\}$$

The above is the joint distribution of variables that are marginally standard Fréchet, but which are perfectly dependent: $X_1 = \ldots = X_d$ with probability 1.

3. Quantifying and representing extremal dependence

Recap

The **extremal dependence is expressed through the exponent function** V, which is a map from $\mathbb{R}_+^d$ to $(0, \infty)$ with the following properties:

- **Homogeneity of order -1:** $V(a\mathbf{z}) = a^{-1}V(\mathbf{z}), \forall a > 0$.
- **Convexity:** $V\left(a\mathbf{z} + (1-a)\mathbf{z}'\right) \leq aV(\mathbf{z}) + (1-a)V(\mathbf{z}'), \forall a \in [0,1], \mathbf{z}, \mathbf{z}' \in \mathbb{R}_+^d$.
- **Ranges between complete dependence and complete independence:**

$\max\left(\frac{1}{z_1}, \ldots, \frac{1}{z_d}\right) \leq V(\mathbf{z}) \leq \frac{1}{z_1} + \ldots + \frac{1}{z_d}$.

Representing the extremal dependence

The **Pickands dependence function** as visual summary of the exponent function.

Definition [Pickands, 1981] The Pickands dependence function is a function $A : \mathbb{W} \to [1/d, 1]$ defined by

$$V(\mathbf{z}) = (z_1^{-1} + \cdots + z_d^{-1}) A(t_1, \ldots, t_d), \quad \mathbf{z} \in \mathbb{R}_+^d.$$

where $t_j = z_j^{-1} / \sum_{i=1}^d z_i^{-1}, j = 1, \ldots, d-1$ and $t_d = 1 - \sum_{i=1}^{d-1} t_i$.

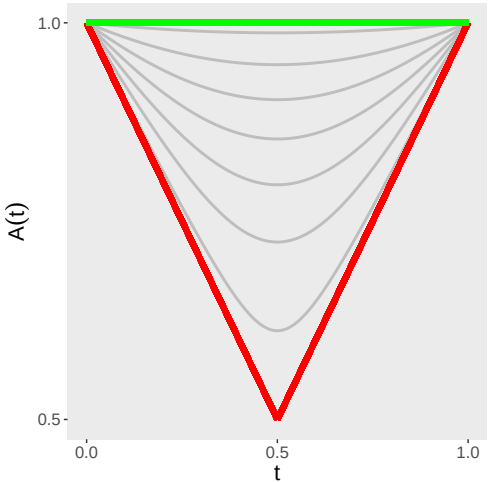
Following on the properties of the exponent function we have that

$$\frac{1}{d} \leq \max(t_1, \ldots, t_d) \leq A(\mathbf{t}) \leq 1,$$

where the lower bound correspond to complete dependence whereas the upper bound corresponds to independence. Furthermore, it is a **convex** function.

Representing the extremal dependence

Example - Bivariate case, $d = 2$



Here $A : \mathbb{W} \to [1/2, 1]$ such that

$$A(t) = \frac{z_1 z_2}{z_1 + z_2} V(z_1, z_2),$$

with $t = z_2/(z_1 + z_2)$.

We also have $\max(t, 1 - t) \leq A(t) \leq 1$.

Example - Bivariate case, $d = 2$

As seen in the previous slide:

$$V(z_1, z_2) = (z_1^{-1} + z_2^{-1}) A\left(\frac{z_2}{z_1 + z_2}\right), \quad z_1, z_2 > 0.$$

Note that we can also write:

$$G(z_1, z_2) = \exp\left\{\log\left(G_1(z_1) G_2(z_2)\right) A\left(\frac{\log\left(G_1(z_1) G_2(z_2)\right)}{\log\left(G_2(z_2)\right)}\right)\right\},$$

where G_1 and G_2 are the marginal distributions (unit-Fréchet).

Quantifying the extremal dependence

Clearly graphical summaries of extremal dependence are ok up until $d = 3$.

We introduce numerical summaries to get a better understanding in higher dimensions.

Definition [Smith, 1990]

The extremal coefficient is given by $\vartheta = V(1, \ldots, 1) = d \times A\left(\frac{1}{d}, \ldots, \frac{1}{d}\right)$.

- ϑ represents the **number of independent components** of the random vector $\mathbf{Y}$.
- $\vartheta \in [1, d]$ with bounds respectively indicating complete dependence and independence.



Can you prove the last comment using a result previously stated?

Quantifying the extremal dependence

Example - Independence when $d = 2$

Let's recall that in case of complete independence $V(z_1, z_2) = 1/z_1 + 1/z_2$, therefore $\vartheta = V(1, \ldots, 1) = 2$.

Alternatively $A(1/2) = 1$ and thus $\vartheta = 2 \times A(1/2) = 2$.

Example - Complete dependence when $d = 2$

Let's recall that in case of complete dependence $V(z_1, z_2) = \max(1/z_1, 1/z_2)$, therefore $\vartheta = V(1, \ldots, 1) = 1$.

Alternatively $A(1/2) = 1/2$ and thus $\vartheta = 2 \times A(1/2) = 1$.



Reproduce the above examples when $d = 3$.

Quantifying the extremal dependence

Definition [dHF06; Nikoloulopoulos et al., 2009]

For any $\mathbf{z} > \mathbf{0}$ and failure region

$$\mathcal{F}_{\mathbf{z}} = \left(\mathbf{v} \in \mathbb{R}_+^d \backslash \{\mathbf{0}\} : v_1 > z_1 \text{ and } \ldots \text{ and } v_d > z_d\right),$$

the *tail dependence function* $R(\boldsymbol{z})$ focuses on the observations that fall in the failure region (all the coordinates are simultaneously large).

It is related to the exponent function by the inclusion-exclusion principle.

Quantifying the extremal dependence

Example - Independence when $d = 2$

When $d = 2$, the tail dependence function is given by

$$R(z_1, z_2) = \frac{1}{z_1} + \frac{1}{z_2} - V(z_1, z_2).$$



Write down the tail dependence function as function of the exponent function when $d = 3$.

Quantifying the extremal dependence

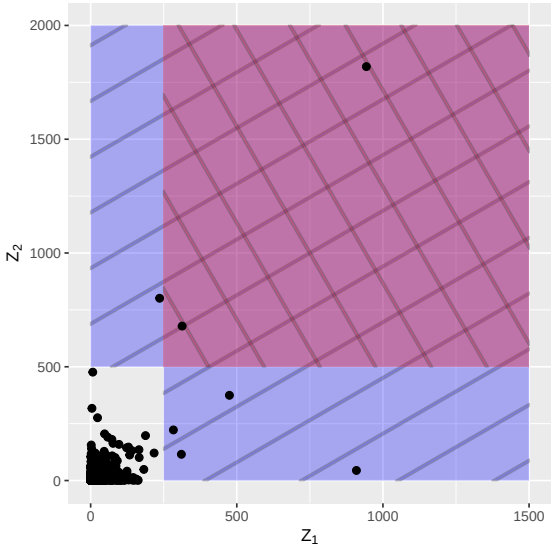


Illustration of the tail dependence function in a 2-dimensional context.

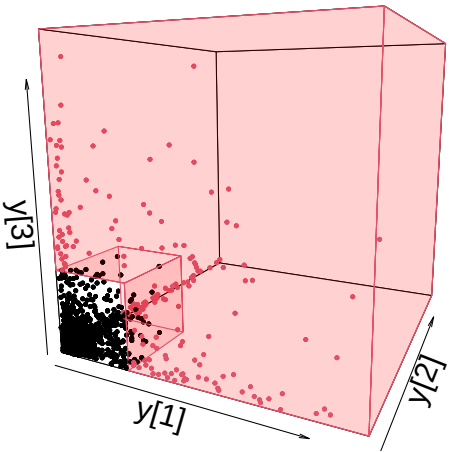
Exponent function relates to the observations that have at least one large component: $Z_1 > 250$ or $Z_2 > 500$ (**blue striped region**).

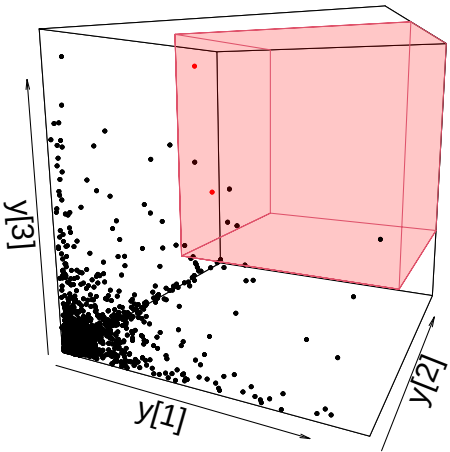
Tail dependence function relates to the observations that have all components large: $Z_1 > 250$ and $Z_2 > 500$ (red striped region).

Quantifying the extremal dependence

Illustration of the tail dependence function in a 3-dimensional context.

Weak dependence

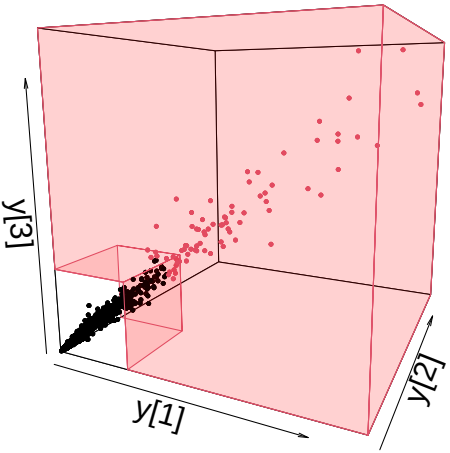


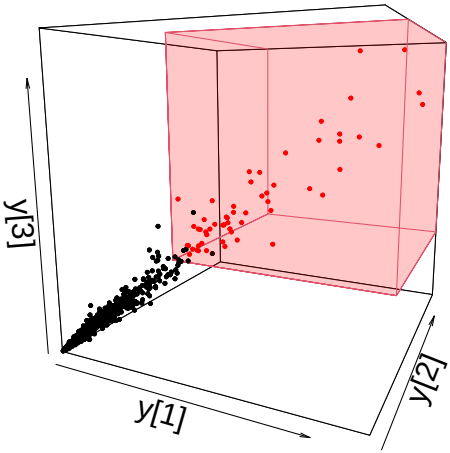


Quantifying the extremal dependence

Illustration of the tail dependence function in a 3-dimensional context.

Strong dependence





Quantifying the extremal dependence

Another summary of the dependence between the elements of $\mathbf{Y}$ is defined as follows.

Definition [Joe, 1997, Chap. 2; Li, 2009]
The coefficient of upper tail dependence is given by

$$\chi = R(1, \ldots, 1).$$

It measures the strength of dependence in the tail distribution of $\mathbf{Y}$ or in other terms the probability that all the components of $\mathbf{Y}$ are simultaneously large.

Quantifying the extremal dependence

Example - $d = 2$

We have established that $R(z_1, z_2) = \frac{1}{z_1} + \frac{1}{z_2} - V(z_1, z_2)$, therefore

$$\chi = R(1,1) = 2 - V(1,1) = 2 - \vartheta.$$

- In the case of **independence**, we have shown that $\vartheta = 2$ and thus $\chi = 0$.
- In the case of **complete dependence**, we have shown that $\vartheta = 1$ and thus $\chi = 1$.



Reproduce the above example when $d = 3$.

Quantifying the extremal dependence

The **exponent** and **tail dependence** functions can be used to approximate the probability of certain types of extreme events occurring.

Let $\mathbf{Z}$ be a random vector with unit-Pareto margins. The joint d.f. F is in the domain of attraction of a MEVD with Fréchet margins.

1. Exponent function:

We have the approximation $1 - F(n\mathbf{z}) \approx V(n\mathbf{z})$.

We can therefore use the approximation

$$\mathbb{P}\left(Z_1 > z_1 \text{ or } \ldots \text{ or } Z_d > z_d\right) \approx V(z_1, \ldots, z_d),$$

where $z_1, \ldots, z_d$ are high enough thresholds.

Quantifying the extremal dependence

2. Tail dependence function:

We have the approximation $\mathbb{P}(Z_1 > nz_1, \ldots, Z_d > nz_d) \approx R(n\mathbf{z})$.

We can therefore use the approximation

$$\mathbb{P}\left(Z_1 > z_1 \text{ and } \ldots \text{ and } Z_d > z_d\right) \approx R(z_1, \ldots, z_d),$$

where $z_1, \ldots, z_d$ are high enough thresholds.



Justify in details the blue approximations stated in the last two slides.

Returns levels

Similarly to the univariate case we are interested in estimating some return levels for some given return periods (or vice versa). Different definitions are possible but here we use the following.

Definition

The tail dependence function can be used for estimating **joint return levels**. Let $J \subset \{1, \dots, d\}$, $\{x_i, i \in \{1, \dots, d\} \backslash J\}$ be a sequence of fixed high thresholds and $p \in (0, 1)$ be a fixed probability.

Given a **return period** $1/p$, we define **joint return levels** the quantiles $\{z_{j;p}, j \in J\}$ satisfying the equation

$$p = \mathbb{P}\left(Z_j > z_{j;p}, Z_i > x_i, j \in J, i \in \{1, \dots, d\} \backslash J\right)$$

4. Parametric models for extremal dependence



Asymmetric logistic (Tawn, 1990)

Construction:

Let $\mathbb{S}$ be the power set of $\{1, \ldots, d\}$ minus the empty set, and $\mathcal{S}$ the index set that takes values in $\mathbb{S}$.

1. Let $N_{\mathcal{S}}$ be a Poisson random variable with rate $1/\tau_{\mathcal{S}}$. One can think of it as the number of storm events that take place on the sites $\mathcal{S}$ in a time interval.

2. Given $n_{\mathcal{S}}$, $j \in \mathcal{S}$, let $\{X_{j,\mathcal{S};i}, i = 1, \ldots, n_{\mathcal{S}}\}$ be a dependent sequence of i.i.d. random variables that describe an episode (e.g., rain) and define $X_{j,\mathcal{S}} = \max_{i=1,\ldots,n_{\mathcal{S}}}\{X_{j,\mathcal{S};i}\}$.

3. Let $\alpha_{\mathcal{S}}$ be a random effect with a positive stable distribution and stability parameter $\alpha_{\mathcal{S}} \geq 1$ representing an unrecorded additional piece of information on storm events.

4. Consider the independent sequence $\{X_{j,\mathcal{S}}\}_{j\in\mathcal{S}}|\alpha_{\mathcal{S}}$ and define $Y_j = \max_{\mathcal{S}\in\mathbb{S}_j}\{X_{j,\mathcal{S}}\}$ where $\mathbb{S}_j \subset \mathbb{S}$ contains all the sets including $j \Rightarrow$ max over all the storms involving j.

Asymmetric logistic

After marginal transformation to **unit exponential** random variables, the exponent function is given by

$$V(\mathbf{z};\boldsymbol{\theta}) = \sum_{\mathcal{S}\in\mathbb{S}} \left\{ \sum_{j\in\mathcal{S}} \left(\frac{\beta_{j,\mathcal{S}}}{z_j} \right)^{\alpha_{\mathcal{S}}} \right\}^{1/\alpha_{\mathcal{S}}}, \quad \mathbf{z} \in \mathbb{R}_+^d,$$

where $\boldsymbol{\theta} = \{\alpha_{\mathcal{S}}, \boldsymbol{\beta_{j,\mathcal{S}}}\}_{\mathcal{S}\in\mathbb{S}}$, $\alpha_{\mathcal{S}} \geq 1$, $\beta_{\mathcal{S}} = \tau_{\mathcal{S}} / \sum_{\mathcal{S}\in\mathbb{S}_j} \tau_{\mathcal{S}}$ and $\beta_{j,\mathcal{S}} = 0$ if $j \notin \mathcal{S}$ and for $j \in \{1,\ldots,d\}$, $0 \leq \beta_{j,\mathcal{S}} \leq 1$ and $\sum_{\mathcal{S}\in\mathbb{S}} \beta_{j,\mathcal{S}} = 1$.

$\beta_{j,\mathcal{S}}$ represents the probability that the maximum value value observed at site j is attributed to a storm event involving the sites of $\mathcal{S}$.

The number of model parameters is $2^{d-1}(d+2) - (2d+1)$.

Asymmetric logistic

Example - $d = 2$

The model has $2(2+2)-(2\times 2+1) = 3$ parameters. Here $\mathbb{S} = \{\{1\},\{2\},\{1,2\}\}$

and

- $\beta_{1,\{1\}} + \beta_{1,\{2\}} + \beta_{1,\{1,2\}} = \beta_{1,\{1\}} + \beta_{1,\{1,2\}} = 1$
  $\Longrightarrow \beta_{\mathbf{1},\{1\}} = \mathbf{1} - \beta_{\mathbf{1}}$
- $\beta_{2,\{1\}} + \beta_{2,\{2\}} + \beta_{2,\{1,2\}} = \beta_{2,\{2\}} + \beta_{2,\{1,2\}} = 1$
  $\Longrightarrow \beta_{\mathbf{2},\{2\}} = \mathbf{1} - \beta_{\mathbf{2}}$

Vector of parameters is $\boldsymbol{\theta} = (\alpha, \beta_1, \beta_2)$.

Asymmetric logistic

Example - $d = 2$

The exponent function is given by

$$V(z_1, z_2; \boldsymbol{\theta}) = \frac{1-\beta_1}{z_1} + \frac{1-\beta_2}{z_2} + \left\{ \left(\frac{\beta_1}{z_1} \right)^{\alpha} + \left(\frac{\beta_2}{z_2} \right)^{\alpha} \right\}^{1/\alpha},$$

and the Pickands dependence function

$$A(t) = (1-\beta_1)t + (1-\beta_2)(1-t) + \{(\beta_1 t)^{\alpha} + (\beta_2(1-t))^{\alpha}\}^{1/\alpha}, \quad 0 \leq t \leq 1.$$



Plot the Pickands dependence function of the asymmetric logistic model for various values of θ. What condition is needed to have a symmetric dependence structure?

Hüsler-Reiss (Hüsler & Reiss, 1989)

Construction:

1. $\mathbf{X}_1, \ldots, \mathbf{X}_n$ i.i.d. replicates of a Gaussian random vector with zero mean, unit variance and pairwise correlation $\rho_{i,j;n}, i, j \in \{1, \ldots, d\}$ satisfying

$$\lim_{n\to\infty} \log n(1 - \rho_{i,j;n}) = \lambda_{i,j}^2 \in [0, \infty).$$

2. For $\boldsymbol{b}_n = (b_n, \ldots, b_n)$, the normalised vector $\boldsymbol{b}_n(\boldsymbol{M}_n - \boldsymbol{b}_n)$ has limiting distribution with exponent function

$$V(\mathbf{z}; \boldsymbol{\theta}) = \sum_{j=1}^{d} \frac{1}{z_j} \Phi_{d-1} \left\{ \left(\lambda_{i,j} + \frac{\log z_i / z_j}{2\lambda_{i,j}} \right)_{i \in \{1,\ldots,d\} \setminus \{j\}} ; \bar{\Lambda}_j \right\},$$

where Φ_{d-1} is the cdf of $d-1$ dimensional Gaussian distribution with partial correlation $\bar{\Lambda}_j$.

Comments

- For all $j = 1, \ldots, d$, the elements of $\bar{\Lambda}_j$ are

$$\lambda_{k,i;j} = \frac{\lambda_{k,j}^2 + \lambda_{i,j}^2 - \lambda_{k,i}^2}{2\lambda_{k,j}\lambda_{i,j}}, \quad k, i \in \{1, \ldots, d\} \backslash \{j\}; k < i.$$

Note the these define in row order the upper triangle (or in column order the lower triangle) of the matrix and symmetry needs to be used to complete it.

- The parameter vector is $\boldsymbol{\theta} = \left(\{\lambda_{i,j}\}_{i,j \in \{1,\ldots,d\}}\right)$
- $\lambda_{i,j}$: controls the dependence between the i-th and j-th elements of a vector f d extremes.
  - $\star$ $\lambda_{i,j} = 0 \Longrightarrow$ Complete dependence.
  - $\star$ $\lambda_{i,j} \to \infty \Longrightarrow$ Independence.

Hüsler-Reiss

Example - $d = 2$

There is only 1 parameter given by $\theta = \lambda_{1,2} = \lambda > 0$.

The exponent function is given by

$$V(z_1, z_2) = \frac{1}{z_1}\Phi\left(\lambda + \frac{\log(z_2/z_1)}{2\lambda}\right) + \frac{1}{z_2}\Phi\left(\lambda + \frac{\log(z_1/z_2)}{2\lambda}\right),$$

and the Pickands dependence function

$$A(t) = t\Phi\left(\lambda + \frac{\log(t/(1-t))}{2\lambda}\right) + (1-t)\Phi\left(\lambda + \frac{\log((1-t)/t)}{2\lambda}\right), \quad 0 \leq t \leq 1.$$



Plot the Pickands dependence function of the Hüsler-Reiss model for various values of θ.

Extremal-t (Nikoloulopoulos et al., 2009)

Construction:

1. $\mathbf{X}_1, \ldots, \mathbf{X}_n$ i.i.d. replicates of a Student-t random vector with zero mean, dispersion matrix Σ and ν degrees of freedom (DoF).

2. For $\boldsymbol{a}_n = (a_n, \ldots, a_n)$, the normalised vector $\boldsymbol{M}_n / \boldsymbol{a}_n$ has limiting distribution with exponent function

$$V(\mathbf{z};\boldsymbol{\theta}) = \sum_{j=1}^{d} \frac{1}{z_j} T_{d-1;\nu+1} \left\{ \left(\sqrt{\frac{\nu+1}{1-\rho_{i,j}^2}} \left(\left(\frac{z_i}{z_j} \right)^{1/\nu} - \rho_{i,j} \right) \right)_{i \in \{1,\ldots,d\} \setminus \{j\}} ; \bar{\Sigma}_j \right\},$$

where $T_{d-1;\nu+1}$ is the cdf of $d-1$ dimensional Student-t with $\nu+1$ DoF and partial correlation $\bar{\Sigma}_j$.

Extremal-t

Comments

- For all $j = 1, \ldots, d$, the elements of $\bar{\Sigma}_j$ are

$$\sigma_{k,i;j} = \frac{\rho_{k,i} - \rho_{k,j}\rho_{i,j}}{\sqrt{(1-\rho_{k,j}^2)(1-\rho_{i,j}^2)}}$$

- The parameter vector is $\boldsymbol{\theta} = \left(\{\rho_{i,j}\}_{i,j\in\{1,\ldots,d\}}, \nu\right)$
- $\rho_{i,j}$: controls the dependence between the i-th and j-th elements of a vector f d extremes. Dependence increases with increasing $\rho_{i,j}$.
- ν: controls the overall dependence among all the variables. Dependence increases with decreasing value of ν.
- The Hüsler-Reiss is a special case of the Extremal-t: set $\rho_{i,j} = 1 - \lambda_{i,j}^2/\nu$ and then the Extremal-t converges weakly to the Hüsler-Reiss as $\nu \to \infty$.

Example - $d = 2$

There are 2 parameters: $\boldsymbol{\theta} = (\rho_{1,2}, \nu) = (\rho, \nu)$ where $\rho \in [-1, 1]$ and $\nu > 0$.
The exponent function is given by

$$V(z_1, z_2) = \frac{1}{z_1} T_{\nu+1}\left\{\sqrt{\frac{\nu+1}{1-\rho^2}}\left(\left(\frac{z_2}{z_1}\right)^{1/\nu} - \rho\right)\right\} + \frac{1}{z_2} T_{\nu+1}\left\{\sqrt{\frac{\nu+1}{1-\rho^2}}\left(\left(\frac{z_1}{z_2}\right)^{1/\nu} - \rho\right)\right\},$$

and the Pickands dependence function

$$A(t) = t T_{\nu+1}\left\{\sqrt{\frac{\nu+1}{1-\rho^2}}\left(\left(\frac{t}{1-t}\right)^{1/\nu} - \rho\right)\right\} + (1-t) T_{\nu+1}\left\{\sqrt{\frac{\nu+1}{1-\rho^2}}\left(\left(\frac{1-t}{t}\right)^{1/\nu} - \rho\right)\right\}.$$



Plot the Pickands dependence function of the Extremal-t model for various values of θ.

References



Sources and recommending reading

1. **[BGST04]** Beirlant, J., Goegebeur, Y., Segers, J. & Teugels J. (2004) *Statistics of Extremes: Theory and Applications*. John Wiley & Sons. Chapter 8 & 9

2. **[C01]** Coles, S. (2001) An Introduction to Statistical Modeling of Extreme Values. Springer. Chapter 7.