Multivariate Sparse Clustering for Extremes

Olivier WINTENBERGER

LPSM, Sorbonne université

Joint work with Nicolas MEYER, LMG, Montpellier

March 24 2023

StatMath Research Seminar @WU, Vienna

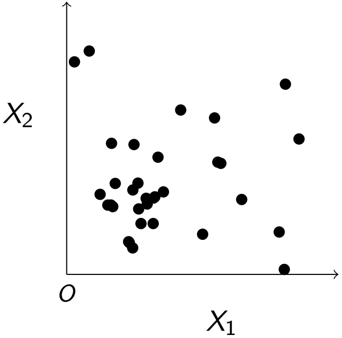
LPSM
LABORATOIRE DE PROBABILITÉS
STATISTIQUE & MODÉLISATION

Some introductory words

General goal: Developing a procedure to learn the dependence structure of multivariate extremes

- ▷ How to model the dependence structure of multivariate extremes?
- ▷ In practice how to reduce the dimension (curse of dimensionality)?
- ▷ How many data points should be considered as extreme?

Multivariate extremes

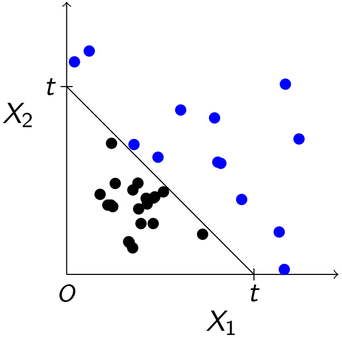


- Threshold exceedances $\mathbf{X} \mid |\mathbf{X}| > t$:

$$\mathbf{X}/t \mid |\mathbf{X}| > t \underset{t\to\infty}{\to} \text{GPD}$$

GPD: Generalized Pareto distribution

Multivariate extremes



- Threshold exceedances $\mathbf{X} \mid |\mathbf{X}| > t$:

$$\mathbf{X}/t \mid |\mathbf{X}| > t \underset{t\to\infty}{\to} \text{GPD}$$

GPD: Generalized Pareto distribution

Regular variation

- We assume regular variation:

$$\mathbb{P}(\mathbf{X}/t \in \cdot \mid |\mathbf{X}| > t) \xrightarrow{w} \mathbb{P}(\mathbf{Y} \in \cdot)\,, \quad t \to \infty\,.$$

- Study of the angular component:

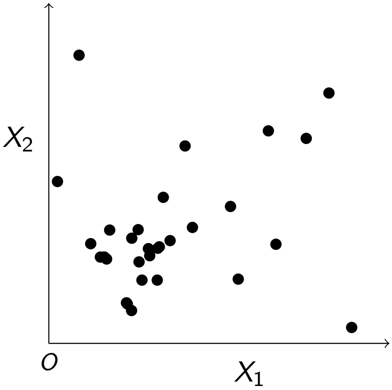
$$\mathbb{P}(\mathbf{X}/|\mathbf{X}| \in \cdot \mid |\mathbf{X}| > t) \xrightarrow{w} \mathbb{P}(\mathbf{\Theta} \in \cdot)\,, \quad t \to \infty\,.$$

The vector $\mathbf{\Theta} = \mathbf{Y}/|\mathbf{Y}|$ in

$$\mathbb{S}_+^{d-1} = \{\mathbf{x} \in \mathbb{R}_+^d,\ |\mathbf{x}| = 1\}\,.$$

is the spectral vector and its distribution is the spectral measure.

The spectral measure

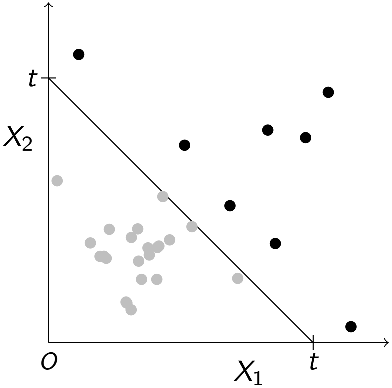


- Convergence to the spectral measure:

$$\mathbb{P}(\mathbf{X}/|\mathbf{X}| \in \cdot \mid |\mathbf{X}| > t) \xrightarrow{w} \mathbb{P}(\mathbf{\Theta} \in \cdot)\,,$$

when $t \to \infty$.

The spectral measure

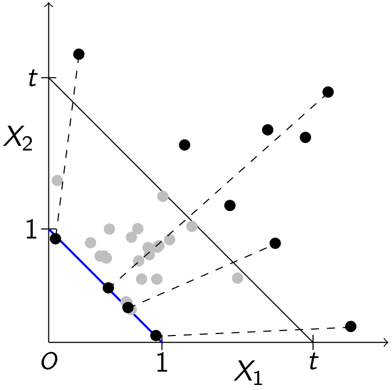


- Convergence to the spectral measure:

$$\mathbb{P}(\mathbf{X}/|\mathbf{X}| \in \cdot \mid |\mathbf{X}| > t) \xrightarrow{w} \mathbb{P}(\mathbf{\Theta} \in \cdot)\,,$$

when $t \to \infty$.

The spectral measure

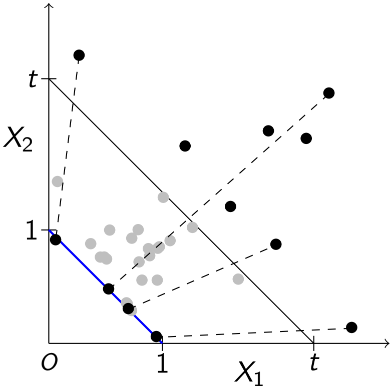


- Convergence to the spectral measure:

$$\mathbb{P}(\mathbf{X}/|\mathbf{X}| \in \cdot \mid |\mathbf{X}| > t) \xrightarrow{w} \mathbb{P}(\mathbf{\Theta} \in \cdot)\,,$$

when $t \to \infty$.

The spectral measure



- ▷ Convergence to the spectral measure:

$$\mathbb{P}(\mathbf{X}/|\mathbf{X}| \in \cdot \mid |\mathbf{X}| > t) \xrightarrow{w} \mathbb{P}(\mathbf{\Theta} \in \cdot)\,,$$

when $t \to \infty$.

$\hookrightarrow$ How to estimate the support of the spectral measure $S(\cdot) := \mathbb{P}(\mathbf{\Theta} \in \cdot)$?

A natural partition for $\mathbb{S}_+^{d-1}$

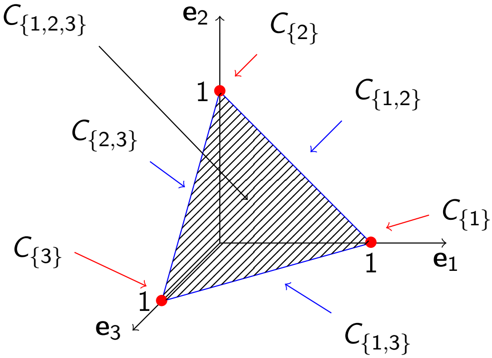
For $\beta \subset \{1, \dots, d\}$ we define

$$C_\beta = \{\mathbf{x} \in \mathbb{S}_+^{d-1}, \text{ for all } j \in \beta,\ x_j > 0, \text{ for all } j \notin \beta,\ x_j = 0\}\,.$$

A natural partition for $\mathbb{S}_+^{d-1}$

For $\beta \subset \{1, \dots, d\}$ we define

$$C_\beta = \{\mathbf{x} \in \mathbb{S}_+^{d-1}, \text{ for all } j \in \beta,\ x_j > 0, \text{ for all } j \notin \beta,\ x_j = 0\}\,.$$



And regarding extremes?

Interpretation of the C_β regarding the spectral measure[1]:

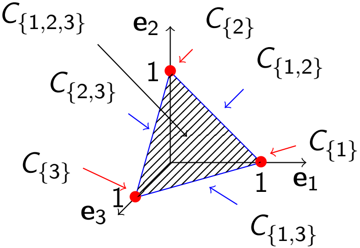
$$\mathbb{P}(\boldsymbol{\Theta} \in C_\beta) > 0 \iff \text{it is likely to observe extremes in the cluster } \beta\,.$$

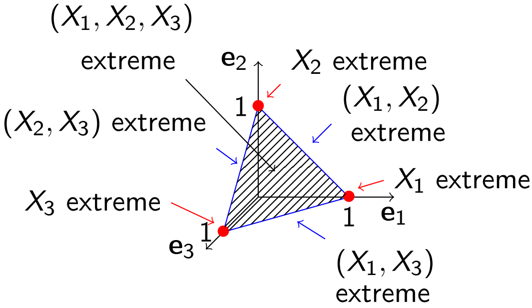
[1] see also *Simpson et al. (2019)*

And regarding extremes?

Interpretation of the C_β regarding the spectral measure[1]:

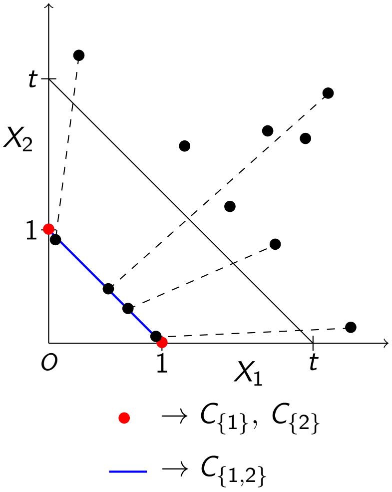
$$\mathbb{P}(\boldsymbol{\Theta} \in C_\beta) > 0 \iff \text{it is likely to observe extremes in the cluster } \beta\,.$$



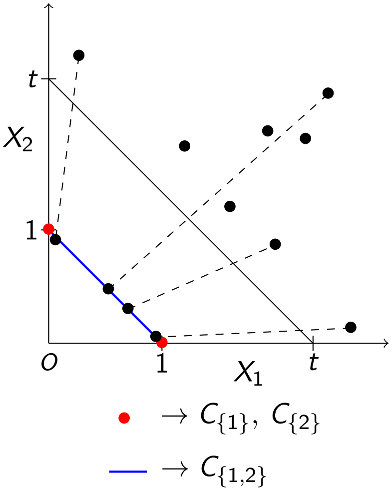


[1] see also *Simpson et al. (2019)*

Some issues with the standard framework

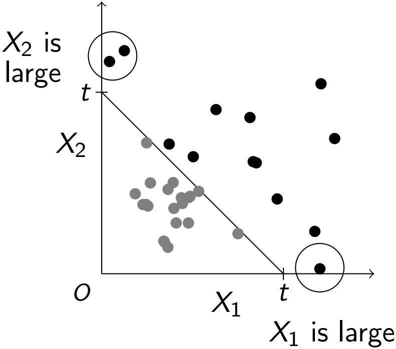


Some issues with the standard framework



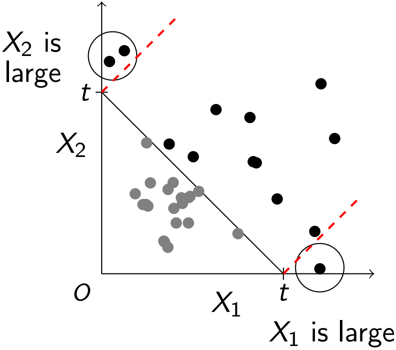
- Statistical issue:
  $\mathbb{P}(\mathbf{X}/|\mathbf{X}| \in C_{\{1,\ldots,d\}}) = 1$
- Topological issue: $\mathbf{\Theta}$ may put mass on the boundary of the C_β (no weak convergence)

How to use the data?



[1] *Goix et al. (2016), Goix et al. (2017), Chiapino and Sabourin (2016), Chiapino et al. (2019)*

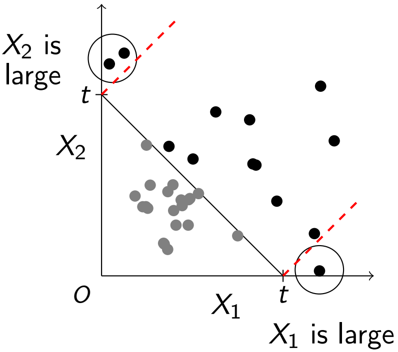
How to use the data?



- ▷ Introduce a classification procedure[1]

[1] *Goix et al. (2016), Goix et al. (2017), Chiapino and Sabourin (2016), Chiapino et al. (2019)*

How to use the data?



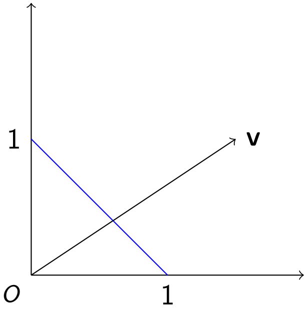
- ▷ Introduce a classification procedure[1]
- ▷ This is done by the Euclidean projection onto the simplex!

[1] *Goix et al. (2016), Goix et al. (2017), Chiapino and Sabourin (2016), Chiapino et al. (2019)*

The Euclidean projection onto the simplex

The Euclidean projection $\pi : \mathbb{R}_+^d \to \mathbb{S}_+^{d-1}$ onto the simplex[1]:

$$\pi(\mathbf{v}) = \underset{\mathbf{w}\geq\mathbf{0},\,|\mathbf{w}|=1}{\arg\min}\ |\mathbf{v}-\mathbf{w}|_2.$$

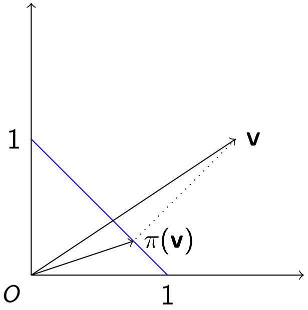


[1] *Michelot (1986), Duchi et al. (2008),...*

The Euclidean projection onto the simplex

The Euclidean projection $\pi : \mathbb{R}_+^d \to \mathbb{S}_+^{d-1}$ onto the simplex[1]:

$$\pi(\mathbf{v}) = \underset{\mathbf{w}\geq\mathbf{0},\,|\mathbf{w}|=1}{\arg\min}\ |\mathbf{v}-\mathbf{w}|_2.$$

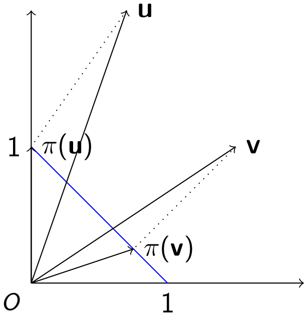


[1] *Michelot (1986), Duchi et al. (2008),...*

The Euclidean projection onto the simplex

The Euclidean projection $\pi : \mathbb{R}_+^d \to \mathbb{S}_+^{d-1}$ onto the simplex[1]:

$$\pi(\mathbf{v}) = \underset{\mathbf{w} \geq \mathbf{0},\, |\mathbf{w}|=1}{\arg\min} |\mathbf{v} - \mathbf{w}|_2 .$$

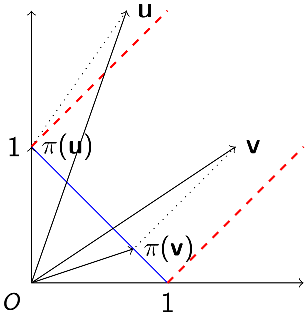


[1] *Michelot (1986), Duchi et al. (2008),...*

The Euclidean projection onto the simplex

The Euclidean projection $\pi : \mathbb{R}_+^d \to \mathbb{S}_+^{d-1}$ onto the simplex[1]:

$$\pi(\mathbf{v}) = \underset{\mathbf{w}\geq \mathbf{0},\, |\mathbf{w}|=1}{\arg\min} |\mathbf{v}-\mathbf{w}|_2 .$$

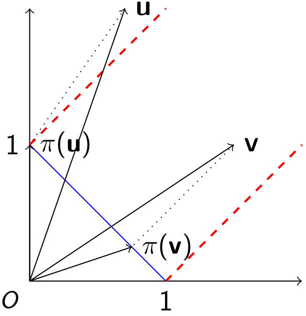


[1] *Michelot (1986), Duchi et al. (2008),...*

The Euclidean projection onto the simplex

The Euclidean projection $\pi : \mathbb{R}_+^d \to \mathbb{S}_+^{d-1}$ onto the simplex[1]:

$$\pi(\mathbf{v}) = \underset{\mathbf{w} \geq \mathbf{0},\, |\mathbf{w}|=1}{\arg\min} |\mathbf{v} - \mathbf{w}|_2.$$



Regular variation:
$$\mathbb{P}(\mathbf{X}/t \in \cdot \mid |\mathbf{X}| > t) \to \mathbb{P}(\mathbf{Y} \in \cdot)$$

So far: $\mathbb{P}(\mathbf{X}/|\mathbf{X}| \in \cdot \mid |\mathbf{X}| > t) \to \mathbb{P}(\boldsymbol{\Theta} \in \cdot)$

Now: $\mathbb{P}(\pi(\mathbf{X}/t) \in \cdot \mid |\mathbf{X}| > t) \to \mathbb{P}(\mathbf{Z} \in \cdot)$, with $\mathbf{Z} = \pi(\mathbf{Y})$

$\hookrightarrow$ **Sparse regular variation**

[1] *Michelot (1986), Duchi et al. (2008),...*

Sparse regular variation

Theorem (M. & Wintenberger (2021+))

1. *$\mathbf{X}$ is regularly varying $\Longrightarrow$ $\mathbf{X}$ is sparsely regularly varying*
2. $\left.\begin{array}{r}\mathbf{X} \textit{ is sparsely regularly varying}\\ \textit{Some assumptions on the limit } \mathbf{Z}\end{array}\right\} \Longrightarrow \mathbf{X}$ *is regularly varying*

Proposition (M. & Wintenberger (2021+))

For $\beta \subset \{1, \dots, d\}$ we have the convergence

$$\mathbb{P}(\pi(\mathbf{X}/t) \in C_\beta \mid |\mathbf{X}| > t) \to \mathbb{P}(\mathbf{Z} \in C_\beta), \quad t \to \infty\,.$$

+ Other results that link both vectors $\mathbf{Z}$ and $\boldsymbol{\Theta}$

Recap: Studying dependence in extremes

- ▷ *Goal:* Identify clusters β in which extremes appear
- ▷ *Standard model:* Regular variation $\mathbb{P}(\mathbf{X}/t \in \cdot \mid |\mathbf{X}| > t) \xrightarrow{w} \mathbb{P}(\mathbf{Y} \in \cdot)$
- ▷ *Direct attempt:* Study the angular component $\boldsymbol{\Theta} = \mathbf{Y}/|\mathbf{Y}|$ via $\mathbf{X}/|\mathbf{X}|$
  $\hookrightarrow$ some issues
- ▷ *Idea:* Study the angular component $\mathbf{Z} = \pi(\mathbf{Y})$ via $\pi(\mathbf{X}/t)$

$\rightsquigarrow$ How to estimate the support of $\mathbf{Z}$ in a statistical framework?

Statistical framework

- $\mathbf{X}_1, \ldots, \mathbf{X}_n$ i.i.d. regularly varying, $\mathbf{Z} = \pi(\mathbf{Y})$.
- A threshold u_n or equivalently a level $k_n = n\mathbb{P}(|\mathbf{X}| > u_n)$.
- *Goal: find the clusters* β *such that* $p^*(\beta) := \mathbb{P}(\mathbf{Z} \in C_\beta) > 0$.

For each $\beta \subset \{1, \ldots, d\}$ we compute

$$T_n(\beta) = \sum_{j=1}^{n} \mathbb{1}\{\pi(\mathbf{X}_j/u_n) \in C_\beta,\ |\mathbf{X}_j| > u_n\}$$

$=$ number of points in C_β among the extremes.

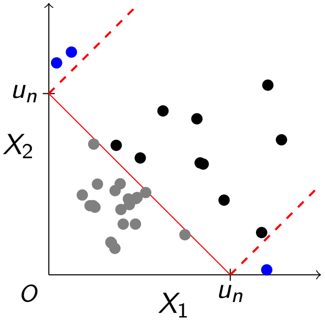
Statistical framework

- $\mathbf{X}_1, \ldots, \mathbf{X}_n$ i.i.d. regularly varying, $\mathbf{Z} = \pi(\mathbf{Y})$.
- A threshold u_n or equivalently a level $k_n = n\mathbb{P}(|\mathbf{X}| > u_n)$.
- *Goal: find the clusters β such that $p^*(\beta) := \mathbb{P}(\mathbf{Z} \in C_\beta) > 0$.*

For each $\beta \subset \{1, \ldots, d\}$ we compute

$$T_n(\beta) = \sum_{j=1}^{n} \mathbb{1}\{\pi(\mathbf{X}_j/u_n) \in C_\beta, |\mathbf{X}_j| > u_n\}$$

$= $ number of points in C_β among the extremes.



The vector $\mathbf{T}_n$

- Consider $\mathbf{T}_n = (T_n(\beta))_\beta$ and $\mathbf{p}^* = (p^*(\beta))_\beta = (\mathbb{P}(\mathbf{Z} \in C_\beta))_\beta$ (in increasing order)

- If $k_n = k$ is fixed and $u_n = |\mathbf{X}|_{(k)}$, then $\mathbf{T}_n$ follows a multinomial distribution:
  - $\mathbf{T}_n$ is a vector in $\mathbb{N}^{2^d-1}$,
  - a linear relation: $T_n(\beta_1) + \ldots + T_n(\beta_{2^d-1}) = k$

The vector $\mathbf{T}_n$

- ▷ Consider $\mathbf{T}_n = (T_n(\beta))_\beta$ and $\mathbf{p}^* = (p^*(\beta))_\beta = (\mathbb{P}(\mathbf{Z} \in C_\beta))_\beta$ (in increasing order)

- ▷ If $k_n = k$ is fixed and $u_n = |\mathbf{X}|_{(k)}$, then $\mathbf{T}_n$ follows a multinomial distribution:
  - $\mathbf{T}_n$ is a vector in $\mathbb{N}^{2^d-1}$,
  - a linear relation: $T_n(\beta_1) + \ldots + T_n(\beta_{2^d-1}) = k$

- ▷ Three types of components for $\mathbf{T}_n$:
  - $T_n(\beta) \gg 0$: relevant clusters
  - $T_n(\beta) \approx 0$: biased clusters
  - $T_n(\beta) = 0$: non-relevant clusters

Convergence results

Define $\mathcal{S}^*(\mathbf{Z}) := \{\beta : \mathbb{E}[T_n(\beta)] \to \infty,\ n \to \infty\}$ with cardinality s^* and $\mathbf{T}_{n,\mathcal{S}^*(\mathbf{Z})} := (T_n(\beta))_{\beta\in\mathcal{S}^*(\mathbf{Z})}$

Theorem

1. *Convergence:* $\mathbf{T}_{n,\mathcal{S}^*(\mathbf{Z})}/k_n \to (\mathbf{p}^*, 0, \dots, 0)$ *in probability when* $n \to \infty$.
2. *Asymptotic normality under hidden regular variation and bias assumptions:*

$$\mathrm{Diag}(\mathbb{E}[\mathbf{T}_{n,\mathcal{S}^*(\mathbf{Z})}])^{-1/2}\Big(\mathbf{T}_{n,\mathcal{S}^*(\mathbf{Z})} - \mathbb{E}[T_n(\beta)]\Big) \xrightarrow{d} \mathcal{N}(0, Id_{s^*}), \quad n \to \infty.$$

Model selection for k_n fixed

- For $k_n = k$ fixed, $\mathbf{T}_n \sim \mathbf{P}_k$ (unknown)
- A multinomial model $\mathbf{M}_k$ with probability vector

$$\mathbf{p} = (\overbrace{p_1, \ldots, p_s, \underbrace{p, \ldots, p}_{r-s}, 0, \ldots, 0}^{2^d-1 \text{ components}}), \quad p_1 \geq \ldots \geq p_s > p, \quad p \approx 0$$

Model selection for k_n fixed

- For $k_n = k$ fixed, $\mathbf{T}_n \sim \mathbf{P}_k$ (unknown)
- A multinomial model $\mathbf{M}_k$ with probability vector

$$\mathbf{p} = (\overbrace{p_1, \dots, p_s, \underbrace{p, \dots, p}_{r-s}, 0, \dots, 0}^{2^d - 1 \text{ components}}), \quad p_1 \geq \dots \geq p_s > p\,, \quad p \approx 0$$

- We address this question by minimizing the KL divergence

$$KL\big(\mathbf{P}_k \| \mathbf{M}_k\big) \propto -\mathbb{E}\big[\log L_{\mathbf{M}_k}(\mathbf{p}; \mathbf{T}_n)\big]$$

- An estimator:

$$-\mathbb{E}\Big[\log L_{\mathbf{M}_k}(\mathbf{p}; \mathbf{T}_n)\Big]\Big|_{\mathbf{p}=\hat{\mathbf{p}}}$$

Model selection for k_n fixed

- For $k_n = k$ fixed, $\mathbf{T}_n \sim \mathbf{P}_k$ (unknown)
- A multinomial model $\mathbf{M}_k$ with probability vector

$$\mathbf{p} = (\overbrace{p_1, \ldots, p_s, \underbrace{p, \ldots, p}_{r-s}, 0, \ldots, 0}^{2^d-1 \text{ components}}), \quad p_1 \geq \ldots \geq p_s > p\,, \quad p \approx 0$$

- We address this question by minimizing the KL divergence

$$KL\big(\mathbf{P}_k \| \mathbf{M}_k\big) \propto -\mathbb{E}\big[\log L_{\mathbf{M}_k}(\mathbf{p}; \mathbf{T}_n)\big]$$

- An estimator:

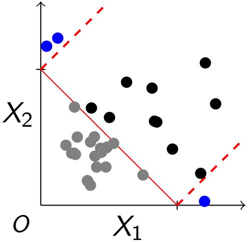
$$-\mathbb{E}\Big[\log L_{\mathbf{M}_k}(\mathbf{p}; \mathbf{T}_n)\Big]\Big|_{\mathbf{p}=\hat{\mathbf{p}}} \approx -\mathbb{E}\big[\log L_{\mathbf{M}_k}(\hat{\mathbf{p}}; \mathbf{T}_n)\big] + s$$

$\hookrightarrow$ Minimize: $-\log L_{\mathbf{M}_k}(\hat{\mathbf{p}}; \mathbf{T}_n) + s$

A method for threshold selection

▷ For a fixed k, a method to choose s

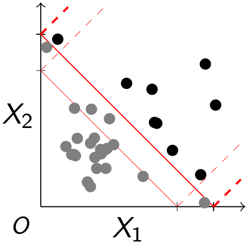
$\hookrightarrow$ How to choose k?



A method for threshold selection

▷ For a fixed k, a method to choose s

$\hookrightarrow$ How to choose k?

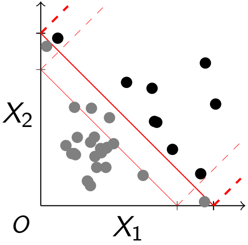


A method for threshold selection

▷ For a fixed k, a method to choose s

$\hookrightarrow$ How to choose k?

▷ $\mathbf{T}'_n := (\mathbf{T}_n, \sum_{j=1}^{n} \mathbb{1}_{\{|\mathbf{X}_j| \le u_n\}}) \sim \mathbf{P}'_n$

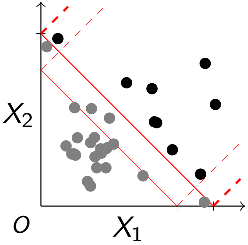


A method for threshold selection

- For a fixed k, a method to choose s

  $\hookrightarrow$ How to choose k?

- $\mathbf{T}'_n := (\mathbf{T}_n, \sum_{j=1}^n \mathbb{1}_{\{|\mathbf{X}_j| \le u_n\}}) \sim \mathbf{P}'_n$



- We add a category and consider the model $\mathbf{M}'_n$ with

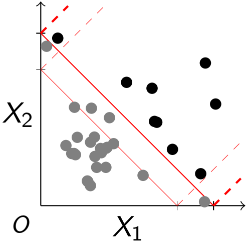
$$\mathbf{p}' = (\overbrace{q' p'_1, \dots, q' p'_{s'}, \underbrace{q' p', \dots, q' p'}_{r' - s'}, 0, \dots, 0}^{2^d - 1 \text{ components}}, 1 - q')\,.$$

A method for threshold selection

- For a fixed k, a method to choose s

  $\hookrightarrow$ How to choose k?

- $\mathbf{T}'_n := (\mathbf{T}_n, \sum_{j=1}^n \mathbb{1}_{\{|\mathbf{X}_j| \le u_n\}}) \sim \mathbf{P}'_n$



- We add a category and consider the model $\mathbf{M}'_n$ with

$$\mathbf{p}' = (\overbrace{q' p'_1, \dots, q' p'_{s'}, \underbrace{q' p', \dots, q' p'}_{r'-s'}, 0, \dots, 0}^{2^d - 1 \text{ components}}, 1 - q') \,.$$

- Similar calculations: minimize the quantity

$$\boxed{\frac{1}{k}\Big(-\log L_{\mathbf{M}_k}(\hat{\mathbf{p}}; \mathbf{T}_n) + (s+1) - k \log(1 - k/n)\Big) \,.}$$

Algorithm: MUSCLE

Algorithm 1: MUltivariate Sparse CLustering for Extremes (MUSCLE)

Data: A sample $\mathbf{X}_1, \ldots, \mathbf{X}_n \in \mathbb{R}_+^d$ and a range of values K for the level

for $k \in K$ **do**

- Compute $u_n = |\mathbf{X}|_{(k+1)}$ the $(k+1)$-th largest norm;
- Assign to each $\pi(\mathbf{X}_j/u_n)$ the subsets C_β it belongs to;
- Compute $\mathbf{T}_n$;
- Compute $\hat{s}(k)$ which minimizes the penalized log-likelihood;

end

Choose $\hat{k}$ which minimizes

$$(-\log L_{\mathbf{M}_k}(\hat{\mathbf{p}}; \mathbf{T}_n) + \hat{s}(k))/k + k/n\,;$$

Output: $\widehat{\mathcal{S}}^* = \{$ the β's associated to $T_{n,j} > 0$ for $j = 1, \ldots, \hat{s}(\hat{k})\}$.

Example: Asymptotic independence

- Asymptotic independence[1] $\iff$ $\mathbf{\Theta}$ places mass only on the axes
  $\iff$ $\mathbf{Z}$ places mass only on the axes
- $n = 30\,000$, $\mathbf{X} \in \mathbb{R}^{40}$, Gaussian dependence, Pareto(1) marginals
  $\hookrightarrow \mathcal{S}^*(\mathbf{Z}) = \{\beta = \{j\}, j = 1, \ldots, 40\}$ and $s^* = 40$

[1] *de Haan & Ferreira (2006), Ledford & Tawn (1996), Heffernan & Tawn (2004)*

Example: Asymptotic independence

- ▷ Asymptotic independence[1] $\iff$ $\boldsymbol{\Theta}$ places mass only on the axes
  $\iff$ $\mathbf{Z}$ places mass only on the axes
- ▷ $n = 30\,000$, $\mathbf{X} \in \mathbb{R}^{40}$, Gaussian dependence, Pareto(1) marginals
  $\hookrightarrow \mathcal{S}^*(\mathbf{Z}) = \{\beta = \{j\}, j = 1, \dots, 40\}$ and $s^* = 40$

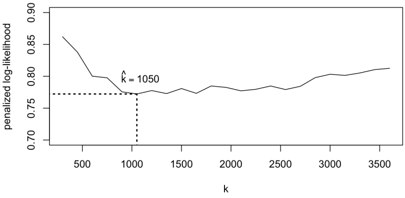


Figure: Evolution of the estimator of $KL(\mathbf{P}_n' \| \mathbf{M}_n')$.

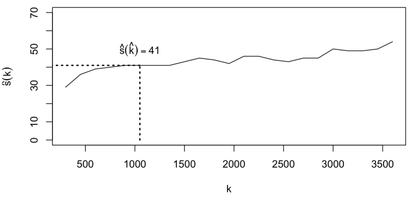


Figure: Evolution of the optimal value of s.

[1] *de Haan & Ferreira (2006), Ledford & Tawn (1996), Heffernan & Tawn (2004)*

Application to extreme variability for financial data

- Data set: value-average daily returns of $d = 49$ industry portfolios in $1970 - 2019$ ($n = 12\,613$)

Application to extreme variability for financial data

▷ Data set: value-average daily returns of $d = 49$ industry portfolios in $1970 - 2019$ ($n = 12\,613$)

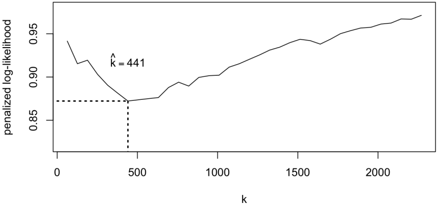


Figure: Evolution of the estimator of $KL(\mathbf{P}'_n \| \mathbf{M}'_n)$.

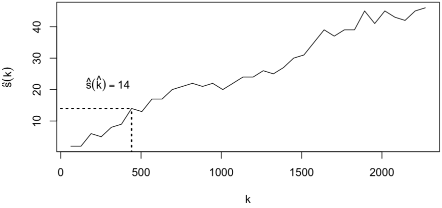
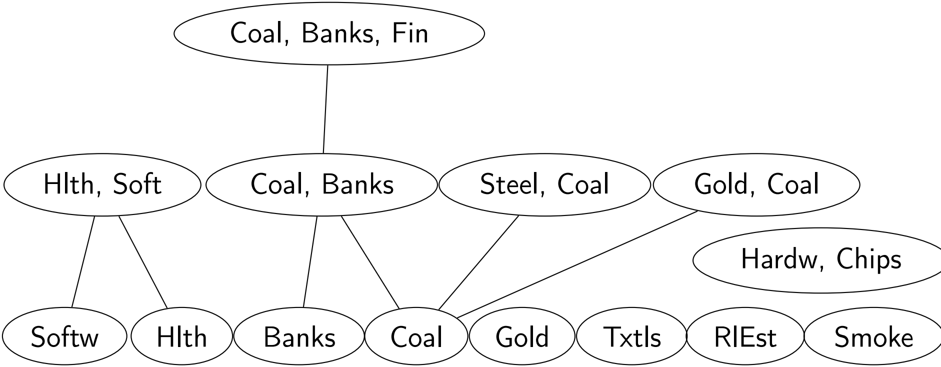


Figure: Evolution of the optimal value of s.

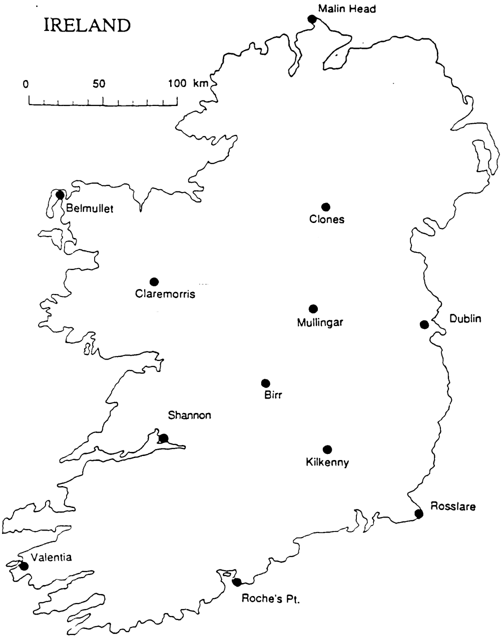
$\hookrightarrow$ We obtain $\hat{k} = 441$ and $\hat{s}(\hat{k}) = 14$ (with $\widehat{\mathcal{S}}^*(\mathbf{Z})$ on the next slide).

Application to extreme variability for financial data



Application to wind speed data

- Data set: daily-average wind speed at $d = 12$ meteorological stations in the Republic of Ireland for 1961-1978 ($n = 6574$)


IRELAND
0
50
100 km
Malin Head
Belmullet
Clones
Claremorris
Mullingar
Dublin
Birr
Shannon
Kilkenny
Rosslare
Valentia
Roche's Pt.

Application to wind speed data

▷ Data set: daily-average wind speed at $d = 12$ meteorological stations in the Republic of Ireland for 1961-1978 ($n = 6574$)

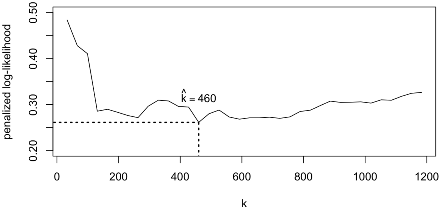


Figure: Evolution of the estimator of $KL(\mathbf{P}_n \| \mathbf{M}_n)$.

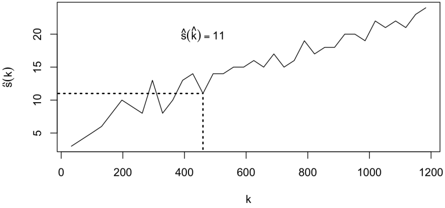


Figure: Evolution of the optimal value of s for the variability.

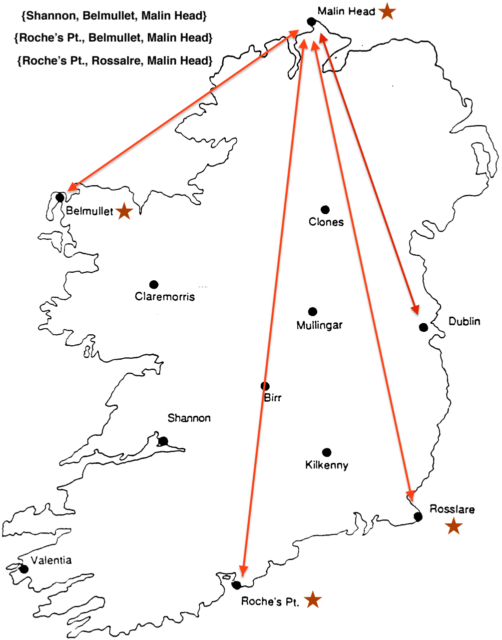
$\hookrightarrow$ We obtain $\hat{k} = 460$ and $\hat{s}(\hat{k}) = 11$ (with $\widehat{\mathcal{S}}^*$ on the next slide)

3-DIMENSIONAL CLUSTERS

{Shannon, Belmullet, Malin Head}

{Roche's Pt., Belmullet, Malin Head}

{Roche's Pt., Rossalre, Malin Head}



Application to wind speed data

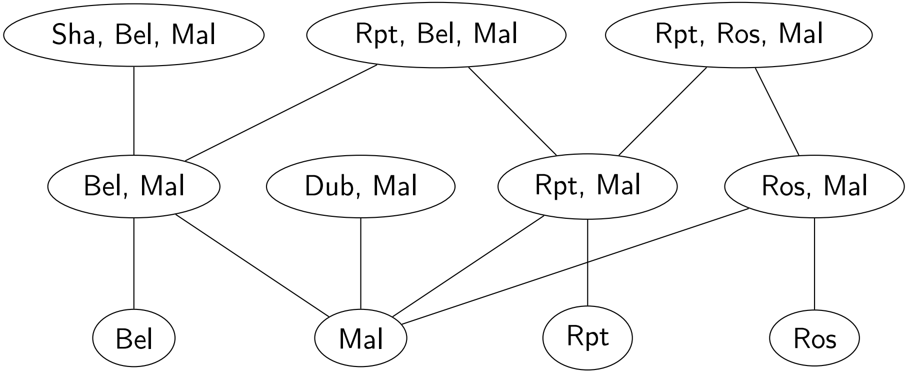


Figure: Representation of the 11 clusters and their inclusions.

Summary and future work

Summary

- An efficient algorithm to study multivariate extremes
- No hyperparameter! (the selection of k is included in the procedure)
- Dimension reduction

Future work

- Extend the work to regularly varying time series
- Study the vector $\mathbf{Z}$ and $\mathbf{\Theta}$ on each cluster

Thank you for your attention!

Meyer & Wintenberger (2021). Sparse regular variation, *Advances in Applied Probability*.
Meyer & Wintenberger (2023+). Tail inference for high-dimensional data, *arXiv:2007.11848*.

https://wintenberger.fr
olivier.wintenberger@sorbonne-universite.fr