

# ÉTUDE DU SUPPORT DE LA MESURE SPECTRALE

Nicolas MEYER <sup>1</sup> & Olivier WINTENBERGER <sup>2</sup>

<sup>1</sup> *Sorbonne Université, LPSM, France, nicolas.meyer@upmc.fr*

<sup>2</sup> *Sorbonne Université, LPSM, France, olivier.wintenberger@upmc.fr*

**Résumé.** Identifier les directions où des événements extrêmes se produisent est un problème majeure de la théorie des valeurs extrêmes multivariées. La mesure exposant de vecteurs à variation régulière permet de déterminer quelles coordonnées contribuent aux extrêmes. Cette mesure est définie par la convergence vague, ce qui ne fournit pas d'estimateur naturel de son support. Un problème analogue apparaît pour la mesure spectrale, qui est l'équivalent de la mesure exposant sur la sphère unité. Dans cette présentation, nous proposons une nouvelle méthode basée sur la projection  $\ell^2$  sur le simplexe. Cette projection permet de détecter sur quelles coordonnées se concentre la masse de la mesure spectrale. À partir de ces résultats et de la théorie de l'apprentissage, un test est proposé pour décider si une direction peut rassembler des événements extrêmes en grande dimension. Nous mettons ainsi en évidence les éventuelles coordonnées sur lesquelles la mesure spectrale ne met pas de masse et réduisons la dimension de l'étude des extrêmes.

**Mots-clés.** Extrêmes multivariés, mesure exposant, mesure spectrale, projection sur le simplexe, réduction de la dimension

**Abstract.** Identifying directions where severe events occur is a major problem of multivariate extreme value theory. The tail measure of regularly varying vectors brings out which coordinates contribute to the extremes. This measure is defined thanks to vague convergence which does not provide a natural estimator of its support. An analogue issue arises for the spectral measure, which is the equivalent of the tail measure on the unit sphere. In this presentation, we propose a new method based on the  $\ell^2$ -projection onto the simplex. This projection allows to detect on which coordinates the spectral measure puts mass. Based on these results and learning theory, a test is proposed to decide if a direction could gather extreme events or not in large dimension. Hence we exhibit the sparsity of the tail measure and reduce the dimension of the extremes' study.

**Keywords.** Dimension reduction, multivariate extremes, projection onto the simplex, spectral measure, tail measure

## 1 Introduction

Dans de nombreuses applications, il est nécessaire d'identifier les directions dans lesquelles les événements extrêmes apparaissent pour évaluer le risque et prédire de futurs événements semblables. En gestion financière du risque, on souhaite détecter les entreprises qui ont une certaine probabilité de subir de grandes pertes. Dans le domaine climatique, il est important d'établir une distinction entre les zones où des précipitations élevées se produisent et celles où ce risque est modéré. Le but de cette présentation est de mettre en évidence deux groupes dans les données : un groupe rassemblant les directions dans lesquelles les événements extrêmes pourraient apparaître, et un autre dans lequel les données ont un comportement plus classique.

On se place dans le cadre de la théorie des valeurs extrêmes qui donne des conditions pour que le maximum d'une suite i.i.d. de vecteurs aléatoires converge. Les trois types de distribution limite (MEV) ont été grandement étudiés. On s'intéresse ici au cas Fréchet qui modélise les phénomènes à queues épaisses, voir par exemple Resnick (2007) ou Embrechts et al. (1997).

## 1.1 Variation régulière multivariée

Un vecteur aléatoire positif  $\mathbf{X}$  converge vers une MEV de type Fréchet si et seulement s'il est à variation régulière : il existe une suite positive  $(a_n)$ ,  $a_n \rightarrow +\infty$  quand  $n \rightarrow \infty$ , et une mesure positive de Radon  $\mu$  sur  $\mathbb{R}_+^d \setminus \{\mathbf{0}\}$  telles que

$$n\mathbb{P}\left(\frac{\mathbf{X}}{a_n} \in \cdot\right) \xrightarrow{v} \mu(\cdot), \quad n \rightarrow +\infty. \quad (1)$$

La mesure limite est appelé *mesure exposant* and décrit le comportement des extrêmes. L'équation (1) correspond à la notion de variation régulière standard. La notion non-standard est quant à elle définie de la manière suivante : il existe une mesure positive de Radon  $\tilde{\mu}$  sur  $\mathbb{R}_+^d \setminus \{\mathbf{0}\}$  telle que

$$n\mathbb{P}((X_i/a_{n,i})_{1 \leq i \leq d} \in \cdot) \xrightarrow{v} \tilde{\mu}(\cdot), \quad n \rightarrow +\infty,$$

où les suites  $(a_{n,i})_n$  vérifient  $n\mathbb{P}(X_i > a_{n,i}) \rightarrow 1$  quand  $n \rightarrow +\infty$ .

La notion de variation régulière standard (1) est plus générale puisque dans ce cas la mesure exposant  $\mu$  peut être "creuse", c'est-à-dire qu'on peut trouver des directions qui ne sont pas chargées par cette mesure. On se place dans ce cadre et le but principal de la présentation est d'identifier ces directions pour réduire la dimension de l'étude.

Une manière équivalente de définir la variation régulière est la suivante. Un vecteur aléatoire positif  $\mathbf{X}$  satisfait la condition (1) si et seulement s'il existe  $\alpha > 0$  et un vecteur aléatoire  $\Theta$  sur la sphère unité  $\mathbb{S}_+^{d-1} := \{\mathbf{x}/|\mathbf{x}|, \mathbf{x} \in \mathbb{R}_+^d \setminus \{\mathbf{0}\}\}$  telle que

$$\mathbb{P}\left(\left(\frac{|\mathbf{X}|}{t}, \frac{\mathbf{X}}{|\mathbf{X}|}\right) \in \cdot \mid |\mathbf{X}| > t\right) \xrightarrow{d} \mathbb{P}((Y, \Theta) \in \cdot), \quad t \rightarrow +\infty, \quad (2)$$

où  $Y$  suit une loi de Pareto de paramètre  $\alpha$  et est indépendant de  $\Theta$ . Le vecteur aléatoire  $\Theta$  est appelé le *vecteur spectral* et sa loi est appelée la *mesure spectrale*. On note alors  $\mathbf{X} \in \text{RV}(\alpha, \Theta)$ .

## 1.2 Estimation de la mesure spectrale : restriction de la projection classique

En considérant uniquement la deuxième coordonnée dans l'équation (2), on obtient

$$\mathbb{P}(\mathbf{X}/|\mathbf{X}| \in \cdot \mid |\mathbf{X}| > t) \xrightarrow{d} \mathbb{P}(\Theta \in \cdot), \quad t \rightarrow +\infty. \quad (3)$$

Cela signifie que le vecteur spectral  $\Theta$  peut être approché par la donnée extrême  $\mathbf{X}/|\mathbf{X}| \mid |\mathbf{X}| > t$ , pour  $t$  "assez grand". La loi de  $\Theta$  rassemble donc l'information sur le comportement des événements extrêmes. C'est pourquoi, estimer la mesure spectrale est une question cruciale, mais difficile, de la théorie des extrêmes multivariés. Différentes méthodes ont été proposées par Chauvru (2015), Chiapino et Sabourin (2016), Cooley et Thibaud (2017), ou encore Goix et al. (2017).

**Sparsity pattern** Dans le but d'étudier le support de  $\Theta$ , on s'intéresse aux directions  $\{1, \dots, d\}$  de  $\mathbb{S}_+^{d-1}$  chargées par la mesure spectrale. En effet, il arrive souvent, essentiellement en grande dimension, que toutes les coordonnées ne contribuent pas aux extrêmes, de sorte que la mesure spectrale est creuse. Les directions chargées par la mesure spectrale sont celles qui rassemblent les événements extrêmes. Identifier ces directions est donc une première étape importante pour comprendre la localisation des extrêmes. Ainsi, il est utile de calculer les probabilités  $\mathbb{P}(\Theta_i = 0)$  ou  $\mathbb{P}(\Theta_i > 0)$ , pour  $i \in \{1, \dots, d\}$  fixé. Ceci mène à la définition suivante.

**Définition 1.** Si  $\mathbf{Z}$  est un vecteur aléatoire de  $\mathbb{R}^d$ , on appelle sparsity pattern de  $\mathbf{Z}$ , et on note  $\text{sp}(\mathbf{Z})$ , l'ensemble des coordonnées  $i \in \{1, \dots, d\}$  telles que  $\mathbb{P}(Z_i > 0) > 0$ . De manière équivalente,

$$i \notin \text{sp}(\mathbf{Z}) \quad \text{if and only if} \quad \mathbb{P}(Z_i = 0) = 1.$$

Le sparsity pattern d'un vecteur aléatoire  $\mathbf{Z}$  rassemble les coordonnées chargées par la loi de  $\mathbf{Z}$ . En grande dimension, connaître le sparsity pattern d'un vecteur est utile puisque cela permet de réduire la dimension en s'intéressant uniquement aux coordonnées qui y appartiennent.

Cependant, l'équation (3) ne peut pas être utilisée directement pour déterminer le sparsity pattern du vecteur spectral  $\Theta$ . En effet, la convergence étroite tombe souvent en défaut pour des ensembles du type  $\{\Theta_i = 0\}$ . Cela signifie que même pour  $t$  grand, la loi de  $X_i/|\mathbf{X}|$  conditionnellement à  $\{|\mathbf{X}| > t\}$  ne permet pas d'estimer  $\Theta_i$ . Illustrons ceci sur un exemple. Supposons qu'une direction, disons  $i$ , n'est pas chargée par la mesure spectrale. Cela signifie que  $i \notin \text{sp}(\Theta) : \mathbb{P}(\Theta_i = 0) = 1$ . Si la marginale  $X_i$  ne charge pas le point  $\{0\}$ , alors elle vérifie  $\mathbb{P}(X_i/|\mathbf{X}| = 0 \mid |\mathbf{X}| > t) = 0$  pour tout  $t > 0$ . Donc le comportement de  $\Theta_i$  et  $X_i/|\mathbf{X}| \mid |\mathbf{X}| > t$  diffèrent complètement.

Ainsi la projection classique  $\mathbf{X}/|\mathbf{X}| \mid |\mathbf{X}| > t$  ne permet pas de déterminer les possibles directions non chargées par la mesure spectrale. L'objectif de la présentation est de contourner ce problème en utilisant une autre projection. On remplace la quantité  $\mathbf{X}/|\mathbf{X}|$  par la projection  $\ell^2$  de  $\mathbf{X}/t$  sur le simplexe. Il s'agit donc d'adapter l'équation (3), ce qui est l'objet de la section suivante.

Dans toute la suite, on considère un entier  $d \geq 2$ . On désigne par  $|\cdot|$  la norme  $\ell^1$  sur  $\mathbb{R}^d$ . On note  $\mathbb{S}^{d-1}$  la sphère unité de  $\mathbb{R}^d$  pour cette norme et  $\mathbb{S}_+^{d-1}$  sa restriction au quadrant positif, c'est-à-dire le simplexe de  $\mathbb{R}^d : \mathbb{S}_+^{d-1} = \mathbb{S}^{d-1} \cap [0, \infty)^d = \{\mathbf{x} \in \mathbb{R}_+^d, x_1 + \dots + x_d = 1\}$ .

## 2 Projection sur le simplexe et variation régulière

On définit la projection  $\ell^2$  sur le simplexe de la manière suivante (voir l'article de Duchi et al. (2008) pour plus de détails) : projeter le vecteur  $\mathbf{x} \in \mathbb{R}_+^d$  sur la sphère  $\ell^1$  de rayon  $s > 0$  consiste à trouver un vecteur  $\mathbf{v}$  sur cette sphère qui minimise la quantité  $\|\mathbf{x} - \mathbf{v}\|_2$ , où  $\|\cdot\|_2$  désigne la norme euclidienne. On peut montrer que le vecteur  $\mathbf{v}$  est alors unique. Notons  $\pi_s$  l'application de projection :  $\pi_s : \mathbb{R}_+^d \rightarrow \mathbb{S}_+^{d-1}$ ,  $\mathbf{x} \rightarrow \mathbf{v}$ . La projection  $\pi_1$  sur la simplexe sera simplement notée  $\pi$ . La Figure 1 illustre la projection  $\pi$  en dimension 2.

La projection  $\pi$  étant continue, on peut adapter l'hypothèse de variation régulière de la manière suivante.

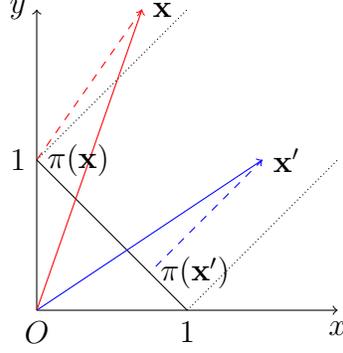


Figure 1: La projection  $\ell^2$  sur le simplexe de  $\mathbb{R}^2$ .

**Lemme 1.** Soit  $\mathbf{X}$  un vecteur aléatoire de  $\mathbb{R}_+^d$  à variation régulière :  $\mathbf{X} \in \text{RV}(\alpha, \Theta)$ . Alors

$$\mathbb{P} \left( \left( \frac{|\mathbf{X}|}{t}, \pi \left( \frac{\mathbf{X}}{t} \right) \right) \in \cdot \mid |\mathbf{X}| > t \right) \xrightarrow{d} \mathbb{P}((Y, \pi(Y\Theta)) \in \cdot), \quad t \rightarrow +\infty.$$

où  $Y$  suit une loi de Pareto de paramètre  $\alpha$  et est indépendante de  $\Theta$ .

De plus,

$$\mathbb{P}(\pi(\mathbf{X}/t)_i = 0 \mid |\mathbf{X}| > t) \xrightarrow{d} \mathbb{P}(\pi(Y\Theta)_i = 0), \quad t \rightarrow +\infty, \quad (4)$$

ou, de manière équivalente,

$$\mathbb{P}(\pi(\mathbf{X}/t)_i > 0 \mid |\mathbf{X}| > t) \xrightarrow{d} \mathbb{P}(\pi(Y\Theta)_i > 0), \quad t \rightarrow +\infty. \quad (5)$$

Si  $\mathbf{X}$  est un vecteur aléatoire de  $\mathbb{R}_+^d$  et  $t$  un réel positif ( $t$  joue le rôle de seuil), on considère donc le projeté  $\pi(\mathbf{X}/t)$ . Le lemme 1 permet de contourner les problèmes qui apparaissent dans la convergence étroite pour des ensembles du type  $\{\Theta_i = 0\}$ . Néanmoins, les limites obtenues en (4) et (5) dépendent de la projection  $\pi$ . Il s'agit donc de montrer que les sparsity pattern de  $\Theta$  et  $\pi(Y\Theta)$  sont les mêmes. C'est l'objet de la proposition suivante.

**Proposition 1.** Soit  $\mathbf{X} \in \text{RV}(\alpha, \Theta)$ , et  $i \in \{1, \dots, d\}$ .

1. Si  $\mathbb{P}(\Theta_i = 0) = 1$ , alors  $\mathbb{P}(\pi(Y\Theta)_i = 0) = 1$ .
2. Si  $\mathbb{P}(\Theta_i = 0) = p < 1$ , alors  $p \leq \mathbb{P}(\pi(Y\Theta)_i = 0) < 1$ .

On en déduit donc que  $\text{sp}(\Theta) = \text{sp}(\pi(Y\Theta))$ . Ainsi, en utilisant le lemme 1, on peut estimer le sparsity pattern de  $\Theta$  via la quantité  $\pi(\mathbf{X}/t)$  pour  $\mathbf{X} > t$ .

### 3 Processus empirique

On considère à présent une suite i.i.d.  $\mathbf{X}^1, \dots, \mathbf{X}^n$  de vecteurs aléatoires à variation régulière de loi  $\mathbf{X} \in \text{RV}(\alpha, \Theta)$ . On fixe une coordonnée  $i \in \{1, \dots, d\}$ . Le but est de déterminer si la coordonnée  $i$  appartient au sparsity pattern de  $\Theta$  ou non. D'après la section précédente, on peut étudier de manière équivalente le sparsity pattern de  $\pi(Y\Theta)$ .

Afin d'estimer la probabilité  $\mathbb{P}(\pi(Y\Theta)_i > 0)$ , on considère la probabilité pré-asymptotique  $\mathbb{P}(\pi(\mathbf{X}/u_n)_i > 0 \mid |\mathbf{X}| > u_n)$ , où  $(u_n)$  est une suite telle que  $u_n \rightarrow +\infty$  quand  $n \rightarrow +\infty$ . En utilisant (5), on sait  $\mathbb{P}(\pi(\mathbf{X}/u_n)_i > 0 \mid |\mathbf{X}| > u_n) \rightarrow \mathbb{P}(\pi(Y\Theta)_i > 0)$  quand  $n \rightarrow +\infty$ . Un estimateur naturel  $\hat{T}_n$  de la probabilité pré-asymptotique  $\mathbb{P}(\pi(\mathbf{X}/u_n)_i > 0 \mid |\mathbf{X}| > u_n)$  est alors défini par

$$\hat{T}_n := \sum_{j=1}^n \mathbf{1}_{\pi(\mathbf{X}^j/u_n)_i > 0, |\mathbf{X}^j| > u_n}.$$

On définit également  $k_n = n\mathbb{P}(|\mathbf{X}| > u_n)$  et  $\hat{k}_n = \sum_{j=1}^n \mathbf{1}_{|\mathbf{X}^j| > u_n}$ , de sorte que  $\mathbb{E}[\hat{T}_n] = k_n\mathbb{P}(\pi(\mathbf{X}/u_n)_i > 0 \mid |\mathbf{X}| > u_n)$ , et  $\hat{k}_n/k_n \rightarrow 1$  en probabilité quand  $n \rightarrow +\infty$ .

On en déduit alors les résultats asymptotiques suivants.

**Théorème 1.** *On considère une suite i.i.d.  $\mathbf{X}^1, \dots, \mathbf{X}^n$  de vecteurs aléatoires à variation régulière de loi  $\mathbf{X} \in \text{RV}(\alpha, \Theta)$ . On fixe une coordonnée  $i \in \{1, \dots, d\}$  et un seuil  $u_n \rightarrow +\infty$ . On suppose que  $k_n = n\mathbb{P}(|\mathbf{X}| > u_n) \rightarrow +\infty$ . Alors*

1. *Convergence de l'estimateur.*

$$\frac{\hat{T}_n}{\hat{k}_n} \rightarrow \mathbb{P}(\pi(Y\Theta)_i > 0), \quad n \rightarrow +\infty, \quad \text{en probabilité.}$$

2. *Normalité asymptotique.* On suppose par ailleurs que  $\mathbb{P}(\pi(Y\Theta)_i > 0) > 0$ . Alors

$$\sqrt{k_n} \frac{\hat{T}_n/\hat{k}_n - \mathbb{P}(\pi(\mathbf{X}/u_n)_i > 0 \mid |\mathbf{X}| > u_n)}{\sqrt{\mathbb{P}(\pi(\mathbf{X}/u_n)_i > 0 \mid |\mathbf{X}| > u_n)[1 - \mathbb{P}(\pi(\mathbf{X}/u_n)_i > 0 \mid |\mathbf{X}| > u_n)]}} \xrightarrow{d} N, \quad n \rightarrow +\infty,$$

où  $N \sim \mathcal{N}(0, 1)$ .

De plus, si on suppose que

$$\sqrt{k_n}[\mathbb{P}(\pi(\mathbf{X}/u_n)_i > 0 \mid |\mathbf{X}| > u_n) - \mathbb{P}(\pi(Y\Theta)_i > 0)] \rightarrow 0,$$

on obtient alors :

$$\sqrt{k_n} \frac{\hat{T}_n/\hat{k}_n - \mathbb{P}(\pi(Y\Theta)_i > 0)}{\sqrt{\mathbb{P}(\pi(\mathbf{X}/u_n)_i > 0 \mid |\mathbf{X}| > u_n)[1 - \mathbb{P}(\pi(\mathbf{X}/u_n)_i > 0 \mid |\mathbf{X}| > u_n)]}} \xrightarrow{d} N, \quad n \rightarrow +\infty, \quad (6)$$

où  $N \sim \mathcal{N}(0, 1)$ .

**Remarque 1.** Si  $\mathbb{P}(\pi(Y\Theta)_i > 0) = 0$ , alors le membre de gauche de l'équation (6) est positif p.s. donc ne converge pas vers une gaussienne. On en déduit donc une différence de comportement selon que  $i$  appartienne à  $\text{sp}(\pi(Y\Theta))$  ou non.

On s'intéresse alors au test suivant :

$$H_0: \mathbb{P}(\pi(Y\Theta)_i > 0) > 0 \quad \text{vs} \quad H_1: \mathbb{P}(\pi(Y\Theta)_i > 0) = 0.$$

Par égalité des sparsity pattern (cf Proposition 1), cela revient à tester :  $i \in \text{sp}(\Theta)$  vs  $i \notin \text{sp}(\Theta)$ . Pour mettre en oeuvre ce test, on utilise une version bootstrap du Théorème 1.

**Théorème 2.** On considère une suite i.i.d.  $\mathbf{X}_1, \dots, \mathbf{X}_n$  de vecteurs aléatoires à variation régulière de loi  $\mathbf{X} \in \text{RV}(\alpha, \Theta)$ , une suite i.i.d. de variables aléatoires de loi de Poisson  $P_1, \dots, P_n$  de paramètre 1, et  $u_n \rightarrow +\infty$ . On définit  $k_n = n\mathbb{P}(|\mathbf{X}| > u_n)$  et  $\hat{k}_n = \sum_{j=1}^n \mathbb{1}_{|\mathbf{X}^j| > u_n}$ , et on suppose que  $k_n \rightarrow +\infty$ . On fixe une coordonnée  $i$  dans  $\{1, \dots, d\}$ . On définit

$$\hat{T}_n^* = \sum_{j=1}^n P_j \mathbb{1}_{\pi(\mathbf{X}^j/u_n)_i > 0, |\mathbf{X}^j| > u_n} \quad \text{and} \quad \hat{k}_n^* = \sum_{j=1}^n P_j \mathbb{1}_{|\mathbf{X}^j| > u_n}.$$

Alors, si  $\mathbb{P}(\pi(Y\Theta)_i > 0) > 0$ ,

$$\sup_{g \in \text{BL}_1(\mathbb{R})} \left| \mathbb{E}_P \left[ g \left( \sqrt{\hat{k}_n} (\hat{T}_n / \hat{k}_n - \hat{T}_n^* / \hat{k}_n^*) \right) \right] - \mathbb{E}[g(N)] \right| \xrightarrow{d} 0, \quad n \rightarrow +\infty, \quad \text{en probabilité,}$$

où  $N \sim \mathcal{N}(0, \mathbb{P}(\pi(Y\Theta)_i > 0)[1 - \mathbb{P}(\pi(Y\Theta)_i > 0)])$ .

Le théorème 2 assure que sous  $H_0$  la loi de  $\sqrt{\hat{k}_n}(\hat{T}_n/\hat{k}_n - \hat{T}_n^*/\hat{k}_n^*)$  conditionnellement à  $\mathbf{X}^1, \dots, \mathbf{X}^n$  converge vers une loi normale centrée réduite de variance  $\mathbb{P}(\pi(Y\Theta)_i > 0)[1 - \mathbb{P}(\pi(Y\Theta)_i > 0)]$ . Le lemme de Slutsky et la Proposition 1 permettent de remplacer  $\sqrt{\hat{k}_n}$  par  $\sqrt{\hat{k}_n}$  et la variance théorique par son estimateur convergent naturel

$$\frac{\hat{T}_n}{\hat{k}_n} \left( 1 - \frac{\hat{T}_n}{\hat{k}_n} \right) = \frac{\sum_{j=1}^n \mathbb{1}_{\pi(\mathbf{X}^j/u_n)_i > 0, |\mathbf{X}^j| > u_n}}{\sum_{j=1}^n \mathbb{1}_{|\mathbf{X}^j| > u_n}} \left( 1 - \frac{\sum_{j=1}^n \mathbb{1}_{\pi(\mathbf{X}^j/u_n)_i > 0, |\mathbf{X}^j| > u_n}}{\sum_{j=1}^n \mathbb{1}_{|\mathbf{X}^j| > u_n}} \right).$$

Ainsi, sous  $H_0$ , on obtient la convergence suivante :

$$\sqrt{\hat{k}_n} \frac{\frac{\hat{T}_n^*}{\hat{k}_n^*} - \frac{\hat{T}_n}{\hat{k}_n}}{\sqrt{\frac{\hat{T}_n}{\hat{k}_n} \left( 1 - \frac{\hat{T}_n}{\hat{k}_n} \right)}} \mid \mathbf{X}^1, \dots, \mathbf{X}^n \xrightarrow{d} N, \quad n \rightarrow +\infty, \quad (7)$$

D'après la Remarque 1, cette convergence tombe en défaut sous  $H_1$ .

La normalité asymptotique (7) sous  $H_0$  permet de construire un test asymptotique. On teste la normalité du membre de gauche de (7) via des tests classiques, par exemple celui de Shapiro-Wilk. La présentation s'attachera à mettre en pratique ce test sur des données simulées, ainsi que sur des données climatiques et financières.

## Bibliographie

- Chautru, E. (2015), *Dimension reduction in multivariate extreme value analysis*, Electronic Journal of Statistics.
- Chiapino, M. Sabourin, A. (2016) *Feature Clustering for Extreme Events Analysis, with Application to Extreme Stream-Flow Data*, ECML-PKDD 2016.
- Cooley, D. Thibaud, E. (2017), *Decompositions of Dependence for High-Dimensional Extremes*, pré-publication.
- Duchi, J. Shalev-Shwartz, S. Singer, Y. Chandra, T. (2008), *Efficient Projections onto the  $\ell_1$ -Ball for Learning in High Dimensions*, ICML.
- Embrechts, P. Klüppelberg, C. Mikosch, T. (1997), *Modelling Extremal Events for Insurance and Finance*, Springer
- Goix, N. Sabourin, A. Cléménçon, S. (2017), *Sparsity in Multivariate Extremes with Applications to Anomaly Detection*, Journal of Multivariate Analysis.
- Resnick, S.I. (2007), *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling*, Springer.