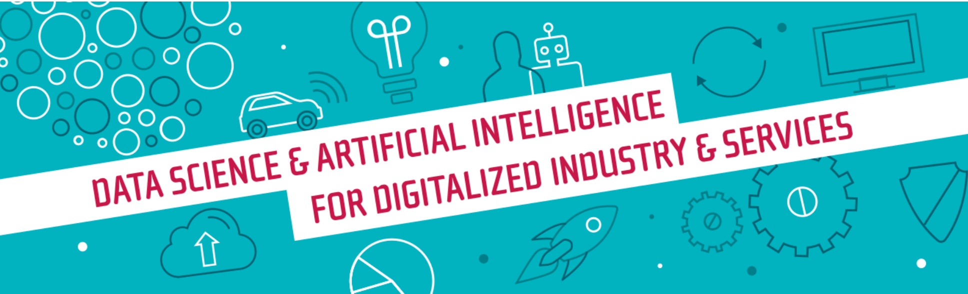
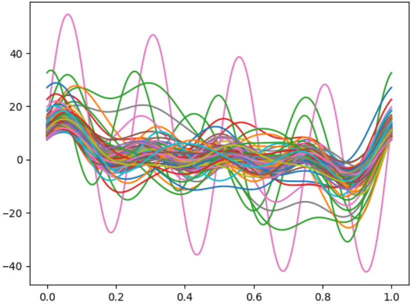
Karhunen-Loève Expansion for Functional Extremes

Nathan Huet
with Anne Sabourin and Stephan Clémençon



Motivations - Context



- **Functional Data**: data depending on continuous variable, especially time and space.

→ Data in many fields increasingly come to us with functional structure.

+ **Extremes**: Data "large" in some sense.

Main issues:

- infinite dimension (or at least high dimension in pratice);
- representation of functional extremes;
- how to work with only a few data?

Goal: developing **functional extreme theory** in order to elaborate practical methods to handle functional issues.

Context

- Object: X a zero-mean and second order random process with sample-path in $L^2([0,1])$, *i.e.* the space of functions f over $[0,1]$ such that $\|f\|_2 := (\int_0^1 |f(t)|^2 dt)^{1/2} < \infty$.

$\Rightarrow$ Dependence structure monitored by covariance function $C(s,t) = \mathbb{E}(X(s)X(t))$ and, *a fortiori*, by covariance operator $T_C(f)(s) = \int_0^1 C(s,t)f(t)dt$.

$\Rightarrow$ Classical tool to study that type of object: **Karhunen-Loève Expansion** (KLE).

Karhunen-Loève Expansion

KLE: allows to reduce dimension of data by selecting directions where the information is the more spread, *i.e.* where the variance is the highest. → analogous to *principal components analysis* in finite dimension.

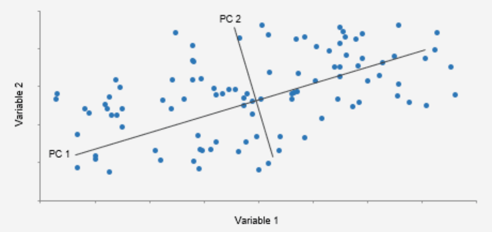


Figure: Illustration of PCA in 2D

Applications: **compression**, **denoising**, **signal estimation**...

Karhunen-Loève Expansion

Mathematical definition

KLE of X is given by:

$$X = \sum_{i=1}^{\infty} Z_i \phi_i$$

where $Z_i = \langle X, \phi_i \rangle$ and ϕ_i are eigenfunctions of T_C forming a Hilbert basis of $L^2([0, 1])$.

$\Rightarrow$ KLE decomposes X into *bi-orthogonal* expansion

- Z_i's are decorrelated: $\mathbb{E}(Z_i Z_j) = 0$ if $i \neq j$;
- ϕ_i's are orthogonal: $\langle \phi_i, \phi_j \rangle = 0$ if $i \neq j$.

Main advantages:

- **best linear approximation** at given dimension;
- **reduces overfitting**.

Karhunen-Loève Expansion

Best linear approximation?

- Reconstruction error: $R(V) = \mathbb{E}\|X - \Pi_V X\|_2^2$
  $\rightarrow$ empirical version: $\hat{R}(V) = \frac{1}{n}\sum_{i=1}^{n} \|X_i - \Pi_V X_i\|_2^2$;

$\Rightarrow$ KLE is the best linear approximation at given dimension N in the sense that

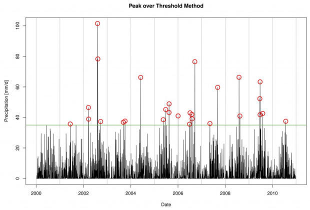
$$\min_{V, dim(V)=N} R(V)$$

is achieved for $V = span\{\phi_1, ..., \phi_N\}$, *i.e.* $\Pi_V X = \sum_{i=1}^{N} Z_i \phi_i$.

Powerful tool but how to extend it to extremal framework?

Multivariate Extreme Value Theory

see [3, Resnick, 1987]



Extreme Value Theory (EVT): study of large data, *i.e.* which exceed a high threshold.

$\to$ modeling rare events and risk monitoring (in hydrology or insurance for instance).

Regular variation (RV) of X (:= classical hypothesis in EVT) := the law of rescaled data given an excess of a high threshold has a limit μ, called *exponent measure*:

$$\mathbb{P}(X/t \in \mathscr{A} \,|\, \|X\| \geqslant t) \xrightarrow[t\to\infty]{} \mu(\mathscr{A})$$

Property of μ: $\mu(t\mathscr{A}) = t^{-\alpha}\mu(\mathscr{A})$

$\Rightarrow$ suggests considering a limiting angular variable Θ_∞, such that

$$\frac{\mathbb{P}(\|X\| \geqslant tx, X/\|X\| \in \cdot)}{\mathbb{P}(\|X\| \geqslant t)} \xrightarrow{v} x^{-\alpha}\mathbb{P}(\Theta_\infty \in \cdot).$$

Regular Variation

Illustration

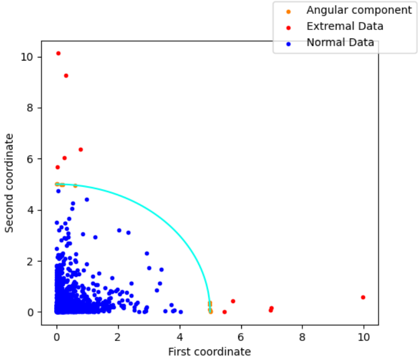


Figure: Plot of Regularly Varying Random Variable in 2D.

$\Rightarrow \Theta_\infty$ characterizes the **dependence structure** in extremes.

What about EVT in the functional case?

Functional Regular Variation

$\Longrightarrow$ same characterization than multivariate RV but weak convergence and measure are less tractable and mentally representable.

Main difference with multivariate case: several representations for functional extremes are possible, including

- high value at one point, measured with $\sup_{t \in T} |x(t)|$;
- high energy over a catchment T, measured with $\int_T x(t)^2 dt$;
- large total amount over a catchment T, measured with $\int_T x(t) dt$.

Our focus: data with high energies.

Our goal: characterizing functional extremes and obtain a suitable representation of finite dimension for limiting measure.

Karhunen-Loève Expansion of Extremes

The limit of KLE is KLE of the limit

- **Working Assumption**: X is regularly varying with tail process Θ_∞ which belongs on a finite-dimensional space of dimension p, noted V_∞.
- **Goals**: characterization of behavior of KLE of largest functional data and estimation of KLE of V_∞.
- working on Θ_t the normalized thresholded process such that $\mathscr{L}(\Theta_t) := \mathscr{L}\left(\frac{X}{\|X\|} \middle| \|X\| \geqslant t\right)$ to avoid moment issues with V_t^p = p-dimensional space given by KLE on Θ_t.

Theorem(*Limit behavior of thresholded spaces*)[HUET2022]

$$\rho(V_t^p, V_\infty) \underset{t\to\infty}{\longrightarrow} 0$$

where $\rho(A, B) := \|\Pi_A - \Pi_B\|_{op}$ is a distance between spaces.

Karhunen-Loève Expansion of Extremes

Statistical guarantees to recover KLE of the limit

- **Second step**: estimating V_∞ thanks to a sample $(X_1, ..., X_n)$ of independent observations following the same distribution as X.

$\Rightarrow$ $\hat{V}^p_{\hat{t}_{n,k}} = p$-dimensional space given by KLE on $(X_1, ..., X_n)$ (using only the k larger data);

$\Rightarrow$ estimator of $V^p_{t_{n,k}}$ only with known quantities.

Theorem(*Convergence rate and consistency*)[HUET2022]

$$\rho(\hat{V}^p_{\hat{t}_{n,k}}, V^p_{t_{n,k}}) \leqslant \frac{C_k}{\sqrt{k}} + o\Big(\frac{1}{\sqrt{k}}\Big)$$

with large probability and where $C_k \underset{k\to+\infty}{\longrightarrow} C \in \mathbb{R}$.

$\Rightarrow$ extension to infinite dimension of [1, Sabourin and Drees, 2021].

Recovering extremal signals with small distortion

Plot of the data

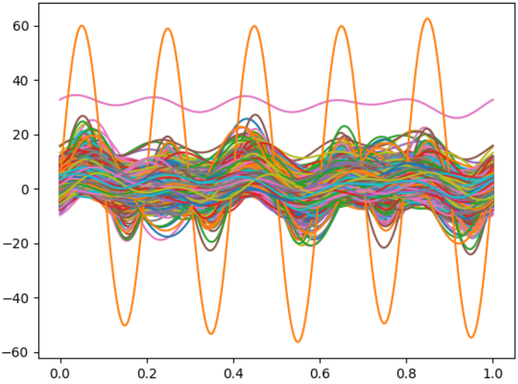


Figure: Plot of simulated data

Recovering extremal signals with small distortion

Settings

Simulation of $n = 10.000$ curves such that
$X_i = \sum_{j \in I_1} Y_{ij} e_j + \sum_{j \in I_2} W_{ij} e_i$ with

- $Y_{ij} \overset{i.i.d.}{\sim} Pareto(3)$ (regularly varying coordinates);
- $W_{ij} \overset{i.i.d.}{\sim} \mathcal{N}(0, 3/4)$ (non-regularly varying coordinates);
- $\{e_j\}_{j \geqslant 1}$ Fourier basis;
- $I_1 = \{1, 6, 10\}$ and $I_2 = \{2, 3, 4, 5, 7, 8, 9, 11\}$ (arbitrary partition of $\{1, ..., 11\}$.

$\Rightarrow Var(Y_{ij}) = Var(W_{ij}), \forall i.$

Recovering extremal signals with small distortion

First KL analysis performed on normalized data

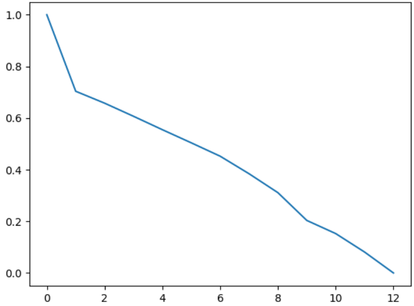


Figure: Mean empirical reconstruction error for PCA, performed on angle of the process without any thresholding, projecting on subspace of dimension $1 \leqslant s \leqslant 11$ versus dimension s.

Recovering extremal signals with small distortion

Plot of thresholded data

Only the $k = 200$ largest curves (w.r.t. $\|\cdot\|_2$) remain.

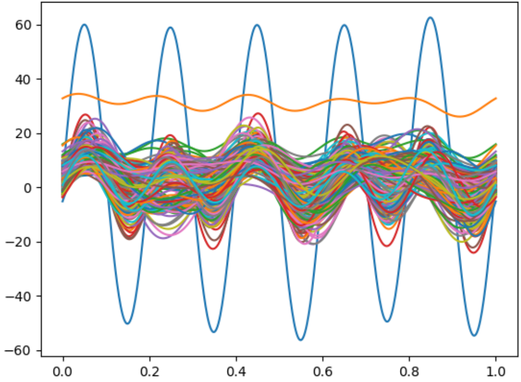


Figure: Plot of the k largest curves.

Recovering extremal signals with small distortion

Second KL analysis performed on normalized and thresholded data

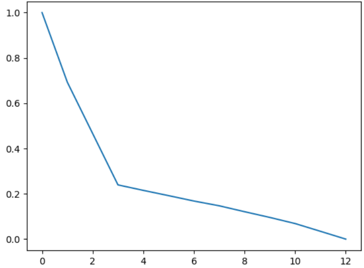


Figure: Mean empirical reconstruction errors for PCA, performed on angle of the thresholded process, projecting on subspace of dimension $1 \leqslant s \leqslant 11$ versus k.

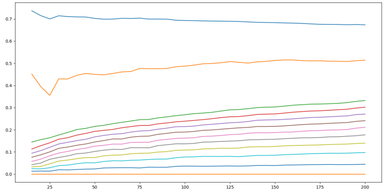


Figure: Mean empirical reconstruction errors for PCA, performed on angle of the thresholded process, projecting on subspace of dimension $1 \leqslant s \leqslant 11$ versus k.

On going/Remaining Work

- Apply those results to **real datasets** with in particular **functional anomaly detection** (as in [2, Goix et al., 2017]);
- Obtain other functional representations by considering different basis (wavelets...);
- Develop in detail Hilbertian regular variation theory.

Thanks for your attention !

Holger Drees and Anne Sabourin.
Principal component analysis for multivariate extremes.
Electronic Journal of Statistics, 15(1):908–943, 2021.

Nicolas Goix, Anne Sabourin, and Stephan Clémençon.
Sparse representation of multivariate extremes with applications to anomaly detection.
Journal of Multivariate Analysis, 161:12–31, 2017.

Sidney I. Resnick.
Extreme Values, Regular Variation and Point Processes.
Springer-Verlag New York, 1987.