

REDUCTION DE DIMENSION POUR L'ESTIMATION DE L'INDICE DES VALEURS EXTRÊMES CONDITIONNEL

Laurent Gardes ¹ & Alex Podgorny ²

¹ *Université de Strasbourg, CNRS, IRMA UMR 7501, F-67000 Strasbourg, France. E-mail : gardes@unistra.fr*

² *Université de Strasbourg, CNRS, IRMA UMR 7501, F-67000 Strasbourg, France. E-mail : apodgorny@unistra.fr*

Résumé. Nous nous intéressons à la relation liant les grandes valeurs d'une variable aléatoire réelle Y à une covariable X prenant ses valeurs dans un sous-ensemble \mathcal{X} de \mathbb{R}^p . Pour ce faire, nous supposons que la loi conditionnelle de Y sachant $X = x$ est une loi à queue lourde d'indice des valeurs extrêmes $\gamma(x)$. L'estimation de cet indice est une étape essentielle pour l'inférence de la loi conditionnelle mais cette tâche est d'autant plus délicate que la dimension p augmente. L'objectif de ce travail est de proposer une méthode de réduction de dimension dans le but d'obtenir un estimateur plus efficace de l'indice des valeurs extrêmes. Plus précisément, nous supposons qu'il existe un sous-espace \mathcal{S}_0 de dimension $q < p$ de base $B_0 \in \mathbb{R}^{p \times q}$ et une fonction positive $g(\cdot)$ tels que pour tout $x \in \mathcal{X}$, $\gamma(x) = g(B_0^\top x)$. Nous proposons une méthode d'estimation de ce sous-espace et en établissons la consistance. Nous illustrons les avantages de cette procédure de réduction de dimension pour l'estimation de l'indice de valeurs extrêmes conditionnel à l'aide de simulations.

Mots-clés. Indice des valeurs extrêmes, réduction de dimension, lois à queue lourde.

Abstract. We are interested in the relationship between the large values of a real random variable Y and its associated covariate X that takes its values in a subset \mathcal{X} of \mathbb{R}^p . To do this, we assume that the conditional distribution of Y given $X = x$ has a heavy-tailed distribution with tail index $\gamma(x) > 0$. Estimating this index is a crucial step for the inference of the conditional distribution, but this task becomes more challenging as the dimension p increases. The objective of this work is to propose a dimension reduction method to obtain a more efficient estimator of the extreme value index. Specifically, we assume the existence of a subspace \mathcal{S}_0 of dimension $q < p$ with basis $B_0 \in \mathbb{R}^{p \times q}$ and a positive function $g(\cdot)$ such that for all $x \in \mathcal{X}$, $\gamma(x) = g(B_0^\top x)$. We propose a method to estimate this subspace and establish its consistency. We illustrate the advantages of this dimension reduction procedure for estimating the extreme value index through simulations.

Keywords. Tail-index, dimension reduction, heavy-tailed distributions.

1 Introduction

On considère un couple aléatoire (X, Y) où Y est une variable aléatoire positive et X une variable aléatoire de support $\mathcal{X} \subset \mathbb{R}^p$. Nous nous plaçons dans le cas où pour tout $x \in \mathcal{X}$, la loi conditionnelle de Y sachant $X = x$ est à queue lourde d'indice des valeurs extrêmes conditionnel $\gamma(x) > 0$. Cet indice contrôle le comportement de la queue de la loi conditionnelle. Son estimation est essentielle dans de nombreux domaines comme en finance (voir, par exemple, Rockafellar et Uryasev [8]) ou en assurance (voir par exemple Brazauskas *et al.* [3] et Read et Vogel [7]). Plusieurs estimateurs de l'indice des valeurs extrêmes conditionnel sont disponibles dans la littérature citons par exemple Daouia, et al. [4], Gardes et Stupfler [5] ou encore Goegebeur et al. [6]. Cependant, il est bien connu que, pour un niveau de précision donné de l'un de ces estimateurs, le nombre n d'observations croît de manière exponentielle avec la dimension p . Ce phénomène est souvent appelé le fléau de la dimension (voir Bellman [2]).

Le point de départ de ce travail est de supposer l'existence d'un sous-espace \mathcal{S} de base $B \in \mathbb{R}^{p \times q}$ tel que, pour tout $x \in \mathcal{X}$, $\gamma(x) = g(B^\top x)$, où $g(\cdot)$ est une fonction positive inconnue. Ce sous-espace est appelé sous-espace TIDR (pour Tail-index dimension reduction en anglais). Par conséquent, si la matrice B est connue (ou peut au moins être estimée), l'estimation de $\gamma(X)$ peut être effectuée en remplaçant X par la covariable $B^\top X$ de dimension réduite $q \leq p$.

La principale contribution de ce travail est l'introduction d'un nouveau sous-espace DR, appelé \mathcal{S} , qui est utile lorsque nous cherchons à estimer l'indice des valeurs extrêmes conditionnel. Nous proposons ensuite une procédure d'estimation de ce sous-espace et nous démontrons sa consistance.

1.1 Définition du sous-espace TIDR

Dans la suite, nous considérons le modèle suivant pour la loi du couple (X, Y) .

- (M) Le support \mathcal{X} de X est supposé compact avec un intérieur non vide et l'extrémité gauche de la distribution de Y est strictement positive. De plus, pour tout $x \in \mathcal{X}$, la distribution conditionnelle de Y sachant $X = x$ est à queue lourde avec un indice de queue $\gamma(x) > 0$.

Rappelons que la loi de Y sachant $X = x$ est à queue lourde si

$$S(y, x) = y^{-1/\gamma(x)} \mathcal{L}(y, x),$$

où $\mathcal{L}(\cdot, x)$ est une fonction à variations lentes c'est-à-dire telle que pour tout $t > 0$ et $x \in \mathcal{X}$,

$$\lim_{y \rightarrow \infty} \frac{\mathcal{L}(ty, x)}{\mathcal{L}(y, x)} = 1.$$

Un premier résultat important en vue de la définition du sous-espace TIDR est le suivant. Pour une matrice donnée $B \in \mathbb{R}^{p \times q}$ avec $1 \leq q < p$, sous le modèle (M) et sous certaines hypothèses de régularité que nous ne détaillons

pas ici, la loi conditionnelle de Y sachant $B^\top X = B^\top x$ est également à queue lourde d'indice des valeurs extrêmes

$$\xi_B(B^\top x) := \max_{z: B^\top z = B^\top x} \gamma(z).$$

Une explication intuitive de ce résultat est la suivante. Notons

$$S_B(y, B^\top x) := \mathbb{P}(Y > y \mid B^\top X = B^\top x),$$

la fonction de survie de Y sachant $B^\top X = B^\top x$. La fonction $S_B(y, \cdot)$ peut être vue comme un mélange de la fonction $S(y, \cdot)$ en ce sens que $S_B(y, B^\top X) = \mathbb{E}[S(y, X) \mid B^\top X]$. Par conséquent, pour tout $x \in \mathcal{X}$, il est naturel de penser que la vitesse de décroissance de $S_B(\cdot, x)$ est donnée par le plus grand indice des valeurs extrêmes impliqué dans le mélange.

Un sous-espace TIDR est alors défini de la façon suivante.

Définition 1. *Un sous-espace vectoriel \mathcal{S} de dimension $q \in \{1, \dots, p\}$ avec $B \in \mathbb{R}^{p \times q}$ est un sous-espace TIDR si $\xi_B(B^\top X) = \gamma(X)$ presque sûrement.*

Lorsque B est la matrice identité de dimension p , il est facile de voir que $\xi_B(B^\top x) = \gamma(x)$ pour tout $x \in \mathcal{X}$ et donc que $\mathcal{S} = \mathbb{R}^p$ est toujours un sous-espace TIDR. Bien entendu, nous cherchons à trouver le plus petit sous-espace.

Définition 2. *Un sous-espace linéaire \mathcal{S}_0 est le sous-espace central tail index (CTI) si \mathcal{S}_0 est un sous-espace TIDR tel que $\mathcal{S}_0 \subset \mathcal{S}$ pour tous les sous-espaces TIDR \mathcal{S} .*

La base $B \in \mathbb{R}^{p \times q}$ telle que $\mathcal{S} = \text{span}(B)$ n'est bien sûr pas unique. Il est plus commode, en particulier pour les besoins de l'estimation, de travailler avec la base canonique de \mathcal{S} . L'ensemble des bases canoniques des sous-espaces de dimension q est noté \mathcal{B}_q .

Dans ce qui suit, nous désignons par $B_0 \in \mathcal{B}_q$ la base canonique du sous-espace CTI \mathcal{S}_0 . La distribution conditionnelle de Y sachant $B_0^\top X$ (qui n'est pas nécessairement égale à celle de Y étant donné X) est à queue lourde d'indice $\xi_{B_0}(B_0^\top X) = \gamma(X)$. Par conséquent, en supposant que B_0 est connu, l'indice des valeurs extrêmes conditionnel peut être estimé en utilisant un échantillon du couple aléatoire $(B_0^\top X, Y) \in \mathbb{R}^q \times \mathbb{R}$ au lieu de $(X, Y) \in \mathbb{R}^p \times \mathbb{R}$. Lorsque $q \ll p$, cela nous permet de construire un estimateur plus efficace de $\gamma(\cdot)$.

1.2 Estimation du CTI

Dans cette partie, nous supposons qu'il existe un sous-espace CTI de dimension connue q . Pour estimer la base canonique $B_0 \in \mathcal{B}_q$ du sous-espace CTI, nous commençons par remarquer que pour tout $x \in \mathcal{X}$ et $B \in \mathcal{B}_q$, on a

$$\gamma(x) = \xi_{B_0}(B_0^\top x) \leq \xi_B(B^\top x),$$

puisque $\gamma(x)$ appartient au support de la distribution conditionnelle de $\gamma(X)$ donnée par $B^\top X = B^\top x$. Ainsi, pour tout $B \in \mathcal{B}_q$ avec $B \neq B_0$, $\mathbb{P}(\xi_{B_0}(B_0^\top X) < \xi_B(B^\top X)) > 0$, ce qui conduit à

$$\arg \min_{B \in \mathcal{B}_q} \mathbb{E}[\xi_B(B^\top X)] = \{B_0\}.$$

En particulier, soit \mathcal{X}_0 un sous-ensemble compact à l'intérieur de \mathcal{X} satisfaisant

(C) pour tout $B \in \mathcal{B}_q$, $B \neq B_0$,

$$\mathbb{P}[\{\xi_{B_0}(B_0^\top X) < \xi_B(B^\top X)\} \cap \{X \in \mathcal{X}_0\}] > 0.$$

Nous avons alors

$$\arg \min_{B \in \mathcal{B}_q} \mathbb{E}[\xi_B(B^\top X) \mathbb{I}_{\mathcal{X}_0}(X)] =: \arg \min_{B \in \mathcal{B}_q} \Psi(B, \mathcal{X}_0) = \{B_0\}.$$

Remarquez que la condition (C) implique que le sous-espace CTI \mathcal{S}_0 existe. L'indicatrice $\mathbb{I}_{\mathcal{X}_0}(X)$ est introduite ici pour garantir que la densité de X sur \mathcal{X}_0 soit suffisamment éloignée de zéro.

Étant donné un échantillon $(X_1, Y_1), \dots, (X_n, Y_n)$ de copies indépendantes du couple aléatoire (X, Y) , la première étape de la procédure d'estimation de B_0 est l'estimation de la fonction $B \mapsto \Psi(B, \mathcal{X}_0)$. La version empirique de $\Psi(B, \mathcal{X}_0)$ est donnée par

$$\tilde{\Psi}_n(B, \mathcal{X}_0) := \frac{1}{n} \sum_{i=1}^n \xi_B(B^\top X_i) \mathbb{I}_{\mathcal{X}_0}(X_i).$$

Nous devons ensuite trouver un estimateur approprié de $\xi_B(B^\top x)$. Pour motiver la définition de notre estimateur, commençons par la situation irréaliste où nous disposons d'un échantillon $\{\check{Z}_i(B, x), i = 1, \dots, m\}$ de $m \in \mathbb{N} \setminus \{0\}$ variables aléatoires indépendantes ayant pour loi commune la loi conditionnelle de Y sachant $B^\top X = B^\top x$. Les statistiques d'ordre associées sont notées $\check{Z}_{(1)}(B, x) < \dots < \check{Z}_{(n)}(B, x)$. L'estimateur le plus connu de l'indice des valeurs extrêmes est l'estimateur de Hill, qui est donné par la formule suivante

$$\check{\xi}_B^{(H)}(B^\top x) := \frac{1}{\lfloor \alpha m \rfloor} \sum_{i=1}^{\lfloor \alpha m \rfloor} \ln \frac{\check{Z}_{(m-i+1)}(B, x)}{\check{Z}_{(m-\lfloor \alpha m \rfloor)}(B, x)},$$

pour tout $\alpha \in]1/m, 1[$. Bien entendu, il ne s'agit pas d'un estimateur puisque les variables aléatoires $\{\check{Z}_i(B, x), i = 1, \dots, m\}$ ne sont pas observées. Nous proposons donc de les remplacer dans l'expression de $\check{\xi}_B^{(H)}(B^\top x)$ par un ensemble de variables aléatoires observées. Plus précisément, introduisons l'ensemble

$$\mathcal{T}(B, x, h) := \{z \in \mathcal{X} \mid \|B^\top z - B^\top x\| \leq h\},$$

où $h = h_n > 0$ et

$$M^* = M(B, x, h) := \sum_{i=1}^n \mathbb{I}_{\mathcal{T}(B, x, h)}(X_i),$$

le nombre aléatoire de covariables dans l'ensemble $\mathcal{T}(B, x, h)$. Nous notons par $\{W_i^* = W_i(B, x, h), i = 1, \dots, M^*\}$ l'ensemble des covariables qui appartiennent à $\mathcal{T}(B, x, h)$. Les variables Y_i associées sont notées $\{Z_i^* = Z_i(B, x, h), i = 1, \dots, M^*\}$. Pour alléger les notations, nous mettons une étoile (\star) pour rappeler la dépendance en B, x et h . Intuitivement, lorsque h est suffisamment proche de zéro, la variable aléatoire Z_i^* est approximativement distribuée comme $\tilde{Z}_i(B, x)$. Pour une suite $(\alpha_n) = (\alpha) \in]0, 1[$, cela nous amène à introduire l'estimateur de l'indice des valeurs extrêmes conditionnel défini ci-dessous.

Définition 3. *Sous le modèle (M), pour tout $(B, x) \in \mathcal{B}_q \times \mathcal{X}$, l'estimateur local de Hill de l'indice de queue $\xi_B(B^\top x)$ est*

$$\tilde{\xi}_B^{(H)}(B^\top x) = \tilde{\xi}_B^{(H)}(B^\top x, \alpha, h) := \frac{1}{\lfloor \alpha M^* \rfloor} \sum_{i=1}^{\lfloor \alpha M^* \rfloor} \ln \frac{Z_{(M^* - i + 1)}^*}{Z_{(M^* - \lfloor \alpha M^* \rfloor)}^*},$$

si $\alpha M^* > 1$ et $\tilde{\xi}_B^{(H)}(B^\top x, \alpha, h) = 0$ sinon.

Nous proposons ainsi d'estimer $\Psi(B, \mathcal{X}_0)$ pour tout $B \in \mathcal{B}_q$ par l'estimateur plug-in

$$\widehat{\Psi}_n^{(H)}(B, \mathcal{X}_0) := \frac{1}{n} \sum_{i=1}^n \tilde{\xi}_B^{(H)}(B^\top X_i) \mathbb{I}_{\mathcal{X}_0}(X_i). \quad (1)$$

La définition de l'estimateur de B_0 est donnée ci-dessous.

Définition 4. *Sous le modèle (M), pour un échantillon $(X_1, Y_1), \dots, (X_n, Y_n)$ de copies indépendantes du couple aléatoire (X, Y) , l'estimateur \widehat{B}_n de la base canonique du sous-espace CTI minimise la fonction $B \mapsto \widehat{\Psi}_n(B, \mathcal{X}_0)$.*

1.3 Consistance de l'estimateur \widehat{B}_n .

Sous des hypothèses supplémentaires techniques que nous ne mentionnons pas ici, nous montrons que

$$\left\| \widehat{B}_n - B_0 \right\| \xrightarrow{\mathbb{P}} 0$$

où $\|\cdot\|$ est une norme quelconque dans $\mathbb{R}^{p \times q}$.

La démonstration de ce résultat passe essentiellement par celle de la consistance uniforme de l'estimateur $\widehat{\Psi}_n^{(H)}(B, \mathcal{X}_0)$ de $\Psi(B, \mathcal{X}_0)$ i.e.,

$$\lim_{n \rightarrow \infty} \sup_{x \in \mathcal{X}_0} \sup_{B \in \mathcal{B}_q} \left| \widehat{\Psi}_n^{(H)}(B, \mathcal{X}_0) - \Psi(B, \mathcal{X}_0) \right| = 0.$$

1.4 Simulations

On se place dans le cadre suivant. La variable aléatoire X suit une loi uniforme sur l'espace $\mathcal{X} := [0, 1]^p$ que l'on munit de la norme uniforme $\|\cdot\|_\infty$. Nous fixons

la dimension de la covariable X à $p = 8$. La variable aléatoire Y est donnée par $Y := Q(U, x)$ avec U suivant une loi uniforme standard et

$$Q(u|x) = u^{-\gamma(x)} [1 + \exp\{B_1^\top x + u^{-1}\}]^{-1},$$

où $B_1 = (0, 0, 5, 5, 0, \dots, 0)^\top$.

Plusieurs modèles pour la fonction $\gamma(\cdot)$ sont considérés. Pour les deux premiers modèles, le sous-espace CTI est de dimension $q = 1$ admettant $B_0 = (2, 1, 0, \dots, 0)^\top / 3 \in \mathbb{R}^8$ pour base canonique.

- **Modèle 1** - Pour tout $x \in \mathcal{X}$,

$$\gamma(x) := \frac{1}{10} + \frac{9[\exp(2B_0^\top x) - 1]}{10[\exp(2) - 1]}.$$

- **Modèle 2** - Pour tout $x \in \mathcal{X}$,

$$\gamma(x) := \frac{1}{10} + \frac{9|\cos(3B_0^\top x)|}{10}.$$

Pour le modèle suivant le sous-espace CTI est de dimension $q = 2$ avec pour base canonique $B_0 = (e_1, e_2) \in \mathbb{R}^{8 \times 2}$ où $e_1 = (1, 0, \dots, 0)^\top$ et $e_2 = (0, 1, 0, \dots, 0)^\top$.

- **Modèle 3** - Pour tout $x \in \mathcal{X}$,

$$\gamma(x) := \frac{1}{10} + \frac{9}{5} [(e_1^\top x - 0, 5)^2 + (e_2^\top x - 0, 5)^2].$$

Nous générons, pour chacun des modèles, $N = 100$ échantillons de taille $n = 2000$. Afin d'évaluer la performance d'un estimateur \tilde{B}_n de la base B_0 du sous-espace CTI, on calcule la distance

$$\|\tilde{B}_n - B_0\|_F$$

où $\|\cdot\|_F$ est norme de Frobenius. Nous comparons l'estimateur TIDR proposé dans ce travail avec les estimateurs TIREX1 et TIREX2 proposés par Aghbalou et al. [1]. Notons toutefois que notre méthode n'est applicable que si nous avons pour tout $x \in \mathcal{X}$, $\gamma(x) > 0$ contrairement aux méthodes TIREX qui sont applicables quelque soit le signe de $\gamma(x)$. A noter également que la méthode TIREX1 n'est justifiée théoriquement que dans le cas $q = 1$.

Nous supposons pour le moment que la dimension q du sous-espace CTI est connue. Une procédure de choix de q sera proposée par la suite. Notre méthode nécessite de choisir deux paramètres h et α . Après plusieurs tests sur de nombreuses simulations, les choix $h = (n^{-1/3}/2^{q-1})^{1/q}$ et $\alpha = n^{-3/10}$ semblent donner de bons résultats dans la majorité des situations. Les méthodes TIREX dépendent elles d'un paramètre k (nombre maximal des plus grandes observations utilisées pour l'estimation). La procédure choix par validation croisée

proposée par [1] est uniquement adaptée dans un contexte d'apprentissage supervisé. Pour donner à notre concurrent l'avantage maximale, nous prenons pour chaque simulation la valeur de k donnant les meilleurs résultats parmi $\mathcal{K} := \{50, 60, \dots, 390, 400\}$.

Comparaison entre modèles – Notons \hat{B}_n^{TIDR} , $\hat{B}_n^{\text{TIREX1}}$ et $\hat{B}_n^{\text{TIREX2}}$ les estimateurs de la base B_0 obtenus respectivement par les méthodes TIDR, TIREX1 et TIREX2. Les boîtes-à-moustaches des erreurs sont donnés dans les Figures 1 et 2. Dans tous les cas, la méthode TIDR donne les meilleurs résultats.

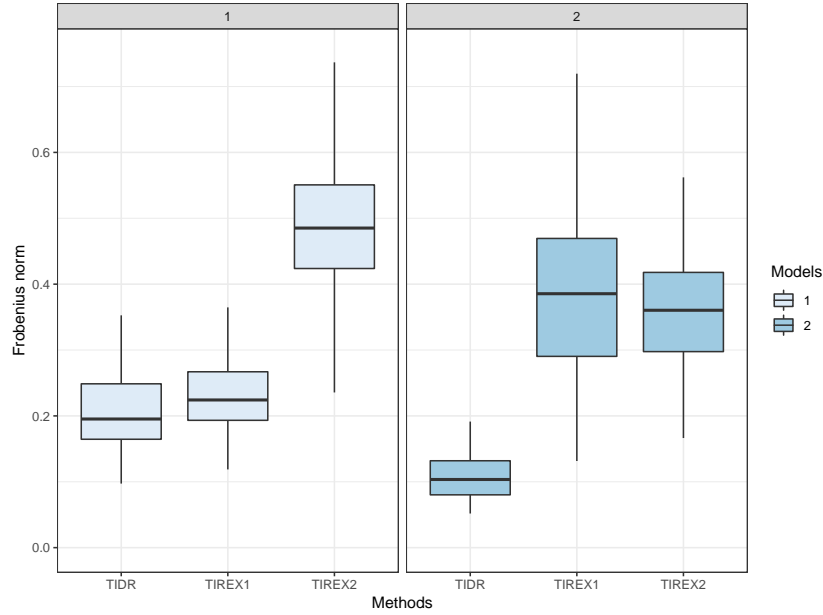


Figure 1: Comparaison des performances des méthodes TIDR et TIREX (1 et 2) sur un échantillon taille $n = 2000$, pour les modèles 1 et 2 avec $p = 8$ et $q = 1$.

Estimation de l'indice des valeurs extrêmes – Nous pouvons utiliser les estimateurs $\check{B}_n \in \{\hat{B}_n^{\text{TIDR}}, \hat{B}_n^{\text{TIREX1}}, \hat{B}_n^{\text{TIREX2}}\}$ du sous-espace CTI pour estimer l'indice des valeurs extrêmes conditionnel $\gamma(x)$ pour un point $x \in \mathcal{X}$. Nous choisissons pour ce faire d'utiliser l'estimateur local de Hill $\hat{\xi}_B^{(H)}(B^\top x)$ introduit dans le paragraphe précédent.

Pour évaluer la qualité de l'estimation, nous calculons pour chaque réplification l'erreur moyenne quadratique

$$E_\gamma(\hat{B}_n) := \frac{1}{1000} \sum_{\ell=1}^{1000} \left[\hat{\xi}_{\hat{B}_n}^{(H)}(\hat{B}_n^\top x_\ell) - \gamma(x_\ell) \right]^2,$$

où les 1000 points x_ℓ sont tirés uniformément sur $[0, 1]^p$.

Les erreurs moyennes sur les $N = 100$ réplifications sont rassemblées dans la Table 1. Comme on pouvait s’y attendre à la vue des résultats précédents, l’estimation de l’indice des valeurs extrêmes basé sur la réduction de dimension TIDR est meilleures pour les trois modèles considérés.

Model	$E_\gamma(\hat{B}_n^{\text{TIDR}})$	$E_\gamma(\hat{B}_n^{\text{TIREX1}})$	$E_\gamma(\hat{B}_n^{\text{TIREX2}})$
1	0.022 (0.018)	0.023 (0.014)	0.046 (0.028)
2	0.042 (0.025)	0.106 (0.043)	0.095 (0.036)
3	0.048 (0.016)	0.194 (0.076)	0.085 (0.022)

Table 1: Comparaison des erreurs moyennes de l’estimation du tail-index (entre parenthèses l’écart type).

Dans la Table 2, nous avons également évalué la performance de l’estimateur local de Hill dans 2 cas : 1) on utilise la vraie base B_0 du sous-espace TIDR et 2) on n’effectue pas de réduction de dimension (i.e., on prend la matrice identité sur \mathbb{R}^p). Comme cela était prévisible l’erreur $E_\gamma(Id)$ est importante justifiant ainsi la nécessité d’appliquer une procédure de réduction de dimension dans le cadre des valeurs extrêmes. Pour tous les modèles, les performances de notre méthode sont proches du cas où B_0 est connue.

Model	$E_\gamma(B_0)$	$E_\gamma(Id)$
1	0.014 (0.009)	0.616 (0.102)
2	0.035 (0.017)	0.925 (0.135)
3	0.047 (0.011)	0.733 (0.109)

Table 2: Comparaison des erreurs moyennes de l’estimation du tail-index (entre parenthèses l’écart type).

Choix de la dimension du TIDR – Nous avons jusqu’à présent supposer que la dimension l’espace CTI q était connue. Évidemment, ce n’est pas le cas en pratique. Nous proposons ici une procédure de choix basée sur le résultat suivant.

Lemme 1. *Soit $d \in \mathbb{N} \setminus \{0\}$. Nous avons*

Model	Choice of dimension			$E_\gamma(\widehat{B}_n^{\text{TIDR}})$		
	$\widehat{q}_n = 1$	$\widehat{q}_n = 2$	$\widehat{q}_n = 3$	$q = 1$	$q = 2$	$q = 3$
1	100 %	0 %	0 %	0.024	0.041	0.064
2	100 %	0 %	0 %	0.042	0.110	0.168
3	3 %	97 %	0 %	0.074	0.048	0.095

Table 3: Résultat de la procédure de choix de la dimension et comparaison des erreurs moyennes selon le choix de q .

$$\min_{B \in \mathcal{B}_{d+1}} \Psi(B) \leq \min_{B \in \mathcal{B}_d} \Psi(B).$$

Posons pour tout d

$$c(d) := \min_{B \in \mathcal{B}_d} \Psi(B).$$

D'après le Lemme 1, les valeurs de $c(d)$ pour $d \in \{1, \dots, q_0\}$ diminue quand d augmente. Nous proposons donc de sélectionner la dimension en comparant successivement $\widehat{c}_n(d)$ à $\widehat{c}_n(d+1)$, où pour tout $d \in \{1, \dots, p\}$,

$$\widehat{c}_n(d) := \widehat{\Psi}_n(\widehat{B}_n^{(d)}),$$

avec $\widehat{B}_n^{(d)}$ l'estimateur donné par la méthode TIDR pour la dimension d . Plus précisément, nous choisissons

$$\widehat{q}_n = \min\{d \mid \widehat{c}_n(d) < \widehat{c}_n(d+1)\}.$$

Comme le montre la Table 3, cette procédure sélectionne la bonne dimension dans quasiment tous les cas. La qualité de l'estimation de $\gamma(x)$ est aussi comparée selon la dimension sélectionnée afin d'évaluer l'impact de ce choix.

Références bibliographiques

- [1] Aghbalou, A., F. Portier, A. Sabourin, and C. Zhou 2024. Tail inverse regression: dimension reduction for prediction of extremes. *Bernoulli* 30(1).
- [2] Bellman, R. 1961. *Adaptive Control Process: A Guided Tour.* : Princeton University Press.
- [3] Brazauskas, V., B. Jones, M. Puri, and R. Zitikis 2008. Estimating conditional tail expectation with actuarial applications in view. *Journal of Statistical Planning and Inference* 138(11):3590–3604.

- [4] Daouia, A., L. Gardes, and S. Girard 2013. On kernel smoothing for extremal quantile regression. *Bernoulli* 19(5B).
- [5] Gardes, L. and G. Stupfler 2014. Estimation of the conditional tail index using a smoothed local hill estimator. *Extremes* 17:45–75.
- [6] Goegebeur, Y., A. Guillo, and G. Stupfler 2015. Uniform asymptotic properties of a nonparametric regression estimator of conditional tails. *Annales de l'IHP Probabilités et statistiques* 51(3):1190–1213.
- [7] Read, L. and R. Vogel 2015. Reliability, return periods, and risk under nonstationarity. *Water Resources Research* 51(8):6381–6398.
- [8] Rockafellar, R. and S. Uryasev 2002. Conditional value-at-risk for general loss distributions. *Journal of banking & finance* 26(7):1443–1471.

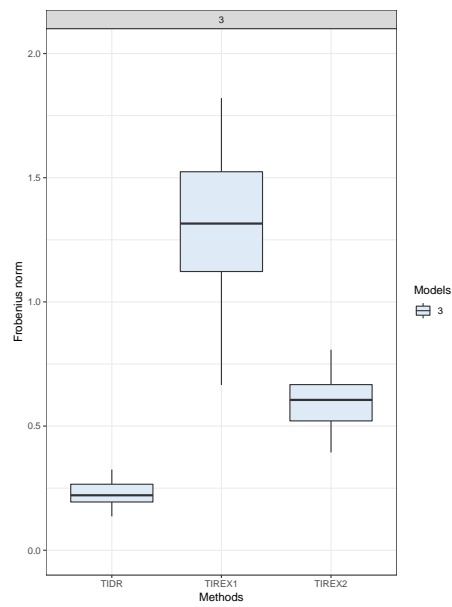


Figure 2: Comparaison des performances des méthodes TIDR et TIREX (1 et 2) avec un échantillon taille $n = 2000$, pour le modèle 3 avec $p = 8$ et $q = 2$.