

Prédiction spatiale sur des champs aléatoires elliptiques

Antoine Usseglio-Carleve¹

Travail joint avec : Véronique Maume-Deschamps¹ et Didier
Rullière²

¹Institut Camille Jordan
Université Lyon 1

²Laboratoire des Sciences Actuarielle et Financière
Université Lyon 1

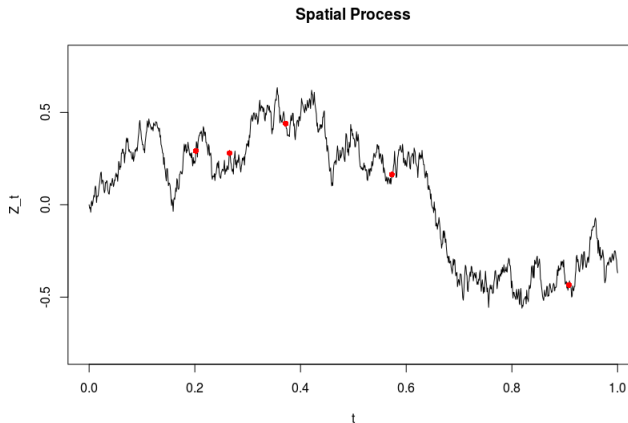
Porquerolles, 5 avril 2017

- 1 Introduction
- 2 Champs elliptiques
- 3 Prédiction de quantile
- 4 Prédiction d'expectile
- 5 Application numérique

- 1 Introduction
- 2 Champs elliptiques
- 3 Prédiction de quantile
- 4 Prédiction d'expectile
- 5 Application numérique

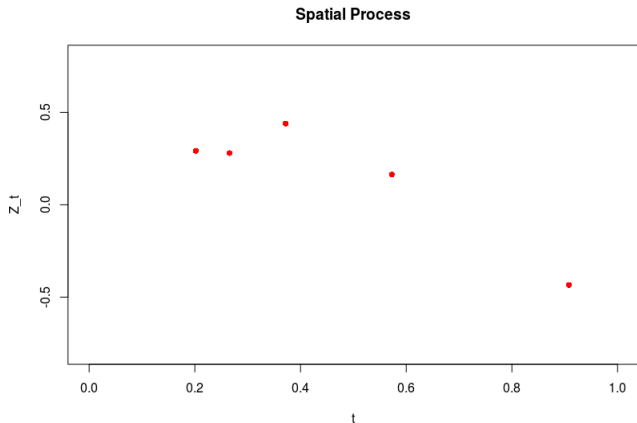
Introduction

Soit $(X_t)_{t \in \mathcal{T}}$ un champ aléatoire observé en n points t_1, \dots, t_n .



Introduction

Soit $(X_t)_{t \in \mathcal{T}}$ un champ aléatoire observé en n points t_1, \dots, t_n .



Introduction

Le krigeage (Krige, 1952 [5], Matheron, 1963 [6]) prédit la valeur du champ au point $t \in T$ de la manière suivante :

$$\hat{\mathbb{E}}[X_t | X] = \beta^{*T} X$$

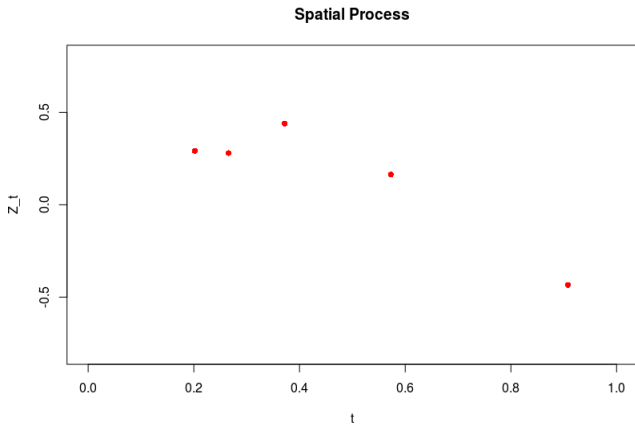
où $X = (X_{t_1}, \dots, X_{t_n})$, et β^* est solution de :

$$\arg \min_{\beta \in \mathbb{R}^n} \mathbb{E}[(X_t - \beta^T X)^2]$$

i.e $\beta^* = \mathbb{E}[X^T X]^{-1} \mathbb{E}[X X_t]$.

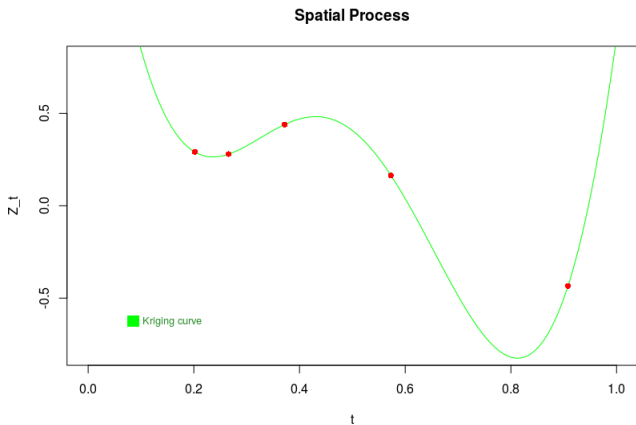
Introduction

Retour à notre exemple :



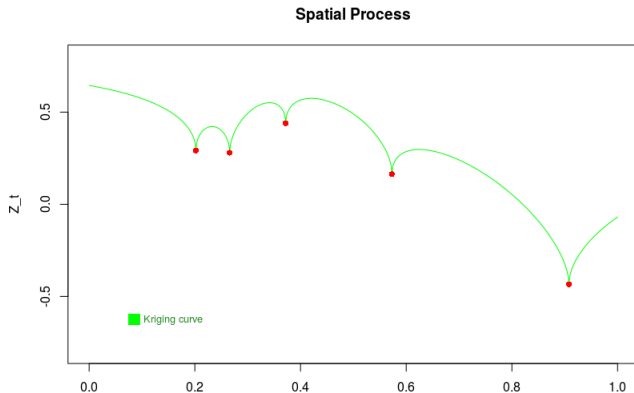
Introduction

Retour à notre exemple :



Introduction

Question : Toujours dans une optique de régression, peut-on prédire la valeur de X_t avec d'autres critères que l'espérance ?



- 1 Introduction
- 2 Champs elliptiques**
- 3 Prédiction de quantile
- 4 Prédiction d'expectile
- 5 Application numérique

Distributions elliptiques

Definition

Soit X un vecteur aléatoire de dimension d . X est elliptique ssi il existe un unique $\mu \in \mathbb{R}^d$, une matrice semi-définie positive $\Sigma \in \mathbb{R}^{d \times d}$, et une variable aléatoire positive R tels que :

$$X \stackrel{d}{=} \mu + R\Lambda U^{(d)}$$

avec $\Lambda\Lambda^T = \Sigma$, $U^{(d)}$ est une loi uniforme sur la sphère de dimension d indépendante de R . De plus, X est dite elliptique consistante si $R \stackrel{d}{=} \frac{\chi_d}{\epsilon}$, où ϵ est une variable aléatoire positive indépendante de χ_d et d .

Distributions elliptiques

Theorem (Densité elliptique)

Si $X \sim \mathcal{E}_d(\mu, \Sigma, R)$, alors X admet comme densité :

$$f_X(x) = \frac{c_d}{|\det(\Sigma)|^{\frac{1}{2}}} g_d((x - \mu)\Sigma^{-1}(x - \mu))$$

où $g_d(t) = \frac{\Gamma(\frac{d}{2})}{2\pi^{\frac{d}{2}}} \sqrt{t}^{-(d-1)} f_R(\sqrt{t})$, et $f_R(t)$ est la densité de R .

- g_d est appelé générateur de X .

Exemples

Distribution	Coefficient c_d	Générateur $g_d(t)$	ϵ
Gaussien	$\frac{1}{(2\pi)^{\frac{d}{2}}}$	$\exp(-\frac{t}{2})$	1
Student, $\nu > 0$	$\frac{\Gamma(\frac{d+\nu}{2})}{\Gamma(\frac{\nu}{2})} \frac{1}{(\nu\pi)^{\frac{d}{2}}}$	$(1 + \frac{t}{\nu})^{-\frac{d+\nu}{2}}$	$\frac{\chi_\nu}{\nu}$
Mélange Gaussien	$\frac{1}{(2\pi)^{\frac{d}{2}}}$	$\sum_{k=1}^n \pi_k \theta_k^d \exp\left(-\frac{\theta_k^2}{2} t\right)$	$\sum_{k=1}^n \pi_k \theta_k$
Slash, $a > 0$	$\frac{2^{\frac{d}{2}-1} a \Gamma(\frac{d+a}{2})}{\pi^{\frac{d}{2}}}$	$\frac{\chi_{d+a}^2(t)}{t^{\frac{d+a}{2}}}$	$\beta(a, 1)$

Propriétés

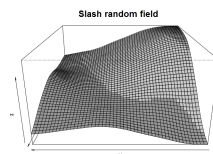
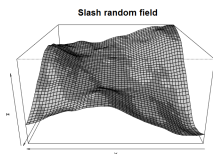
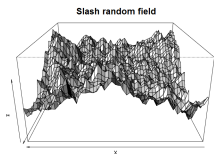
Proposition

- Si X est un vecteur R -elliptique consistant de dimension d , alors tous ses sous-vecteurs de dimension $k \leq d$ sont R -elliptiques de dimension k .
- Si X est un vecteur R -elliptique consistant de dimension d , alors toute combinaison linéaire de ses composantes est R -elliptique de dimension 1.

Champs elliptiques

Definition

Un champ aléatoire $(X_t)_{t \in T}$ est dit R -elliptique si $\forall n \in \mathbb{N}^*$, $\forall t_1, \dots, t_n \in T$, $(X_{t_1}, \dots, X_{t_n})$ est R -elliptique consistant de dimension n .



- 1 Introduction
- 2 Champs elliptiques
- 3 Prédiction de quantile**
- 4 Prédiction d'expectile
- 5 Application numérique

Quantiles conditionnels

Quantile conditionnel théorique :

$$q_\alpha(X_2|X_1 = x_1) = \mu_{2|1} + \sigma_{2|1}\Phi_{R^*}^{-1}(\alpha)$$

avec

$$\begin{cases} \mu_{2|1} = \mu_2 + \Sigma_{21}\Sigma_{11}^{-1}(x_1 - \mu_1) \\ \sigma_{2|1}^2 = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} \end{cases}$$

Régression quantile [4]

$$\hat{q}_\alpha(X_2|X_1 = x_1) = \beta^{*T} x_1 + \beta_0^*$$

où β^* et β_0^* sont solutions du problème de minimisation :

$$(\beta^*, \beta_0^*) = \arg \min_{\beta \in \mathbb{R}^N, \beta_0 \in \mathbb{R}} \mathbb{E}[\mathcal{S}_\alpha(X_2 - \beta^T X_1 - \beta_0)]$$

avec la fonction score \mathcal{S}_α :

$$\mathcal{S}_\alpha(x) = (\alpha - 1)x\mathbb{1}_{\{x < 0\}} + \alpha x\mathbb{1}_{\{x > 0\}}$$

Régression quantile

Theorem (Prédicteur de régression quantile)

Le vecteur de régression quantile (β^*, β_0^*) de $X_2 | (X_1 = x_1)$ est donné par :

$$\begin{cases} \beta^* = & \Sigma_{11}^{-1} \Sigma_{12} \\ \beta_0^* = & \mu_2 - \Sigma_{21} \Sigma_{11}^{-1} \mu_1 + \sigma_{2|1} \Phi_R^{-1}(\alpha) \end{cases}$$

Le prédicteur de régression quantile de niveau $\alpha \in [0, 1]$ est donné par :

$$\hat{q}_\alpha(X_2 | X_1 = x_1) = \mu_{2|1} + \sigma_{2|1} \Phi_R^{-1}(\alpha)$$

De plus,

$$\hat{q}_\alpha(X_2 | X_1) \sim \mathcal{E}_1(\mu_2 + \sigma_{2|1} \Phi_R^{-1}(\alpha), \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12}, R)$$

Quantiles extrêmes

- Prédicteur de régression assez éloigné des quantiles conditionnels théoriques, surtout pour des valeurs de α proches de 0 ou 1.
- On propose alors un prédicteur extrême.

Prédicteurs de quantiles extrêmes

Definition

Une fonction f est une φ -fonction si $f : [0, +\infty[\rightarrow [0, +\infty[$, $f(0) = 0$, f est continue, croissante sur $[0, +\infty[$, et $f \rightarrow +\infty$ quand $x \rightarrow +\infty$. De plus, on appelle K_c l'ensemble des φ -fonctions f vérifiant :

$$\lim_{(x,\lambda) \rightarrow (\infty,1)} \frac{f(\lambda x)}{f(x)} = 1. \quad (1)$$

Prédicteurs de quantiles extrêmes

Assumption

Il existe $0 < \ell < +\infty$ et $\gamma \in \mathbb{R}$ tels que :

$$\lim_{x \rightarrow +\infty} \frac{\bar{\Phi}_{R^*}(x)}{\bar{\Phi}_R(x^\gamma)} = \ell \quad (2)$$

où $\bar{\Phi} = 1 - \Phi$ est la fonction de survie associée à Φ . De plus, soient Ω et Ω_* tels que :

$$\begin{cases} \Omega(x) &= \frac{1}{\ell \bar{\Phi}_R(x^\gamma)} - \frac{2}{\ell} \\ \Omega_*(x) &= \frac{1}{\bar{\Phi}_{R^*}(x)} - 2 \end{cases} \quad (3)$$

Alors Ω^{-1} ou Ω_*^{-1} appartient à la classe K_C .

Exemples

Proposition

Les distributions Gaussienne, Student, Mélange Gaussien et Slash satisfont l'hypothèse, avec des coefficients γ et ℓ donnés dans le tableau ci-dessous.

Distribution	γ	ℓ
Gaussien	1	1
Student, $\nu > 0$	$\frac{N+\nu}{\nu}$	$\frac{\Gamma(\frac{\nu+N+1}{2})\Gamma(\frac{\nu}{2})}{\Gamma(\frac{\nu+N}{2})\Gamma(\frac{\nu+1}{2})} \left(1 + \frac{q_1}{\nu}\right)^{\frac{N+\nu}{2}} \frac{\nu^{\frac{N}{2}+1}}{\nu+N}$
Mélange Gaussien	1	$\frac{\min(\theta_1, \dots, \theta_n)^N \exp\left(-\frac{\min(\theta_1, \dots, \theta_n)^2}{2} q_1\right)}{\sum_{k=1}^n \pi_k \theta_k^N \exp\left(-\frac{\theta_k^2}{2} q_1\right)}$
Slash, $a > 0$	$\frac{N+a}{a}$	$\frac{\Gamma\left(\frac{N+1+a}{2}\right) q_1^{\frac{N+a}{2}}}{\Gamma\left(\frac{N+a}{2}\right) (N+a) \chi_{N+a}^2(q_1) 2^{\frac{a}{2}-1} \Gamma\left(\frac{1+a}{2}\right)}$

Prédicteurs de quantiles extrêmes

Definition (Prédicteurs de quantiles extrêmes)

$$\begin{cases} \hat{q}_{\alpha\uparrow}(X_2|X_1 = x_1) = \mu_{2|1} + \sigma_{2|1} \left[\Phi_R^{-1} \left(1 - \frac{1}{\frac{\ell}{1-\alpha} + 2(1-\ell)} \right) \right]^{\frac{1}{\gamma}} \\ \hat{q}_{\alpha\downarrow}(X_2|X_1 = x_1) = \mu_{2|1} - \sigma_{2|1} \left[\Phi_R^{-1} \left(1 - \frac{1}{\frac{\ell}{\alpha} + 2(1-\ell)} \right) \right]^{\frac{1}{\gamma}} \end{cases}$$

Theorem (Equivalences)

Sous notre hypothèse,

$$\begin{cases} \hat{q}_{\alpha\uparrow}(X_2|X_1 = x_1) \underset{\alpha \rightarrow 1}{\sim} q_{\alpha}(X_2|X_1 = x_1) \\ \hat{q}_{\alpha\downarrow}(X_2|X_1 = x_1) \underset{\alpha \rightarrow 0}{\sim} q_{\alpha}(X_2|X_1 = x_1) \end{cases}$$

- 1 Introduction
- 2 Champs elliptiques
- 3 Prédiction de quantile
- 4 Prédiction d'expectile**
- 5 Application numérique

Expectiles conditionnels

Expectile conditionnel théorique :

$$e_\alpha(X_2|X_1 = x_1) = \mu_{2|1} + \sigma_{2|1} \Psi_{R^*}^{-1} \left(\frac{\alpha}{2\alpha - 1} \right)$$

avec

$$\begin{cases} \mu_{2|1} = \mu_2 + \Sigma_{21} \Sigma_{11}^{-1} (x_1 - \mu_1) \\ \sigma_{2|1}^2 = \Sigma_{22} - \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12} \end{cases}$$

Régression expectile [9]

$$\hat{e}_\alpha(X_2|X_1 = x_1) = \beta^{*T} x_1 + \beta_0^*$$

où β^* et β_0^* sont solutions du problème de minimisation :

$$(\beta^*, \beta_0^*) = \arg \min_{\beta \in \mathbb{R}^N, \beta_0 \in \mathbb{R}} \mathbb{E}[\mathcal{S}_\alpha(X_2 - \beta^T X_1 - \beta_0)]$$

avec la fonction score \mathcal{S}_α :

$$\mathcal{S}_\alpha(x) = (1 - \alpha)x^2 \mathbf{1}_{\{x < 0\}} + \alpha x^2 \mathbf{1}_{\{x > 0\}}$$

Régression expectile

Theorem (Prédicteur de régression expectile)

Le vecteur de régression expectile (β^*, β_0^*) de $X_2 | (X_1 = x_1)$ est donné par :

$$\begin{cases} \beta^* = & \Sigma_{11}^{-1} \Sigma_{12} \\ \beta_0^* = & \mu_2 - \Sigma_{21} \Sigma_{11}^{-1} \mu_1 + \sigma_{2|1} \Psi_R^{-1} \left(\frac{\alpha}{2\alpha-1} \right) \end{cases} \quad (4)$$

Le prédicteur de régression expectile $\alpha \in [0, 1]$ est donné par :

$$\hat{e}_\alpha(X_2 | X_1 = x_1) = \mu_{2|1} + \sigma_{2|1} \Psi_R^{-1} \left(\frac{\alpha}{2\alpha-1} \right) \quad (5)$$

De plus,

$$\hat{e}_\alpha(X_2 | X_1) \sim \mathcal{E}_1 \left(\mu_2 + \sigma_{2|1} \Psi_R^{-1} \left(\frac{\alpha}{2\alpha-1} \right), \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12}, R \right) \quad (6)$$

Expectile extrêmes

Assumption

Il existe $0 < \ell < +\infty$ et $\gamma \in \mathbb{R}$ tels que :

$$\lim_{x \rightarrow +\infty} \frac{\bar{\Psi}_{R^*}(x)}{\bar{\Psi}_R(x^\gamma)} = \ell \quad (7)$$

où $\bar{\Psi} = 1 - \Psi$. De plus, soient Ω et Ω_* tels que :

$$\begin{cases} \Omega(x) &= -(\ell \bar{\Psi}_R(x^\gamma))^{-1} \\ \Omega_*(x) &= -(\bar{\Psi}_{R^*}(x))^{-1} \end{cases} \quad (8)$$

Alors Ω^{-1} ou Ω_*^{-1} appartient à la classe K_c .

Expectile extrêmes

Definition (Prédicteurs d'expectiles extrêmes)

$$\begin{cases} \hat{e}_{\alpha\uparrow}(X_2|X_1 = x_1) = \mu_{2|1} + \sigma_{2|1} \left[\Psi_R^{-1} \left(1 - \frac{\alpha-1}{(2\alpha-1)\ell} \right) \right]^{\frac{1}{\gamma}} \\ \hat{e}_{\alpha\downarrow}(X_2|X_1 = x_1) = \mu_{2|1} - \sigma_{2|1} \left[\Psi_R^{-1} \left(1 - \frac{\alpha}{(2\alpha-1)\ell} \right) \right]^{\frac{1}{\gamma}} \end{cases}$$

Theorem (Equivalences)

Sous notre hypothèse,

$$\begin{cases} \hat{e}_{\alpha\uparrow}(X_2|X_1 = x_1) \underset{\alpha \rightarrow 1}{\sim} e_{\alpha}(X_2|X_1 = x_1) \\ \hat{e}_{\alpha\downarrow}(X_2|X_1 = x_1) \underset{\alpha \rightarrow 0}{\sim} e_{\alpha}(X_2|X_1 = x_1) \end{cases}$$

- 1 Introduction
- 2 Champs elliptiques
- 3 Prédiction de quantile
- 4 Prédiction d'expectile
- 5 Application numérique

Application numérique

On définit les *RMSEs* suivants :

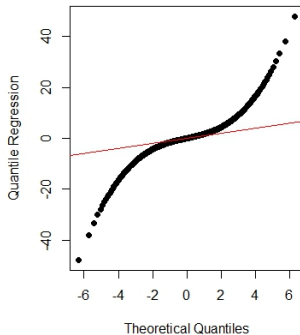
$$\left\{ \begin{array}{l} RMSE(\hat{q}_\alpha) = \sqrt{\frac{1}{n} \sum_{i=1}^n \left(q_\alpha(X_2^{(i)} | X_1 = x_1) - \hat{q}_\alpha(X_2^{(i)} | X_1 = x_1) \right)^2} \\ RMSE(\hat{\hat{q}}_\alpha) = \sqrt{\frac{1}{n} \sum_{i=1}^n \left(q_\alpha(X_2^{(i)} | X_1 = x_1) - \hat{\hat{q}}_\alpha(X_2^{(i)} | X_1 = x_1) \right)^2} \end{array} \right.$$

Application numérique

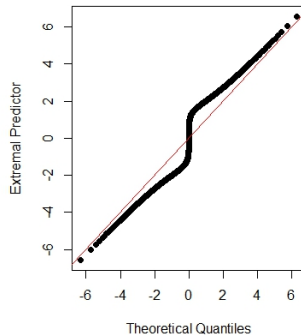
α	Student		Mélange Gaussien		Slash	
	$RMSE(\hat{q}_\alpha)$	$RMSE(\hat{\hat{q}}_\alpha)$	$RMSE(\hat{q}_\alpha)$	$RMSE(\hat{\hat{q}}_\alpha)$	$RMSE(\hat{q}_\alpha)$	$RMSE(\hat{\hat{q}}_\alpha)$
0.5	0	0	0	0	0	0
0.6	0.02717082	0.3716682	0.00146743	0.01868868	0.06691511	0.3335765
0.7	0.05902606	0.3758591	0.003233866	0.03562366	0.1511828	0.3008545
0.8	0.1044727	0.3583675	0.005926889	0.05054183	0.2968964	0.2494973
0.9	0.1952414	0.322185	0.01223556	0.06109096	0.7701059	0.1766719
0.95	0.3136118	0.2888097	0.0236387	0.0610452	1.880313	0.1234383
0.9995	2.880308	0.1478775	0.1628414	0.0003795359	250.1717	0.02022995
0.999995	16.54578	0.08062824	0.1093942	7.130243e-06	25178.53	0.00812057

Application numérique

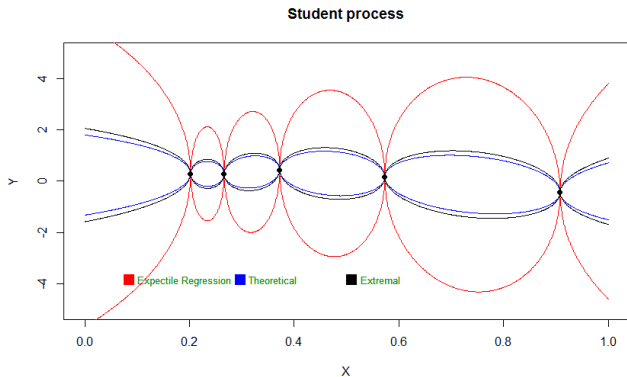
Student Quantile Regression









Student Extremal Predictor






Application numérique



References I

-  Cambanis, S., Huang, S., Simons, G. (1981). *On the theory of elliptically contoured distributions*, Journal of Multivariate Analysis, No. 11, pp. 368-385.
-  Djurčić, D., Torgašev, A. (2001). *Strong Asymptotic Equivalence and Inversion of Functions in the Class Kc* , Journal of Mathematical Analysis and Applications 255, 383–390.
-  Kano, Y. (1994). *Consistency Property of Elliptical Probability Density Functions*, Journal of Multivariate Analysis, Volume 51, Pages 139-147.
-  Koenker, R., Bassett, G. Jr. (1978). *Regression Quantiles*, Econometrica, Vol. 46, No. 1, pp. 33-50.
-  Krige, D. (1951). *A statistical approach to some basic mine valuation problems on the witwatersrand*, Journal of the Chemical, Metallurgical and Mining Society, Volume 52, Pages 119-139.
-  Matheron, G. (1963). *Traité de géostatistique appliquée*, Mémoires du Bureau de Recherches Géologiques et Minières, No. 24.

References II

-  Maume-Deschamps, V., Rullière, D., Usseglio-Carleve, A. (2016). *Spatial expectile predictions for elliptical random fields*, working paper on preprint.
-  Maume-Deschamps, V., Rullière, D., Usseglio-Carleve, A. (2016). *Spatial quantile predictions for elliptical random fields*, en révision à Journal of Multivariate Analysis.
-  Newey, W.K., Powell, J.L., (1987), *Asymmetric Least Squares Estimation and Testing*, Econometrica.

Merci de votre attention