Réduction de dimension en régression spatiale et prédiction spatiale

J.M. Loubes[1] et A.F. Yao[2]

[1]Université Toulouse 3, Institut de Mathématiques de Toulouse.

[2]Université Aix-Marseille 2, Centre d'Océanologie de Marseille.

Motivations

- Estimation de la fonction de régression $m(.)$:

$$Y = m(X) + \varepsilon,$$

$(Y, X) \in \mathbb{R} \times \mathbb{R}^d$ et ε erreur aléatoire.

$\Longrightarrow$ A partir d'observations spatialement dépendantes.

- Proposer un prédicteur spatial basé sur l'estimateur (de la régression) obtenu.

Motivations

- Estimation de la fonction de régression $m(.)$:

$$Y = m(X) + \varepsilon,$$

$(Y, X) \in \mathbb{R} \times \mathbb{R}^d$ et ε erreur aléatoire.

$\Longrightarrow$ A partir d'observations spatialement dépendantes.

- Proposer un prédicteur spatial basé sur l'estimateur (de la régression) obtenu.

Motivations

.... plus précisément...

- Etendre les méthodes réduction de dimension au cas où les observations sont spatialement dépendantes.

- Proposer un prédicteur spatial basé sur le modèle de régression obtenu.

Réduction de dimension

... en Régression (i.i.d.)...

$$Y = m(X) + \varepsilon$$

- $\Longrightarrow$ la qualité de l'estimation non-paramétrique de la fonction $m(\cdot)$ est pénalisée par la dimension du régresseur

Modèle de réduction de dimension:

- $\Longrightarrow$ plusieurs solutions: la méthode de "Projection pursuit regression" (Friedman and Stueltzle, 1981), GAM,"General Additive Model" (Hastie et Tibshirani, 1986), ACE: Adevrage Component Estimation(Brieman et Friedman, 1985),...
- $\Longrightarrow$ SIR ("Sliced Inverse Regression"), Li(1991), OPG("Outer Product Gradient"), Hirstache et al(2001), MAVE("Minimum Average Variance Estimation".), Xia et al(2002)...

Réduction de dimension

... en Régression (i.i.d.)...

$$Y = m(X) + \varepsilon$$

- $\Longrightarrow$ la qualité de l'estimation non-paramétrique de la fonction $m(\cdot)$ est pénalisée par la dimension du régresseur

Modèle de réduction de dimension:

- $\Longrightarrow$ plusieurs solutions: la méthode de "Projection pursuit regression" (Friedman and Stueltzle, 1981), GAM,"General Additive Model" (Hastie et Tibshirani, 1986), ACE: Adevrage Component Estimation(Brieman et Friedman, 1985),...
- $\Longrightarrow$ SIR ("Sliced Inverse Regression"), Li(1991), OPG("Outer Product Gradient"), Hirstache et al(2001), MAVE("Minimum Average Variance Estimation".), Xia et al(2002)...

Réduction de dimension

... en Régression (i.i.d.)...

$$Y = m(X) + \varepsilon$$

- $\Longrightarrow$ la qualité de l'estimation non-paramétrique de la fonction $m(\cdot)$ est pénalisée par la dimension du régresseur

Modèle de réduction de dimension:

- $\Longrightarrow$ plusieurs solutions: la méthode de "Projection pursuit regression" (Friedman and Stueltzle, 1981), GAM,"General Additive Model" (Hastie et Tibshirani, 1986), ACE: Adevrage Component Estimation(Brieman et Friedman, 1985),...
- $\Longrightarrow$ SIR ("Sliced Inverse Regression"), Li(1991), OPG("Outer Product Gradient"), Hirstache et al(2001), MAVE("Minimum Average Variance Estimation".), Xia et al(2002)...

Modèle de réduction de dimension

... on suppose que ...

- on a

$$m(X) = g(\Phi.X)$$

- $g(.) : \mathbb{R}^D \longrightarrow \mathbb{R}$, fonction inconnue, $\Phi : \mathbb{R}^d \longrightarrow \mathbb{R}^D$, $D \leq d$.

- Φ, matrice normale ($\Phi^T \Phi = I_D$, Φ^T matrice orthogonale de Φ) à estimer.
- $g(.)$ pouvant être estimer par régression non-paramétrique de Y sur $\Phi.X$.

- L'espace de dimension D, $\mathrm{Im}(\Phi^T)$,
$\Longrightarrow$appelé *"Effective Dimension Reduction space"* (EDR).

Modèle de réduction de dimension

... on suppose que ...

- on a

$$m(X) = g(\Phi.X)$$

- $g(.) : \mathbb{R}^D \longrightarrow \mathbb{R}$, fonction inconnue, $\Phi : \mathbb{R}^d \longrightarrow \mathbb{R}^D$, $D \leq d$.

- Φ, matrice normale ($\Phi^T\Phi = I_D$, Φ^T matrice orthogonale de Φ) à estimer.
- $g(.)$ pouvant être estimer par régression non-paramétrique de Y sur $\Phi.X$.

- L'espace de dimension D, $\mathrm{Im}(\Phi^T)$,
$\Longrightarrow$appelé *"Effective Dimension Reduction space"* (EDR).

Modèle de réduction de dimension.

$$\ldots m(X) = g(\Phi.X) \ldots$$

- Si $(\phi_1, \ldots, \phi_D)$ est une base de l'EDR,

- le modèle se résume par:
- $\Leftrightarrow$ il suffit de D vecteurs de $\mathbb{R}^d$: $\phi_1, \ldots, \phi_D$, pour estimer $m(X)$.
- $\Rightarrow$ Modèle multi-indices: $m(x) = g(\phi_1' x, \ldots, \phi_D' x)$.
- ...Ainsi, estimer l'espace l'EDR$\Leftrightarrow$
- estimer une base (ϕ_i) de l'EDR ...
- ...et sa dimension, D.
- Divers méthodes d'estimation de l'EDR. $\Longrightarrow$ La Régression Inverse.

Modèle de réduction de dimension.

$$\ldots m(X) = g(\Phi.X) \ldots$$

- Si $(\phi_1, \ldots, \phi_D)$ est une base de l'EDR,

- le modèle se résume par:
- $\Leftrightarrow$ il suffit de D vecteurs de $\mathbb{R}^d$: $\phi_1, \ldots, \phi_D$, pour estimer $m(X)$.

- $\Rightarrow$ Modèle multi-indices: $m(x) = g(\phi_1' x, \ldots, \phi_D' x)$.

- ...Ainsi, estimer l'espace l'EDR$\Leftrightarrow$
- estimer une base (ϕ_i) de l'EDR ...
- ...et sa dimension, D.
- Divers méthodes d'estimation de l'EDR. $\Longrightarrow$ La Régression Inverse.

Modèle de réduction de dimension.

$$\ldots m(X) = g(\Phi.X) \ldots$$

- Si $(\phi_1, \ldots, \phi_D)$ est une base de l'EDR,

- le modèle se résume par:
- $\Leftrightarrow$ il suffit de D vecteurs de $\mathbb{R}^d$: $\phi_1, \ldots, \phi_D$, pour estimer $m(X)$.
- $\Rightarrow$ Modèle multi-indices: $m(x) = g(\phi_1' x, \ldots, \phi_D' x)$.
- ...Ainsi, estimer l'espace l'EDR$\Leftrightarrow$
- estimer une base (ϕ_i) de l'EDR ...
- ...et sa dimension, D.
- Divers méthodes d'estimation de l'EDR. $\Longrightarrow$ La Régression Inverse.

Modèle de réduction de dimension

... cadre général loi conditionnelle...

- Quantile conditionnelle: $Q(Y|X) = Q(Y|\Phi.X)$ (Gannoum et al. 2004),...
- Plus généralement:
- La loi de Y sachant X est la même que celle de Y sachant $\Phi.X$:

$$F(Y|X) = F(Y|\Phi.X)$$

- il suffit de D vecteurs de $\mathbb{R}^d$: $\phi_1, \ldots, \phi_D$, pour entièrement connaitre $Y|X$.

Modèle de réduction de dimension

... cadre général loi conditionnelle...

- Quantile conditionnelle: $Q(Y|X) = Q(Y|\Phi.X)$ (Gannoum et al. 2004),...
- Plus généralement:
- La loi de Y sachant X est la même que celle de Y sachant $\Phi.X$:

$$F(Y|X) = F(Y|\Phi.X)$$

- il suffit de D vecteurs de $\mathbb{R}^d$: $\phi_1, \dots, \phi_D$, pour entièrement connaitre $Y|X$.

Estimation de l'EDR

... par la régression inverse....

- **Condition.** (Li, 1991) $\forall b \in \mathbf{R}^p$, $E(b'X|\phi_1'X, \ldots, \phi_K'X)$ est une combinaison linéaire des $\phi_i'X$.

Théorème. (Li, 1991) L'EDR contient le sous-espace propre associé aux valeurs propres non-nulles de $\Sigma^{-1}\boldsymbol{var}(E(X|Y))$ $(\Sigma = \mathbf{var}(X))$.

- $\Longrightarrow$Si $\Sigma^{-1}\boldsymbol{var}(E(X|Y))$ est de plein rang,
- $\Longrightarrow$ l'estimation de l'EDR s'obtient par diagonalisation d'un estimateur de la matrice $\Sigma^{-1}\mathbf{var}(E(X|Y))$.

Estimation de l'EDR (cas i.i.d.)

... par estimation de $\Sigma^{-1} var(E(X|Y))$...

- Estimation de Σ par estimateur empirique.

- Estimation de **var**$(E(X|Y))$

  par régressogramme de $E(X|Y)$
  $\Rightarrow$SIR(Li, 1991): "Sliced Inverse Regression"

  par méthode à noyau de $E(X|Y)$
  $\Rightarrow$ (Fang and Zu, 1997): "kernel inverse regression"

Estimation de l'EDR (cadre i.i.d.)

... alternatives la régression inverse....

Quid: lorsque $\Sigma^{-1}\boldsymbol{var}(E(X|Y))$ n'est pas de plein rang?... voir dégénérée?

- SIR II (LI, 1991), SIR$-\alpha$ Saracco(1999,...) moment d'ordre $\geq$2, PHD (Li, 1992) : la matrice hessienne de la fonction $g(.)$,$(m(x) = g(\phi_1' x, ..., \phi_D' x))$.
- MAVE, OPG (Hirstache et al; 2001,)... developpement de Taylor de $g(.)$.

Problème d'estimation de Σ^{-1} en grande dimension.

- Réduction de dimension en régression en dimension infinie Ferré and Yao (2005), Amato, Feiss et Antoniadis (2006), Cook et al (2007, 2008), Bernard-Michel, Gardes and Girard(2008)...

Estimation de l'EDR (cadre i.i.d.)

... alternatives la régression inverse....

Quid: lorsque $\Sigma^{-1}\boldsymbol{var}(E(X|Y))$ n'est pas de plein rang?... voir dégénérée?

- SIR II (LI, 1991), SIR$-\alpha$ Saracco(1999,...) moment d'ordre $\geq$2, PHD (Li, 1992) : la matrice hessienne de la fonction $g(.)$,$(m(x) = g(\phi_1' x, ..., \phi_D' x))$.
- MAVE, OPG (Hirstache et al; 2001,)... developpement de Taylor de $g(.)$.

Problème d'estimation de Σ^{-1} en grande dimension.

- Réduction de dimension en régression en dimension infinie Ferré and Yao (2005), Amato, Feiss et Antoniadis (2006), Cook et al (2007, 2008), Bernard-Michel, Gardes and Girard(2008)...

Estimation de l'EDR

- L'estimation du modèle de réduction de dimension
  $\Rightarrow$ a été étudié dans le cadre i.i.d.

- Que devient cet estimation par régression inverse
  $\Rightarrow$ dans le cas de dépendance spatiale?

Régression spatiale

$\Longrightarrow$ Estimation de la fonction de régression $m(.)$:

$$Y = m(X) + \varepsilon,$$

$\Longrightarrow$ A partir d'observations spatialement.

Régression spatiale, prédiction spatiale

- Estimation paramétrique: large littérature surtout en géostatistique, krigeage (Cressie; 1993, Wackernagel; 1995, ...), krigeage,...

- Moins de résultats dans un cadre non-paramétrique:

Régression et prédiction non-paramétrique: Lu &Chen (2004), Biau & Cadre (2004), Dabo-Niang & Yao (2007).

$\Rightarrow$Estimateur étudié est une extension de l'estimateur à noyau du cas i.i.d.

$\Rightarrow$ Sa vitesse de convergence est également pénalisée par la dimension du régresseur: $h^k + \left(\frac{\log n}{n h_n^d}\right)^{1/2}$

Régression spatiale, prédiction spatiale

- Estimation paramétrique: large littérature surtout en géostatistique, krigeage (Cressie; 1993, Wackernagel; 1995, ...), krigeage,...

- Moins de résultats dans un cadre non-paramétrique:

Régression et prédiction non-paramétrique: Lu &Chen (2004), Biau & Cadre (2004), Dabo-Niang & Yao (2007).

$\Rightarrow$Estimateur étudié est une extension de l'estimateur à noyau du cas i.i.d.

$\Rightarrow$ Sa vitesse de convergence est également pénalisée par la dimension du régresseur: $h^k + \left(\frac{\log n}{nh_n^d}\right)^{1/2}$

Régression spatiale, prédiction spatiale

- Estimation paramétrique: large littérature surtout en géostatistique, krigeage (Cressie; 1993, Wackernagel; 1995, ...), krigeage,...

- Moins de résultats dans un cadre non-paramétrique:

Régression et prédiction non-paramétrique: Lu &Chen (2004), Biau & Cadre (2004), Dabo-Niang & Yao (2007).

$\Rightarrow$Estimateur étudié est une extension de l'estimateur à noyau du cas i.i.d.

$\Rightarrow$ Sa vitesse de convergence est également pénalisée par la dimension du régresseur: $h^k + \left(\frac{\log n}{n h_{\mathbf{n}}^d}\right)^{1/2}$

Réduction de dimension

... en régression spatiale...

- Objectif: Etendre les méthodes d'estimation (en regression inverse) au cadre spatiale.

- Que devient l'estimation de l'EDR?$\Longrightarrow$ $\Sigma^{-1}\mathbf{var}(E(X|Y))$**?**

Réduction de dimension

... en régression spatiale...

- Objectif: Etendre les méthodes d'estimation (en regression inverse) au cadre spatiale.

- Que devient l'estimation de l'EDR?$\Longrightarrow$ $\Sigma^{-1}\mathbf{var}(E(X|Y))$**?**

Notations

- $\mathbf{n} = (n_1, ..., n_N) \Longrightarrow \mathcal{I}_{\mathbf{n}} = \left\{\mathbf{i} \in \mathbb{Z}^N : 1 \leq i_k \leq n_k, k = 1, ..., N\right\}$.
- On dira que $\mathbf{n} \to +\infty$ si $\min_{i=1,...,N} n_i \to +\infty$ et $|n_i/n_k| < C$, $C > 0$.
- $\widehat{\mathbf{n}} = n_1 \times ... \times n_N$.

$\Longrightarrow$Estimation de $\Sigma^{-1}\mathbf{var}(E(X|Y))$ à partir des observations:

$$(Z_{\mathbf{i}}, \mathbf{i} \in \mathcal{I}_{\mathbf{n}}),\ Z_{\mathbf{i}} = (X_{\mathbf{i}}, Y_{\mathbf{i}}).$$

Notations

- $\mathbf{n} = (n_1, \ldots, n_N) \Longrightarrow \mathscr{I}_{\mathbf{n}} = \left\{\mathbf{i} \in \mathbb{Z}^N : 1 \leq i_k \leq n_k, k = 1, \ldots, N\right\}$.
- On dira que $\mathbf{n} \to +\infty$ si $\min_{i=1,\ldots,N} n_i \to +\infty$ et $|n_i/n_k| < C$, $C > 0$.
- $\widehat{\mathbf{n}} = n_1 \times \ldots \times n_N$.

$\Longrightarrow$Estimation de $\Sigma^{-1}\mathbf{var}(E(X|Y))$ à partir des observations:

$$(Z_{\mathbf{i}}, \mathbf{i} \in \mathscr{I}_{\mathbf{n}}),\ Z_{\mathbf{i}} = (X_{\mathbf{i}},\ Y_{\mathbf{i}})\,.$$

Estimateur en régression inverse

....cadre spatial...

On construit l'estimateur:

$$r_{\mathbf{n}}(y) = \frac{\varphi_{\mathbf{n}}(y)}{f_{\mathbf{n}}(y)}, y \in \mathbb{R}$$

avec $f_{\mathbf{n}}(y) = \frac{1}{\widehat{\mathbf{n}} h_{\mathbf{n}}} \sum_{\mathbf{i} \in \mathscr{I}_{\mathbf{n}}} K\left(\frac{y - Y_{\mathbf{i}}}{h_{\mathbf{n}}}\right)$ et $\varphi_{\mathbf{n}}(y) = \frac{1}{\widehat{\mathbf{n}} h_{\mathbf{n}}} \sum_{\mathbf{i} \in \mathscr{I}_{\mathbf{n}}} X_{\mathbf{i}} K\left(\frac{y - Y_{\mathbf{i}}}{h_{\mathbf{n}}}\right)$

où $\lim_{\mathbf{n} \to +\infty} h_{\mathbf{n}} = 0(+)$ and $K : \mathbb{R}^d \to \mathbb{R}$ est un noyau.

Pour éviter les valeurs trop petite de $f_{\mathbf{n}}$, on choisit de travailler avec:
$r_{e,\mathbf{n}}(y) = \frac{\varphi_{\mathbf{n}}(y)}{f_{e,\mathbf{n}}(y)}$ où $f_{e,\mathbf{n}}(y) = \max(e_{\mathbf{n}}, f_{\mathbf{n}}(y))$ et $\lim_{\mathbf{n} \to \infty} e_{\mathbf{n}} = 0$.

On construit l'estimateur de $\mathbf{Var}(\mathbf{E}(X|Y))$:

$$\Sigma_{e,\mathbf{n}} = \frac{1}{\hat{\mathbf{n}}} \sum r_{e,\mathbf{n}}(Y_{\mathbf{i}}) r_{e,\mathbf{n}}(Y_{\mathbf{i}})^T - \overline{X}\,\overline{X}^T.$$

Mesure de dépendance spatiale

....Hypothèses...

- Condition de mélange:

$$\alpha(v) = \sup_{\mathbf{u},\mathbf{u}' \in \mathbb{R}^N,\, \|\mathbf{u}-\mathbf{u}'\| = v} \alpha\left(\sigma\left(Z_{\mathbf{u}}\right), \sigma\left(Z_{\mathbf{u}'}\right)\right), \qquad v \geq 0,$$

où pour $\mathscr{B}$ et $\mathscr{C}$ deux sous $\sigma-$algèbre de $\mathscr{A}$,
$\alpha(\mathscr{B},\mathscr{C}) = \sup_{B \in \mathscr{B},\, C \in \mathscr{C}} |P(B \cap C) - P(B)P(C)|$.
$\lim_{u \to \infty} \alpha(u) = 0$.

- Dépendance locale : Pour $\ell = 1, ..., d$, il existe $\Delta > 0$ tel que pour tout $(X_{\mathbf{i}}^{(\ell)}, X_{\mathbf{j}})$ et $((X_{\mathbf{i}}^{(\ell)}, Y_{\mathbf{i}}), (X_{\mathbf{j}}^{(\ell)}, Y_{\mathbf{j}}))$ tel que $\mathrm{dist}(\mathbf{i}, \mathbf{j}) > \Delta$ et de densités $f_{\mathbf{i},\mathbf{j}}$ et $g_{\mathbf{i},\mathbf{j}}$ alors

$$|f_{\mathbf{i},\mathbf{j}}(x,y) - f(x) f(y)| \leq C, \ \forall x, y \in \mathbb{R}$$

$$|g_{\mathbf{i},\mathbf{j}}(u,v) - g(u) g(v)| \leq C, \ \forall u, v \in \mathbb{R}^2$$

Mesure de dépendance spatiale

....Hypothèses...

- Condition de mélange:

$$\alpha(v) = \sup_{\mathbf{u},\mathbf{u}' \in \mathbb{R}^N,\, \|\mathbf{u}-\mathbf{u}'\|=v} \alpha\left(\sigma\left(Z_{\mathbf{u}}\right), \sigma\left(Z_{\mathbf{u}'}\right)\right), \qquad v \geq 0,$$

où pour $\mathscr{B}$ et $\mathscr{C}$ deux sous $\sigma-$algèbre de $\mathscr{A}$,
$\alpha(\mathscr{B},\mathscr{C}) = \sup_{B \in \mathscr{B},\, C \in \mathscr{C}} |P(B \cap C) - P(B)P(C)|.$
$\lim_{u\to\infty} \alpha(u) = 0.$

- Dépendance locale : Pour $\ell = 1, ..., d$, il existe $\Delta > 0$ tel que pour tout $(X_{\mathbf{i}}^{(\ell)}, X_{\mathbf{j}})$ et $((X_{\mathbf{i}}^{(\ell)}, Y_{\mathbf{i}}), (X_{\mathbf{j}}^{(\ell)}, Y_{\mathbf{j}}))$ tel que $\text{dist}(\mathbf{i}, \mathbf{j}) > \Delta$ et de densités $f_{\mathbf{i},\mathbf{j}}$ et $g_{\mathbf{i},\mathbf{j}}$ alors

$$|f_{\mathbf{i},\mathbf{j}}(x,y) - f(x)f(y)| \leq C, \ \forall x, y \in \mathbb{R}$$

$$|g_{\mathbf{i},\mathbf{j}}(u,v) - g(u)g(v)| \leq C, \ \forall u, v \in \mathbb{R}^2$$

Estimateur à noyau de la régression

....cadre spatial...

Etudier par Carbon et al. (1997) Lu &Chen (2004), Biau & Cadre (2004), Dabo & Yao (2007)
Sous certaines hypothèses.

Théorème.
$\sup_{y\in\mathbb{R}}|f_{\mathbf{n}}(y)-f(y)| = \mathcal{O}_p\left(h_{\mathbf{n}}^k + \frac{\sqrt{\log \hat{\mathbf{n}}}}{\sqrt{\hat{\mathbf{n}} h_{\mathbf{n}}}}\right)$,
$\sup_{y\in\mathbb{R}}||\varphi_{\mathbf{n}}(y)-\varphi(y)|| = \mathcal{O}_p\left(h_{\mathbf{n}}^k + \frac{\sqrt{\log \hat{\mathbf{n}}}}{\sqrt{\hat{\mathbf{n}} h_{\mathbf{n}}}}\right)$ avec k entier lié à la régularité des fonction f et φ.

Estimateur à noyau de la régression

....cadre spatial...

Etudier par Carbon et al. (1997) Lu &Chen (2004), Biau & Cadre (2004), Dabo & Yao (2007)
Sous certaines hypothèses.

Théorème.
$\sup_{y\in\mathbb{R}}|f_{\mathbf{n}}(y)-f(y)| = \mathcal{O}_p\left(h_{\mathbf{n}}^k + \frac{\sqrt{\log\hat{\mathbf{n}}}}{\sqrt{\hat{\mathbf{n}}h_{\mathbf{n}}}}\right)$,
$\sup_{y\in\mathbb{R}}||\varphi_{\mathbf{n}}(y)-\varphi(y)|| = \mathcal{O}_p\left(h_{\mathbf{n}}^k + \frac{\sqrt{\log\hat{\mathbf{n}}}}{\sqrt{\hat{\mathbf{n}}h_{\mathbf{n}}}}\right)$ avec k entier lié à la régularité des fonction f et φ.

Estimateur de EDR

... Estimation de $\Sigma_e = \mathbf{var}(\mathbf{E}X|Y)$...

Sous certaines conditions:

Théorème: si $\alpha(t) \leq Ct^{-\theta}$, $t > 0$, $\theta > 2N$ et si $\widehat{\mathbf{n}} h_{\mathbf{n}}^3 (\log \widehat{\mathbf{n}})^{-1} \to 0$, $\widehat{\mathbf{n}} h_{\mathbf{n}}^{\theta_1} (\log \widehat{\mathbf{n}})^{-1} \to \infty$ avec $\theta_1 = \frac{4N+\theta}{\theta-2N}$, alors:

$$\Sigma_{e,\mathbf{n}} - \Sigma_e = \mathcal{O}_p\left(e_{\mathbf{n}} + h_{\mathbf{n}}^k + \frac{\Psi_{\mathbf{n}}^2}{e_{\mathbf{n}}^2}\right)$$

Corollaire: Prennant $h \simeq \hat{\mathbf{n}}^{-c_1}$, $e_n \simeq \hat{\mathbf{n}}^{-c_2}$ pour des c_1 et c_2 tel que $\frac{c_2}{k} + \frac{1}{2k} < c_1 < \frac{1}{4} - c_2$, on a:

$$\Sigma_{e,\mathbf{n}} - \Sigma_e = o_p\left(\frac{1}{\sqrt{\hat{\mathbf{n}}}}\right) \text{ et } \sqrt{\hat{\mathbf{n}}}\left(\Sigma_{e,\mathbf{n}} - \Sigma_e\right) \xrightarrow{\mathcal{L}} \Lambda.$$

- Et par conséquent vitesse similaire pour $\Sigma_{\mathbf{n}}^{-1}\Sigma_{e,\mathbf{n}}$ tant que la dimension du régresseur reste raisonnable.

Predicteur spatial.

Soit $(\xi_{\mathbf{n}}, \mathbf{n} \in (\mathbb{N}^*)^N)$, $\xi_{\mathbf{n}} \in \mathbb{R}$, stationnaire et observé dans $\mathcal{O}_{\mathbf{n}} \subset \mathcal{I}_{\mathbf{n}}$

- Objectif: prédire la valeur $\xi_{\mathbf{i}_0}$, en $\mathbf{i}_0 \in \mathcal{I}_{\mathbf{n}} \backslash \mathcal{O}_{\mathbf{n}}$.
- On suppose que: $\xi_{\mathbf{i}_0}$ ne dépend que des valeurs $\xi_{\mathbf{i}}$, $\mathbf{i} \in \mathcal{V}_{\mathbf{i}_0}$, $\mathcal{V}_{\mathbf{i}_0} \subset \mathcal{O}_{\mathbf{n}}$
- Meilleur prédicteur de $\xi_{\mathbf{i}_0}$ au sens des moindres carré: $E(\xi_{\mathbf{i}_0} | \xi_{\mathbf{i}}, \mathbf{i} \in \mathcal{V}_{\mathbf{i}_0})$.

$\Longrightarrow$Prédiction à partir d'une **estimation** de $E(\xi_{\mathbf{i}_0} | \xi_{\mathbf{i}}, \mathbf{i} \in \mathcal{V}_{\mathbf{i}_0})$.

Predicteur spatial.

Soit $(\xi_{\mathbf{n}}, \mathbf{n} \in (\mathbb{N}^*)^N)$, $\xi_{\mathbf{n}} \in \mathbb{R}$, stationnaire et observé dans $\mathscr{O}_{\mathbf{n}} \subset \mathscr{I}_{\mathbf{n}}$

- Objectif: prédire la valeur $\xi_{\mathbf{i}_0}$, en $\mathbf{i}_0 \in \mathscr{I}_{\mathbf{n}} \backslash \mathscr{O}_{\mathbf{n}}$.

- On suppose que: $\xi_{\mathbf{i}_0}$ ne dépend que des valeurs $\xi_{\mathbf{i}}$, $\mathbf{i} \in \mathscr{V}_{\mathbf{i}_0}$, $\mathscr{V}_{\mathbf{i}_0} \subset \mathscr{O}_{\mathbf{n}}$

- Meilleur prédicteur de $\xi_{\mathbf{i}_0}$ au sens des moindres carré: $E(\xi_{\mathbf{i}_0} | \xi_{\mathbf{i}}, \mathbf{i} \in \mathscr{V}_{\mathbf{i}_0})$.

  $\Longrightarrow$Prédiction à partir d'une **estimation** de $E(\xi_{\mathbf{i}_0} | \xi_{\mathbf{i}}, \mathbf{i} \in \mathscr{V}_{\mathbf{i}_0})$.

Predicteur spatial.

Soit $(\xi_{\mathbf{n}}, \mathbf{n} \in (\mathbb{N}^*)^N)$, $\xi_{\mathbf{n}} \in \mathbb{R}$, stationnaire et observé dans $\mathscr{O}_{\mathbf{n}} \subset \mathscr{I}_{\mathbf{n}}$

- Objectif: prédire la valeur $\xi_{\mathbf{i}_0}$, en $\mathbf{i}_0 \in \mathscr{I}_{\mathbf{n}} \backslash \mathscr{O}_{\mathbf{n}}$.

- On suppose que: $\xi_{\mathbf{i}_0}$ ne dépend que des valeurs $\xi_{\mathbf{i}}$, $\mathbf{i} \in \mathscr{V}_{\mathbf{i}_0}$, $\mathscr{V}_{\mathbf{i}_0} \subset \mathscr{O}_{\mathbf{n}}$

- Meilleur prédicteur de $\xi_{\mathbf{i}_0}$ au sens des moindres carré: $E(\xi_{\mathbf{i}_0} | \xi_{\mathbf{i}}, \mathbf{i} \in \mathscr{V}_{\mathbf{i}_0})$.

$\Longrightarrow$Prédiction à partir d'une **estimation** de $E(\xi_{\mathbf{i}_0} | \xi_{\mathbf{i}}, \mathbf{i} \in \mathscr{V}_{\mathbf{i}_0})$.

Predicteur spatial.

Soit $(\xi_{\mathbf{n}}, \mathbf{n} \in (\mathbb{N}^*)^N)$, $\xi_{\mathbf{n}} \in \mathbb{R}$, stationnaire et observé dans $\mathcal{O}_{\mathbf{n}} \subset \mathcal{I}_{\mathbf{n}}$

- Objectif: prédire la valeur $\xi_{\mathbf{i}_0}$, en $\mathbf{i}_0 \in \mathcal{I}_{\mathbf{n}} \backslash \mathcal{O}_{\mathbf{n}}$.
- On suppose que: $\xi_{\mathbf{i}_0}$ ne dépend que des valeurs $\xi_{\mathbf{i}}$, $\mathbf{i} \in \mathcal{V}_{\mathbf{i}_0}$, $\mathcal{V}_{\mathbf{i}_0} \subset \mathcal{O}_{\mathbf{n}}$
- Meilleur prédicteur de $\xi_{\mathbf{i}_0}$ au sens des moindres carré: $E(\xi_{\mathbf{i}_0} | \xi_{\mathbf{i}}, \mathbf{i} \in \mathcal{V}_{\mathbf{i}_0})$.

$\Longrightarrow$Prédiction à partir d'une **estimation** de $E(\xi_{\mathbf{i}_0} | \xi_{\mathbf{i}}, \mathbf{i} \in \mathcal{V}_{\mathbf{i}_0})$.

Predicteur spatial.

Soit $(\xi_{\mathbf{n}}, \mathbf{n} \in (\mathbb{N}^*)^N)$, $\xi_{\mathbf{n}} \in \mathbb{R}$, stationnaire et observé dans $\mathscr{O}_{\mathbf{n}} \subset \mathscr{I}_{\mathbf{n}}$

- Objectif: prédire la valeur $\xi_{\mathbf{i}_0}$, en $\mathbf{i}_0 \in \mathscr{I}_{\mathbf{n}} \backslash \mathscr{O}_{\mathbf{n}}$.

- On suppose que: $\xi_{\mathbf{i}_0}$ ne dépend que des valeurs $\xi_{\mathbf{i}}$, $\mathbf{i} \in \mathscr{V}_{\mathbf{i}_0}$, $\mathscr{V}_{\mathbf{i}_0} \subset \mathscr{O}_{\mathbf{n}}$

- Meilleur prédicteur de $\xi_{\mathbf{i}_0}$ au sens des moindres carré: $E(\xi_{\mathbf{i}_0} | \xi_{\mathbf{i}}, \mathbf{i} \in \mathscr{V}_{\mathbf{i}_0})$.

$\Longrightarrow$Prédiction à partir d'une **estimation** de $E(\xi_{\mathbf{i}_0} | \xi_{\mathbf{i}}, \mathbf{i} \in \mathscr{V}_{\mathbf{i}_0})$.

Predicteur spatial.

Soit $(\xi_{\mathbf{n}}, \mathbf{n} \in (\mathbb{N}^*)^N)$, $\xi_{\mathbf{n}} \in \mathbb{R}$, stationnaire et observé dans $\mathcal{O}_{\mathbf{n}} \subset \mathcal{I}_{\mathbf{n}}$

- Pose: $d = \text{card}\{\mathcal{V}\}$, $\xi_{\mathbf{i}_0}^d = (\xi_{\mathbf{i}}, \mathbf{i} \in \mathcal{V}_{\mathbf{i}_0})$.
- Le problème de prédiction $\Longleftrightarrow$ estimation de $m(x) = E(\xi_{\mathbf{i}_0} | \xi_{\mathbf{i}_0}^d = x)$.
- Cas linéaire $\Longrightarrow$Krigeage (Cressie; 1991, Wackernagel; 1995),
- Cas non paramétrique Biau et Cadre (2004) ,
- $\Longrightarrow$Application de la méthode précédente pour estimer $m(x) = g(\Phi.X)$.
- Permet d'estimer le nombre de voisins intéressant à la prédiction car puisque le processus est supposé markovien, la régression inverse: $r(y) = \mathbf{E}(\xi_{\mathbf{i}_0}^d | \xi_{\mathbf{i}_0} = y)$ permettra d'éliminer les éléments du voisinage (choisi) indépendants de $\xi_{\mathbf{i}_0}$. Nous proposons un algorithme permettant de le faire.

Predicteur spatial.

Soit $(\xi_{\mathbf{n}}, \mathbf{n} \in (\mathbb{N}^*)^N)$, $\xi_{\mathbf{n}} \in \mathbb{R}$, stationnaire et observé dans $\mathcal{O}_{\mathbf{n}} \subset \mathcal{I}_{\mathbf{n}}$

- Pose: $d = \mathrm{card}\{\mathcal{V}\}$, $\xi_{\mathbf{i}_0}^d = (\xi_{\mathbf{i}}, \mathbf{i} \in \mathcal{V}_{\mathbf{i}_0})$.

- Le problème de prédiction $\Longleftrightarrow$ estimation de $m(x) = E(\xi_{\mathbf{i}_0} | \xi_{\mathbf{i}_0}^d = x)$.

- Cas linéaire $\Longrightarrow$Krigeage (Cressie; 1991, Wackernagel; 1995),
- Cas non paramétrique Biau et Cadre (2004) ,

- $\Longrightarrow$Application de la méthode précédente pour estimer $m(x) = g(\Phi.X)$.

- Permet d'estimer le nombre de voisins intéressant à la prédiction car puisque le processus est supposé markovien, la régression inverse: $r(y) = \mathbf{E}(\xi_{\mathbf{i}_0}^d | \xi_{\mathbf{i}_0} = y)$ permettra d'éliminer les éléments du voisinage (choisi) indépendants de $\xi_{\mathbf{i}_0}$. Nous proposons un algorithme permettant de le faire.

Predicteur spatial.

Soit $(\xi_{\mathbf{n}}, \mathbf{n} \in (\mathbb{N}^*)^N)$, $\xi_{\mathbf{n}} \in \mathbb{R}$, stationnaire et observé dans $\mathscr{O}_{\mathbf{n}} \subset \mathscr{I}_{\mathbf{n}}$

- Pose: $d = \text{card}\{\mathscr{V}\}$, $\xi^d_{\mathbf{i}_0} = (\xi_{\mathbf{i}}, \mathbf{i} \in \mathscr{V}_{\mathbf{i}_0})$.

- Le problème de prédiction $\Longleftrightarrow$ estimation de $m(x) = E(\xi_{\mathbf{i}_0} | \xi^d_{\mathbf{i}_0} = x)$.

- Cas linéaire $\Longrightarrow$Krigeage (Cressie; 1991, Wackernagel; 1995),
- Cas non paramétrique Biau et Cadre (2004) ,

- $\Longrightarrow$Application de la méthode précédente pour estimer $m(x) = g(\Phi.X)$.

- Permet d'estimer le nombre de voisins intéressant à la prédiction car puisque le processus est supposé markovien, la régression inverse: $r(y) = \mathbf{E}(\xi^d_{\mathbf{i}_0} | \xi_{\mathbf{i}_0} = y)$ permettra d'éliminer les éléments du voisinage (choisi) indépendants de $\xi_{\mathbf{i}_0}$. Nous proposons un algorithme permettant de le faire.

Predicteur spatial.

Soit $(\xi_{\mathbf{n}}, \mathbf{n} \in (\mathbb{N}^*)^N)$, $\xi_{\mathbf{n}} \in \mathbb{R}$, stationnaire et observé dans $\mathcal{O}_{\mathbf{n}} \subset \mathcal{I}_{\mathbf{n}}$

- Pose: $d = \text{card}\{\mathcal{V}\}$, $\xi_{\mathbf{i_0}}^d = (\xi_{\mathbf{i}}, \mathbf{i} \in \mathcal{V}_{\mathbf{i_0}})$.
- Le problème de prédiction $\Longleftrightarrow$ estimation de $m(x) = E(\xi_{\mathbf{i_0}} | \xi_{\mathbf{i_0}}^d = x)$.
- Cas linéaire $\Longrightarrow$Krigeage (Cressie; 1991, Wackernagel; 1995),
- Cas non paramétrique Biau et Cadre (2004) ,
- $\Longrightarrow$Application de la méthode précédente pour estimer $m(x) = g(\Phi.X)$.
- Permet d'estimer le nombre de voisins intéressant à la prédiction car puisque le processus est supposé markovien, la régression inverse: $r(y) = \mathbf{E}(\xi_{\mathbf{i_0}}^d | \xi_{\mathbf{i_0}} = y)$ permettra d'éliminer les éléments du voisinage (choisi) indépendants de $\xi_{\mathbf{i_0}}$. Nous proposons un algorithme permettant de le faire.

Predicteur spatial.

Soit $(\xi_{\mathbf{n}}, \mathbf{n} \in (\mathbb{N}^*)^N)$, $\xi_{\mathbf{n}} \in \mathbb{R}$, stationnaire et observé dans $\mathcal{O}_{\mathbf{n}} \subset \mathcal{I}_{\mathbf{n}}$

- Pose: $d = \text{card}\{\mathcal{V}\}$, $\xi_{\mathbf{i_0}}^d = (\xi_{\mathbf{i}}, \mathbf{i} \in \mathcal{V}_{\mathbf{i_0}})$.
- Le problème de prédiction $\Longleftrightarrow$ estimation de $m(x) = E(\xi_{\mathbf{i_0}} | \xi_{\mathbf{i_0}}^d = x)$.
- Cas linéaire $\Longrightarrow$Krigeage (Cressie; 1991, Wackernagel; 1995),
- Cas non paramétrique Biau et Cadre (2004) ,
- $\Longrightarrow$Application de la méthode précédente pour estimer $m(x) = g(\Phi.X)$.
- Permet d'estimer le nombre de voisins intéressant à la prédiction car puisque le processus est supposé markovien, la régression inverse: $r(y) = \mathbf{E}(\xi_{\mathbf{i_0}}^d | \xi_{\mathbf{i_0}} = y)$ permettra d'éliminer les éléments du voisinage (choisi) indépendants de $\xi_{\mathbf{i_0}}$. Nous proposons un algorithme permettant de le faire.

Predicteur spatial.

Soit $(\xi_{\mathbf{n}}, \mathbf{n} \in (\mathbb{N}^*)^N)$, $\xi_{\mathbf{n}} \in \mathbb{R}$, stationnaire et observé dans $\mathcal{O}_{\mathbf{n}} \subset \mathcal{I}_{\mathbf{n}}$

- Pose: $d = \text{card}\{\mathcal{V}\}$, $\xi^d_{\mathbf{i_0}} = (\xi_{\mathbf{i}}, \mathbf{i} \in \mathcal{V}_{\mathbf{i_0}})$.
- Le problème de prédiction $\Longleftrightarrow$ estimation de $m(x) = E(\xi_{\mathbf{i_0}} | \xi^d_{\mathbf{i_0}} = x)$.
- Cas linéaire $\Longrightarrow$Krigeage (Cressie; 1991, Wackernagel; 1995),
- Cas non paramétrique Biau et Cadre (2004) ,
- $\Longrightarrow$Application de la méthode précédente pour estimer $m(x) = g(\Phi.X)$.
- Permet d'estimer le nombre de voisins intéressant à la prédiction car puisque le processus est supposé markovien, la régression inverse: $r(y) = \mathbf{E}(\xi^d_{\mathbf{i_0}} | \xi_{\mathbf{i_0}} = y)$ permettra d'éliminer les éléments du voisinage (choisi) indépendants de $\xi_{\mathbf{i_0}}$. Nous proposons un algorithme permettant de le faire.

Perspectives

- Applications sur des données.
- Continuer cette extension en regardant les alternatives: les problèmes en grande dimension.
    - comme alternative Régression fonctionnelle spatial (Dabo-Niang, Rachdi & Yao, 2008).
- Le cadre spatial continu.

Merci de votre attention!

Perspectives

- Applications sur des données.
- Continuer cette extension en regardant les alternatives: les problèmes en grande dimension.
    - comme alternative Régression fonctionnelle spatial (Dabo-Niang, Rachdi & Yao, 2008).
- Le cadre spatial continu.

Merci de votre attention!

Algorithm for estimation of d, the number of neighbors.

1. Initialization: specify a parameter $\delta > 0$ (small) and fix a site $\mathbf{j}_0$; set $k = 1$.
2. compute $r_{\mathbf{n}}^{(k)}(y) = \dfrac{\displaystyle\sum_{\mathbf{i}\in\mathcal{O}_{\mathbf{n}},\mathcal{V}_{\mathbf{j}_0}\subset\mathcal{O}_{\mathbf{n}}} \xi_{\mathbf{i}(k)} K_{h_{\mathbf{n}}}(y-\xi_{\mathbf{i}})}{\displaystyle\sum_{\mathbf{i}\in\mathcal{O}_{\mathbf{n}},\mathcal{V}_{\mathbf{j}_0}\subset\mathcal{O}_{\mathbf{n}}} K_{h_{\mathbf{n}}}(y-\xi_{\mathbf{i}})}$, the kernel estimate of $r^{(k)}(y) = \mathbf{E}(X^{(k)}|Y = y)$
3. if $|(r_{\mathbf{n}}^{(k)}(y)| > \delta$, then $k = k+1$ and continue with Step 2; otherwise terminate and $d = k$.

Then, we can compute a predictor based on $d = k$:

The dimension reduction predictor

To get the predictor, we suggest the following algorithm:

1. compute

$$r_{\mathbf{n}}^*(y) = \frac{\sum_{\mathbf{i}\in\mathcal{O}_{\mathbf{n}},\mathcal{V}_{\mathbf{i}_0}\subset\mathcal{O}_{\mathbf{n}}} \xi_{\mathbf{i}}^d K_{h_{\mathbf{n}}}(y-\xi_{\mathbf{i}})}{\sum_{\mathbf{i}\in\mathcal{O}_{\mathbf{n}},\mathcal{V}_{\mathbf{i}_0}\subset\mathcal{O}_{\mathbf{n}}} K_{h_{\mathbf{n}}}(y-\xi_{\mathbf{i}})}$$

2. compute

$$\Sigma_{e,\mathbf{n}} = \frac{1}{\hat{\mathbf{n}}}\sum_{\mathbf{i}\in\mathcal{O}_{\mathbf{n}},\mathcal{V}_{\mathbf{i}_0}\subset\mathcal{O}_{\mathbf{n}}} r_{e,\mathbf{n}}^*(Y_{\mathbf{i}})\, r_{e,\mathbf{n}}^*(Y_{\mathbf{i}})^T - \overline{X}\,\overline{X}^T.$$

3. Do the PCA of $\Sigma_{\mathbf{n}}^{-1}\Sigma_{e,\mathbf{n}}$ both to get a basis of $\mathrm{Im}(\Sigma_{\mathbf{n}}^{-1}\Sigma_{e,\mathbf{n}})$ and estimation of the D, the dimension of $\mathrm{Im}(\Phi)$ as suggested in the next remark
4. compute the predictor:

$$\hat{\xi}_{\mathbf{i}_0} = g_{\mathbf{n}}^*(\Phi_{\mathbf{n}}^*.X_{\mathbf{i}_0}).$$

based on data $(Z_{\mathbf{i}}, \mathbf{i}\in\mathcal{O}_{\mathbf{n}})$; where $g_{\mathbf{n}}^*$ is the kernel estimate: