
Remerciements

*En premier lieu, je tiens à remercier mon **Dieu** qui ma donné le courage et la volonté pour réaliser ce travail. Je tiens à exprimer ma reconnaissance à ma directrice de mémoire M^{me} **Fatiha Mokhtari**, pour la confiance qu'elle ma accordé. Je remercie profondément pour sa disponibilité, ses orientations et ses remarques fructueuses. Qu'elle trouve ici l'expression de ma profonde gratitude.*

*Mais remerciements les plus profondes s'adressent également à monsieur **T. Guendouzi**, M^{me} **F. Benziadi** et monsieur **F. Madani**, pour l'intérêt qu'ils ont a porté à ce travail en acceptant de le juger.*

Nous tenons à remercier à titre individuel mes enseignants :

***A. Kandouci**, M^{elle} **F. Benziadi**, **M. Kadi**, M^{elle} **N. Ait**, M^{elle} **R. Rouane**, **S. Ouakkas**,
S. Abbas, **K. Djerfi**, **M. Belmekki**, **A. Azzouz**, **F. Hathout**, **O. Bennihi**, **H. Dida**.*

merci.

Dédicaces

Je dédie ce travail au l'âme de mon Père qui a laissé un vide immense. A ma très chère Mère qui a sacrifié sa vie pour que je puisse atteindre mon objectif que Dieu la garde pour moi.

A mon chere frère "Mohamed".

A mes chères soeurs "Sameh, Abir, Manel, et Hiba" pour leur soutien et leur accueil chaleureux.

A ma grand père.

A mes oncles, mes tentes et mes cousins.

A mes collègues "Djemaa, Bahia, Amina, Nacira, Bouazza, Mohamed et Sadek".

Finalemnt, je dédie ce modeste travail à tous les familles "Bendada", "Meziane" et "Khaldi" et ceux qui ont contribués de prêt ou de loin l'aboutissement de ce travail.

merci.

Table des matières

1	Processus Autorégessif Hilbertien	7
1.1	Notations et définitions	7
1.2	Estimation et Prévission	10
1.2.1	Estimation des opérateurs de covariance Γ et Δ	11
1.2.2	Estimation des éléments propres de Γ	12
1.2.3	Estimation de ρ et Prévission de X_{n+1}	14
2	La méthode du noyau	15
2.1	Rappels	15
2.1.1	Les variables fonctionnelles	15
2.1.2	La méthode du noyau	15
2.2	Régression scalaire par noyau	16
2.2.1	Etude de la convergence presque complète	17
2.3	Régression fonctionnelle par noyau	19
2.3.1	Présentation du modèle	19
2.3.2	Présentation de l'estimateur à noyau fonctionnel	19
2.4	Prédiction du modèles autorégressifs Hilbertiens	22
2.5	Ecritures matricielles	22
2.5.1	Spline de lissage	22
2.5.2	Régression fonctionnelle par noyau	23
2.5.3	Estimation d'un ARH(1)	24
2.5.4	ARH(1) Local	24
2.5.5	Interpolation linéaire	25
3	Simulation	27
3.1	Présentation des données	27
3.1.1	Simulation d'un H-bruit blanc	28
3.1.2	Définition d'un opérateur ρ	29
3.1.3	Une valeur initiale de X_0	30
3.2	Les fonctions utilisés dans R	30

3.3	Application	35
3.3.1	Sur des données simulés	35
3.3.2	Sur des données réels	37

Introduction

L'emploi de modèles en temps discret sur les observations d'un processus dont le vrai modèle est à temps continu peut négliger des caractéristiques essentielles de ce processus et nuire aux prévisions. Il faut garder à l'esprit que si les observations sont toujours en temps discret, il n'en n'est pas de même pour le processus qui les a générées. L'étude d'observations en temps discret n'est pas alors une justification valable pour l'emploi d'un modèle en temps discret.

Les variables aléatoires à valeurs dans des espaces fonctionnels sont considérées naturellement dans plusieurs domaines des probabilités et statistiques. Aussi l'interprétation d'un processus stochastique à temps continu comme une variable aléatoire à valeurs dans un espace fonctionnel a trouvé un grand intérêt aussi bien dans les théorèmes limites pour des variables aléatoires à valeurs dans des espaces de Banach que dans l'inférence et la prévision des processus stochastiques. Nous étudions la prédiction d'un processus aléatoire à temps continu par la méthode du noyau sur tout un intervalle par une modélisation autorégressive fonctionnelle.

Le type d'espace fonctionnel nous semble intéressant dans l'application : est l'espace de Hilbert.

Dans le premier chapitre, on donne un rappel sur les espaces de Hilbert. Historiquement, *D.Bosq* et *T.Mourid*(1990) sont les premiers à introduire les processus autorégressifs à valeurs dans un espace de Hilbert que nous notons *ARH*. En particulier, nous nous intéressons aux processus autorégressifs d'ordre un, notés *ARH*(1), on donne quelques solutions des problèmes liés à la prévision de ces processus.

Dans le deuxième chapitre, on traite la prévision de processus fonctionnels hilbertien d'ordre un par des méthodes du noyau (scalaire et fonctionnelle) un, on se propose d'étudier l'estimateur à noyau du processus autorégressif hilbertien dans le cas où la variable explicative est une courbe, en considérant les deux cas pour déterminer les courbes : cas où la variable explicative réel, et le cas où la variable explicative fonctionnelle. On trouve également un résultat de convergence presque complète d'estimateurs à noyau.

Dans le troisième chapitre, nous présentons les études de simulation. Nous reprenons l'exemple numérique du processus *ARH*(1) donnée par Julien Damon qui prend en considération le travail de Pumo(1993). On présente le modèle utilisé pour simuler $n+1$ observations d'un processus *ARH*(1) et on donne quelques

détails sur le calcul du prédicteur à noyau. Pour simuler le processus et les prédicteurs on utilise le logiciel **R** avec spécification d'un modèle particulier. Nous appliquons cette méthode de prévision à la série chronologique climatologique décrivant le phénomène (ENSO) pendant une période de douze mois de l'année 1986. De même, on s'intéresse à la prévision sur une année de la température d'Alger sur la base de données mensuelle.

Chapitre 1

Processus Autorégessif Hilbertien

1.1 Notations et définitions

Définition 1.1.1 Soit E un espace vectoriel sur \mathbb{K} ($\mathbb{K} = \mathbb{R}$ ou \mathbb{C}). Un produit scalaire sur E est une application, notée $\langle \cdot, \cdot \rangle$, de $E \times E$ à valeurs dans \mathbb{K} telle que pour tout $x, y, z \in E$ et $\alpha \in \mathbb{K}$, on ait :

1. $\langle x, x \rangle \geq 0$ et si $\langle x, x \rangle = 0 \iff x = 0$;
2. $\langle x, y \rangle = \overline{\langle y, x \rangle}$
3. $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$ et $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$

Dans le deuxième point, $\overline{\langle y, x \rangle}$ est le conjugué dans \mathbb{C} de $\langle y, x \rangle$. Les deux derniers points nous donnent $\langle x, \alpha y \rangle = \overline{\alpha} \langle x, y \rangle$. Lorsque $\mathbb{K} = \mathbb{R}$, ces barres de conjugaison sont inutiles.

Dans la suite de ce chapitre, on considérera le cas où $\mathbb{K} = \mathbb{C}$.

Définition 1.1.2 On appelle espace hermitien un espace vectoriel muni d'un produit scalaire.

On pose $\|x\| = \sqrt{\langle x, x \rangle}$.

Nous verrons que cette quantité est une norme sur un espace hermitien \mathcal{H} lorsque nous aurons énoncé les propriétés :

- Inégalité de Schwarz : pour tout $x, y \in H$, on a

$$|\langle x, y \rangle| \leq \|x\| \cdot \|y\|;$$

- Identité du parallélogramme :

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2).$$

Proposition 1.1.1 Un espace hermitien H est un espace vectoriel normé avec la norme $\|x\| = \sqrt{\langle x, x \rangle}$ induite par le produit scalaire.

Démonstration. Le seul axiome non évident à obtenir est l'inégalité triangulaire.

On a

$$\|x + y\|^2 = 2\|x\|^2 + 2\|y\|^2 + 2\operatorname{Re}\langle x, y \rangle.$$

De l'inégalité de Schwarz, on déduit que

$$2\operatorname{Re}\langle x, y \rangle \leq 2\|x\| \cdot \|y\|,$$

d'où

$$\|x + y\|^2 \leq (\|x\| + \|y\|)^2.$$

Définition 1.1.3 Lorsqu'un espace hermitien \mathcal{H} muni de la norme induite par le produit scalaire est complet, on dit que H est un espace de Hilbert.

Exemples 1.1.1 L'espace $L^2(\Omega, \mathcal{A}, \mathbb{P})$ de toutes les variables aléatoires complexes X définies sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ telles que

$$\mathbb{E}(|X|^2) < \infty,$$

muni du produit scalaire

$$\langle X, Y \rangle = \mathbb{E}(X\bar{Y}) = \int_{\Omega} X(\omega)\bar{Y}(\omega)d\mathbb{P}(\omega),$$

est un espace de Hilbert.

Définition 1.1.4 Opérateurs de Hilbert Shmidts.

T est un opérateur linéaire sur un espace de Hilbert séparable $(H, \|\cdot\|)$ est de Hilbert Schmidt si :

$$\|T\|_S^2 = \sum_{i=1}^{\infty} \|Te_i\|^2 = \sum_{i,j=1}^{\infty} \langle Te_i, e_j \rangle_H^2 < \infty$$

où $(e_i; i \in \mathbb{N})$ est une base Hilbertienne dans H . l'espace S des opérateurs aléatoires de Hilbert Schmidt sur H est un espace de Hilbert pour le produit $\langle \cdot, \cdot \rangle_S$ suivant : si $T_1, T_2 \in S$

$$\langle T_1, T_2 \rangle_S = \sum_{i=1}^{\infty} \langle T_1(e_i), T_2(e_i) \rangle$$

Les opérateurs $\Gamma(x) = E(\langle X_0, x \rangle X_0)$ et $\Delta(x) = E(\langle X_0, x \rangle X_1)$ sont des opérateurs de Hilbert-Schmidt.

Définition 1.1.5 On dit que l'espace de Hilbert H est l'espace de Sobolev W^d , des fonctions définies sur l'intervalle $[t_1, t_p]$ si :

$$W^d = \{f; f, f', \dots, f^{(d-1)}, \text{ sont absolument continus, } f^d \in L^2([t_1, t_p])\};$$

Par la suite on considèrera le cas particulier $d = 2$, très employé en pratique.

• **Lissage spline.**

Le mot anglais "spline" désigne une latte flexible utilisée par les dessinateurs pour matérialiser des lignes à courbure variable et passant par des points fixés a priori ou à "proximité" de ceux-ci. Le tracé ainsi réalisé minimise l'énergie de déformation de la latte. Par analogie, ce mot désigne également des familles de fonctions d'interpolation ou de lissage présentant des propriétés "optimales" de régularité. L'idée originale est attribuée à Whittaker(1923), puis reformulée par Schoenberg(1964), Atteia(1965) et Reinsch(1967). Elle connaît ses premières applications en Statistique avec Kimeldorf et Wahba(1970).

Les familles de splines constituent un ensemble d'outils d'approximation, de lissage et d'interpolation largement utilisés dans des domaines d'application des mathématiques très variés (problèmes de régularisation, d'équations intégrales, de restauration d'images,...).

Définition 1.1.6 *Pour un intervalle $[a, b]$, un entier $r \geq 1$, un entier k , et une suite de k points z_1, \dots, z_k dans $[a, b]$, on appelle spline polynomiale d'ordre r ayant pour noeuds simples les points z_1, \dots, z_k toute fonction f de $[a, b]$ dans \mathbb{R} telle que :*

- f est continument dérivable jusqu'à l'ordre $r - 2$.
- La restriction de f aux intervalles inter-noeuds :

$$[a, z_1], \dots, [z_i, z_{i+1}], \dots, [z_k, b].$$
- Coïncide avec un polynôme de degré inférieur ou égal à $r - 1$.
- Cet ensemble de fonctions sera noté $S_r(z_1, \dots, z_k)$. Il contient l'ensemble des polynômes de degré inférieur ou égal à $r - 1$.
- Pour $r = 2$, une spline d'ordre 2 est donc une fonction continue et linéaire par morceaux.
- Les splines les plus fréquemment utilisées sont les splines d'ordre 4 dites splines cubiques.

• **Les hypothèses de dépendance.**

Pour obtenir des estimateurs convergents il est nécessaire de faire des hypothèses d'indépendance asymptotique sur le processus observé.

Les hypothèses de mélangeance sont les plus utilisées. Nous allons énoncer les principales.

On considère un processus réel $(X_t, t \in I)$ où $I = \mathbb{Z}$ ou \mathbb{R} et on désigne par $\mathfrak{B}_{-\infty}^t$ (resp. $\mathfrak{B}_{t+k}^{+\infty}$) la tribu engendrée par $(X_s, s \leq t)$ (resp. $X_s, s \geq t + k$).

Définition 1.1.7 *On dit que (X_t) est ϕ -mélangeant si*

$$\sup_{A \in \mathfrak{B}_{-\infty}^t, B \in \mathfrak{B}_{t+k}^{+\infty}} |P(B | A) - P(B)| \leq \phi_k; \quad t \in I, \quad t + k \in I, \quad k \geq 0.$$

où $\lim_{k \rightarrow +\infty} \downarrow \phi_k = 0$ et avec la convention $P(B | A) = P(B)$ si $P(A) = 0$.

Définition 1.1.8 *On dit que (X_t) est α -mélangeant (ou fortement mélangeant) si*

$$\sup_{A \in \mathfrak{B}_{-\infty}^t, B \in \mathfrak{B}_{t+k}^{+\infty}} |P(A \cap B) - P(A)P(B)| \leq \alpha_k; \quad t \in I, \quad t + k \in I, \quad k \geq 0.$$

où $\lim_{k \rightarrow +\infty} \downarrow \alpha_k = 0$.

1.2 Estimation et Prévision

Définition 1.2.1 Soit $(X_n, n \in \mathbb{Z})$ un processus défini sur $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans un espace de Hilbert réel et séparable $(H, \mathcal{H}, \|\cdot\|)$ vérifiant :

$$(1.1) \quad X_{n+1} = \rho(X_n) + \epsilon_{n+1}, \quad \text{ou} \quad X_{n+1}^\omega = \rho X_n(\omega) + \epsilon_{n+1}(\omega); \quad n \in \mathbb{Z}, \quad \omega \in \Omega$$

où les variables $X_i; \epsilon_i$ sont en fait des éléments aléatoires à valeurs dans un espace de Hilbert séparable $(H, \langle \cdot, \cdot \rangle)$ et ρ est un opérateur linéaire continu de H vers H ; sur lequel on peut être amené à faire des hypothèses supplémentaires du type symétrie, appartenance à des sous-classes (comme les opérateurs compacts ou de Hilbert- Schmidt), etc. On suppose que les variables $(\epsilon_n)_{n \in \mathbb{Z}}$ constituent un H -bruit blanc. Elles sont centrées, indépendantes des $(X_j)_{j \leq n}$ constituent un processus d'innovation pour le processus $(X_n)_{n \in \mathbb{Z}}$ qui est par conséquent stationnaire.

Le processus X_n est ainsi centré et markovien.

Des exemples de processus $ARH(1)$ son présentés dans [1].

Il faut également mentionner le fait que la plupart des auteurs ayant étudié les propriétés théoriques des autorégressifs fonctionnels ou leurs généralisations s'appuient sur la version vectorielle du modèle (1.1) (délaissant l'optique traditionnelle des processus à temps continu c'est à dire une écriture du type

$$X_{n+1}(t) = [\rho(X_n)](t) + \epsilon_{n+1}(t) \quad n \in \mathbb{Z}; \quad t \in T$$

qui semble peu pratique dans ce contexte).

Si φ est un vecteur propre de ρ associé à λ alors $(\langle X_i, \varphi \rangle, i \in \mathbb{Z})$ est un $AR(1)$.

Pour les variables aléatoires Hilbertiennes, les opérateurs de covariance et de covariance croisée comme des opérateurs linéaires sur H sont définis par :

$$C_{x_0}(x) = \mathbb{E}(\langle X_0, x \rangle X_0)$$

$$C_{X_0 X_1}(x) = \mathbb{E}(\langle X_0, x \rangle X_1)$$

pour un $ARH(1)$ on a $C_{X_0 X_1} = \rho C_{X_0}$

Proposition 1.2.1 Si (X_i) est un $ARH(1)$ associé à ρ et ϵ_i on a :

$$\mathbb{E}(X_0) = 0$$

$$\begin{aligned}
C_{X_0} &= \rho C_{X_0} \rho + C_\varepsilon \\
&= \sum_{j=0}^{\infty} \rho^j C_\varepsilon \rho^j \\
C_{X_0 X_{-k}} &= C_{X_0} \rho^k, \quad k = 1, 2, \dots \\
C_{X_{-k}, X_0} &= \rho^k C_{X_0}, \quad k = 1, 2, \dots
\end{aligned}$$

Dans la suite on note

$$\Gamma = C_{X_0}, \quad \Delta^* = C_{X_1 X_0}$$

Les deux dernières équations de la proposition s'écrivent donc

$$\Delta^* = \Gamma_\rho \quad \text{et} \quad \Delta = \rho \Gamma$$

Comme cela a été signalé dans le livre de [6], l'estimation de Γ est un problème difficile même si l'existence de l'inverse de Γ entraîne $\rho = \Gamma^{-1} \Delta$. Cela est dû au fait que H étant de dimension infinie Γ^{-1} n'est pas un opérateur borné.

1.2.1 Estimation des opérateurs de covariance Γ et Δ

En fait si X et Y sont des variables aléatoires Hilbertiennes du second ordre, alors C_{XY} est un opérateur de Hilbert Schmidt car

$$\begin{aligned}
\sum_j \|C_{XY}(\varphi_j)\|^2 &= \sum_j \sum_l \langle C_{XY}(\varphi_j), \varphi_l \rangle^2 \\
&= \sum_j \sum_l (E\langle X, \varphi_j \rangle \langle Y, \varphi_l \rangle)^2 \\
&\leq \sum_j \sum_l E\langle X, \varphi_j \rangle^2 E\langle Y, \varphi_l \rangle^2 \\
&\leq E\|X\|^2 E\|Y\|^2 < \infty
\end{aligned}$$

où (φ_j) est une base orthonormée de H . Leurs normes de Hilbert Schmidt sont respectivement :

$$\begin{aligned}
\|\Gamma\|_S &= \left(\sum_{j,l} |E(\langle X_0, \varphi_j \rangle \langle X_0, \varphi_l \rangle)|^2 \right)^{\frac{1}{2}} \\
\|\Delta\|_S &= \left(\sum_{j,l} |E(\langle X_0, \varphi_j \rangle \langle X_1, \varphi_l \rangle)|^2 \right)^{\frac{1}{2}}.
\end{aligned}$$

Un estimateur naturel de Γ est l'opérateur de la covariance empirique, noté Γ_n

$$(1.2) \quad \Gamma_n(x) = \frac{1}{n} \sum_{i=1}^n \langle X_i, x \rangle X_i, \quad x \in H;$$

Cet opérateur est de rang fini et donc de Hilbert Schmidt. De façon analogue on définit un estimateur de Δ par :

$$(1.3) \quad \Delta_n(x) = \frac{1}{n-1} \sum_{i=1}^{n-1} \langle X_i, x \rangle X_{i+1}, \quad x \in H.$$

On l'appelle opérateur de la covariance croisée empirique de X_0, X_1 . Il est de Hilbert Schmidt. On montre que Γ_n est un estimateur symétrique sans biais de Γ . La proposition suivante donne la convergence p.s de Γ_n vers Γ et une vitesse de convergence de l'ordre $\frac{1}{n}$.

Proposition 1.2.2 *Supposons que $\mathbb{E}\|X_0\|^4 < \infty$. Alors :*

$$(1.4) \quad E\|\Gamma_n - \Gamma\|_S^2 \leq \frac{A}{n}$$

$$\text{où } A = \mathbb{E}\|X_0\|^4 \left[1 + 4\|\rho\|^2 (1 - \|\rho\|^2)^{-1} \right]$$

$$\|\Gamma_n - \Gamma\|_S \rightarrow 0 \quad p.s$$

Proposition 1.2.3 *Pour $n \geq 2$*

$$(1.5) \quad \mathbb{E}\|\Delta_n - \Delta\|_S^2 = E\|\Delta_n^*\|_S^2 \leq \frac{B}{n-1}$$

$$\text{où } B = 2\mathbb{E}\|X_0\|^2 E\|\varepsilon_0\|^2 + 2A \text{ et } A \text{ est défini dans la proposition (1.2.2).}$$

De plus on a :

$$\|\Delta_n - \Delta\|_S = \|\Delta_n^* - \Delta^*\|_S \rightarrow 0 \quad p.s.$$

1.2.2 Estimation des éléments propres de Γ

Soient $\lambda_{1n} > \lambda_{2n} > \dots > \lambda_{nn} \geq 0$ les valeurs propres de Γ_n et $\varphi_{1n}, \varphi_{2n}, \dots$ le système complet de vecteurs propres de Γ_n tels que :

$$\Gamma_n \varphi_{jn} = \lambda_{jn} \varphi_{jn}, \quad j = 1, 2, \dots$$

Alors des estimateurs naturels de λ_j et φ_j sont respectivement λ_{jn} et φ_{jn} .

Lemme 1.2.1 *Pour tout entier positif j nous avons :*

$$(1.6) \quad |\lambda_{jn} - \lambda_j| \leq \|\Gamma_n - \Gamma\|.$$

et si $\lambda_1 > \lambda_2$

$$(1.7) \quad \|\varphi_{1n} - \varphi'_1\| \leq a_1 \|\Gamma_n - \Gamma\|$$

où $a_1 = 2\sqrt{2}(\lambda_1 - \lambda_2)^{-1}$ et si pour un $j > 1$, $\lambda_{j-1} > \lambda_j > \lambda_{j+1}$ alors :

$$(1.8) \quad \|\varphi_{jn} - \varphi'_j\| \leq a_j \|\Gamma_n - \Gamma\|,$$

où $a_j = 2\sqrt{2}[\min(\lambda_{j-1} - \lambda_j, \lambda_j - \lambda_{j+1})]^{-1}$. Ici $\varphi'_j = (sg < \varphi_{jn}, \varphi_j >) \varphi_j$, $j \geq 1$,

$$\text{où } sg(x) = \begin{cases} 1, & \text{si } x \geq 0; \\ -1, & \text{sinon.} \end{cases}$$

Proposition 1.2.4 *Si $\|X_0\| \leq d$ p.s, (X_i) est géométriquement α -mélangeant et*

$$\lambda_i = ar^j \quad ; a > 0, \quad 0 < r < 1; \quad j = 1, 2, \dots$$

alors pour $0 < \gamma < \frac{1}{4}$, nous avons pour tout $\varepsilon > 0$ et $n > n_\gamma$

$$(1.9) \quad P(\|\Gamma_n - \Gamma\|_S > \varepsilon) \leq 2n^\gamma \exp\left(-\frac{\varepsilon n^\gamma}{10d^2\varepsilon\sqrt{2}}\right) + \frac{2ad^2r}{\varepsilon^2(1-\sqrt{r})^2} \exp\left(-\log\frac{1}{r}n^\gamma\right)$$

et

$$(1.10) \quad P(\|\Delta_n - \Delta\|_S > \varepsilon) \leq 2n^\gamma \exp\left(-\frac{\varepsilon(n-1)^\gamma}{10(d^2 + \|\Delta\|_S)\varepsilon\sqrt{2}}\right) + \frac{2ad^2r}{\varepsilon^2(1-\sqrt{r})^2} \exp\left(-\log\frac{1}{r}(n-1)^\gamma\right)$$

Proposition 1.2.5 *Nous avons*

$$(1.11) \quad E\left(\sup_{j \geq 1} |\lambda_{jn} - \lambda_j|^2\right) \leq \frac{A}{n}$$

$$\sup_{j \geq 1} |\lambda_{jn} - \lambda_j| \rightarrow 0 \quad \text{p.s}$$

et sous les hypothèses de la proposition (1.2.4) la vitesse de convergence est de l'ordre $O(\exp(-cn^\gamma))$;

$0 < \gamma < \frac{1}{4}$

Proposition 1.2.6 *pour tout j tel que $\lambda_{j-1} > \lambda_j > \lambda_{j+1}$ alors on a*

$$E\|\varphi_{jn} - \varphi'_j\| \leq a_j^2 \frac{A}{n}$$

où A est défini dans Prop (1.3.2) est $\|\varphi_{j_n} - \varphi'_j\|^2 \rightarrow 0$ p.s et la vitesse de convergence est $O(\exp(-cn^\gamma))$, où $0 < \gamma < \frac{1}{4}$.

Proposition 1.2.7 Pour $n \geq 2$

$$(1.12) \quad E\|\Delta_n - \Delta\|_S^2 = E\|\Delta_n^*\|_S^2 \leq \frac{B}{n-1}$$

où $B = 2E\|X_0\|^2 E\|\varepsilon_0\|^2 + 2A$ et défini dans la proposition (1.3.2) De plus on a :

$$\|\Delta_n - \Delta\|_S = \|\Delta_n^* - \Delta^*\|_S \rightarrow 0 \quad p.s$$

1.2.3 Estimation de ρ et Prévision de X_{n+1}

Passons maintenant à l'estimation de ρ et la prévision de X_{n+1} en considérant d'abord le cas où les vecteurs propres de l'opérateur Γ sont connus.

Estimation et prévision dans le cas général

Soit $\hat{\Pi}^{k_n}$ le projecteur orthogonal sur le sous-espace vectoriel engendré par les k_n premiers vecteurs propres $(\phi_{j_n}, j = 1, \dots, k_n)$ de Γ_n et $\hat{Y}_{in} = \hat{\Pi}^{k_n}(X_i), i = 1, \dots, n$.

Les opérateurs de la covariance de \hat{Y}_{0n} et de la covariance croisée de $\hat{Y}_{0n}, \hat{Y}_{1n}$ sont :

$$\hat{G}_n = \hat{\Pi}^{k_n} \Gamma \hat{\Pi}^{k_n}, \quad \hat{H}_n = \hat{\Pi}^{k_n} \Delta \hat{\Pi}^{k_n}, \quad \hat{H}_n^* = \hat{\Pi}^{k_n} \Delta^* \hat{\Pi}^{k_n},$$

et les opérateurs respectifs de la covariance empirique :

$$\hat{C}_n = \hat{\Pi}^{k_n} \Gamma_n \hat{\Pi}^{k_n}, \quad \hat{D}_n = \hat{\Pi}^{k_n} \Delta_n \hat{\Pi}^{k_n}, \quad \hat{D}_n^* = \hat{\Pi}^{k_n} \Delta_n^* \hat{\Pi}^{k_n}$$

Pour construire un estimateur de $\hat{\rho}$ nous ajoutons aux hypothèses de la proposition précédente, l'hypothèse suivante :

H_2^* Le déterminant de $(\sum_{i=1}^n \langle X_i, \phi_{j_n} \rangle \langle X_i, \phi_{l_n} \rangle)_{1 \leq j_n, l_n \leq k_n}$ ne s'annule pas P.s.

Lemme 1.2.2 Sous H_1 et H_2 pour tout $\varepsilon > 0$

$$P(\|\rho_n - \Pi^{k_n} \rho \Pi^{k_n}\| > \varepsilon) \leq \frac{C_\varepsilon}{n\lambda_{k_n}^4}$$

$$C_\varepsilon = 2B\lambda_1^4 + \left(A + \frac{8B}{\varepsilon^2}\right) \lambda_1^2 + \frac{4A(\|\Delta\| + 1)}{\varepsilon^2} \left(1 + \lambda_1 \frac{\varepsilon}{2}\right)^2$$

Proposition 1.2.8 Sous H_1, H_2 et si k_n et $n\lambda_{k_n}^4$ tendent vers l'infini alors :

$$\|\rho_n(X_n) - \rho(X_n)\| \rightarrow 0$$

Chapitre 2

La méthode du noyau

2.1 Rappels

2.1.1 Les variables fonctionnelles

Avant d'aller plus loin, qu'appelle-t-on précisément variable aléatoire fonctionnelle. On appelle variable aléatoire fonctionnelle (*v.a.f.*), notée X , toute application mesurable définie sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ et à valeurs dans un espace probabilisable (E, \mathcal{B}_E) où E est un espace de dimension infinie, appelé génériquement espace fonctionnel, \mathcal{B}_E étant la tribu des boréliens de E (tribu engendrée par les ouverts de E). Au delà de l'espace fonctionnel E dans lequel X prend ses valeurs, la terminologie *fonctionnelle* fait essentiellement référence à la dimension infinie de E .

En effet, la principale difficulté dans l'étude des modèles pour *v.a.f.* provient justement de cette dimension infinie. L'idée consiste alors à généraliser les modèles classiques pour variable aléatoire vectorielle en se plaçant dans un cadre purement fonctionnel. Le cas où l'espace E est de dimension finie se ramène naturellement au cadre vectoriel ci-dessus. Par conséquent, nous nous limiterons de manière implicite à des espaces E de dimension infinie, et nous parlerons d'espace fonctionnel.

La fonction r est une fonction de E dans \mathbb{R} , et nos modèles statistiques s'écrivent :

$$r \in \mathcal{C}$$

où \mathcal{C} est une classe de fonctions.

Dans notre cas, E est un espace de Hilbert.

2.1.2 La méthode du noyau

Les estimateurs de type noyau sont introduits indépendamment par Nadaraya (1964) et Watson (1964). Pour comprendre les idées qui ont amené à l'introduction de ces estimateurs, peut-être faut-il remonter

au régressogramme de *Tukey* (1961) défini de la manière suivante :

$$(2.1) \quad \hat{r}_{reg}(x) = \frac{\sum_{i=1}^n Y_i \mathbb{I}(X_i \in \mathcal{B}_j)}{\sum_{i=1}^n \mathbb{I}(X_i \in \mathcal{B}_j)}, \quad \forall x \in \mathcal{B}_j,$$

où $\mathcal{B}_j, j = 1, \dots, J$ est une partition du support de X fixée a priori. Cet estimateur primitif présente comme inconvénient d'avoir à choisir à la fois la finesse de la discrétisation (*i.e.* le nombre J de découpages) et la position exacte des bornes des intervalles \mathcal{B}_j .

Afin de résoudre ce second problème, un nouvel estimateur peut être construit en remplaçant la discrétisation a priori en intervalles \mathcal{B}_j par un seul intervalle mais qui varie de manière continue. Concrètement, cela donne l'estimateur de la fenêtre mobile défini de la manière suivante :

$$(2.2) \quad \hat{r}_{FM}(x) = \frac{\sum_{i=1}^n Y_i \mathbb{I}(X_i \in [x-h; x+h])}{\sum_{i=1}^n \mathbb{I}(X_i \in [x-h; x+h])}, \quad \forall x$$

où h est un paramètre réel strictement positif.

L'estimateur précédent présente encore le désavantage d'être discontinu par nature. Ainsi sa généralisation naturelle est l'estimateur à noyau, appelé aussi estimateur de Nadaraya-Watson, défini de la manière suivante :

$$(2.3) \quad \hat{r}_{NW} = \frac{\sum_{i=1}^n Y_i K\left(\frac{\mathbf{x}-X_i}{h}\right)}{\sum_{i=1}^n K\left(\frac{\mathbf{x}-X_i}{h}\right)}, \quad \forall x$$

Dans cette définition K est une fonction de \mathbb{R} dans \mathbb{R} (dont nous verrons qu'elle n'est pas nécessairement positive), et h est un paramètre réel strictement positif (dont nous verrons qu'il sera intéressant de le faire dépendre de n).

2.2 Régression scalaire par noyau

La prévision non-paramétrique d'un processus réel à partir de son r -historique (ses r valeurs précédentes) est une approche possible pour l'étude de la série $(X_k)_{k \in \mathbb{Z}}$. Notons

$$\mathbf{X}_t^r = \{X_t, X_{t-1}, \dots, X_{t-r+1}\} \in \mathbb{R}^r$$

le vecteur r -historique et s l'horizon de prévision ($0 < s \leq r$).

L'autorégression à horizon s à partir de l'historique $(X_k)_{k=1, \dots, r}$ est définie par

$$f_s(x) = \mathbb{E}(X_{T+s} \mid \mathbf{X}_t^r = x)$$

L'estimateur à noyau de f_s à partir des observations (x_1, \dots, x_T) est alors

$$(2.4) \quad f_{(T,s)}(\mathbf{x}) = \frac{\sum_{t=r}^{T-s} x_{t+s} K\left(\frac{\mathbf{x}-\mathbf{x}_t^r}{h_T}\right)}{\sum_{t=r}^{T-s} K\left(\frac{\mathbf{x}-\mathbf{x}_t^r}{h_T}\right)}$$

où h_T est la taille de la fenêtre, et K est un noyau de dimension r . La pratique montre que le choix du noyau n'est pas crucial dans ce type de modèle (cf. Bosq, 1998, [9]), prenons donc pour K le noyau gaussien

$$K(\mathbf{x}) = (2\pi)^{-r/2} \exp\left(-\frac{\|\mathbf{x}\|^2}{2}\right), \quad \mathbf{x} \in \mathbb{R}^r$$

La prévision à l'horizon s est alors donnée par

$$\hat{X}_{T+s|T} = f_{T,s}(\mathbf{x}_T^r)$$

Le choix de la valeur h_T est déterminé par validation croisée, selon

$$h_T = \arg \min \sum_{s=1}^p CV_s(h)$$

avec

$$(2.5) \quad CV_s(h) = \sum_{k=r, \dots, T-pm-s} (x_{k+s} - f_{s,h,-k}(\mathbf{x}_{k(r)}))^2$$

où $f_{s,h,-k}$ est la fonction d'autorégression à l'horizon s obtenue à partir des observations $\{x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_{T-mp}\}$ et m est le nombre de périodes gardées afin de tester la qualité de prévision.

Cet estimateur dispose de bonnes propriétés asymptotiques avec une vitesse de convergence optimale sous des hypothèses de stationnarité, de mélange, de propriétés de Markov ou ergodicité.

2.2.1 Etude de la convergence presque complète

Commençons par donner les hypothèses communes aux résultats asymptotiques auxquels nous allons nous intéresser :

- supposons que pour un point fixé $x \in H$, on ait :

$$(2.6) \quad \lim_{\alpha \rightarrow 0^+} \frac{\mathbb{P}(X \in \mathcal{B}(x, \alpha))}{\alpha^{\delta(x)}} = c(x)$$

où $\delta(x)$ et $c(x)$ sont deux réels strictement positifs et $\mathcal{B}(x, \alpha)$ désigne la boule de centre x et de rayon α pour la topologie associée à la semi-norme $\|\cdot\|$.

- nous devons renforcer l'hypothèse (2.6) par l'existence d'un réel $b(x)$ strictement positif tel que pour α au voisinage de 0 on ait :

$$(2.7) \quad \mathbb{P}(Y \in \mathcal{B}(x, \alpha)) = \alpha^{\delta(x)} c(x) + O(\alpha^{\delta(x)+b(x)}).$$

- la largeur de fenêtre h_n est telle que :

$$(2.8) \quad \lim_{n \rightarrow \infty} h_n = 0 \text{ et } \lim_{n \rightarrow \infty} \frac{nh_n^{\delta(x)}}{\log n} = \infty,$$

- le noyau K est tel que

$$(2.9) \quad K \text{ Lipschitzien d'ordre 1 et support}(K) = [0, \xi] \text{ avec } \xi \in \mathbb{R}_*^+,$$

- la variable aléatoire réelle X vérifie :

$$(2.10) \quad |X| \leq M < \infty \text{ p.s.}$$

Théorème 2.2.1 *Convergence presque complète ponctuelle.*

Soit x un point fixé de H . Sous les hypothèses (2.6), et (2.8) – (2.10) et si

$$(2.11) \quad \rho \text{ est continu en } x,$$

alors on a :

$$(2.12) \quad \lim_{n \rightarrow \infty} \hat{\rho}(x) = \rho(x), \text{ p.co.}$$

Démonstration. Posons

$$\hat{f}(x) = \frac{1}{nh\delta(x)} \sum_{i=1}^{n-1} K\left(\frac{\|X_i - x\|}{h_n}\right) \text{ et } \hat{g}(x) = \frac{1}{nh\delta(x)} \sum_{i=1}^{n-1} Y_i K\left(\frac{\|X_i - x\|}{h_n}\right)$$

de sorte que

$$\hat{\rho}(x) = \frac{\hat{f}(x)}{\hat{g}(x)}.$$

La démonstration de ce résultat est alors basée sur la décomposition suivante :

$$(2.13) \quad \hat{\rho}(x) - \rho(x) = \frac{1}{\hat{f}(x)} \{\hat{g}(x) - \rho(x)C_\delta(x)\}$$

où $C_\delta(x)$ est un réel strictement positif (dépendant de x ainsi que de la dimension fractale $\delta(x)$) qui sera précisé en cours de démonstration sur les lemmes suivants :

Lemme 2.2.1 *Sous les conditions (2.8) et (2.9), et si :*

– (2.6) est vérifiée, alors on a :

$$(2.14) \quad \lim_{n \rightarrow \infty} \mathbb{E}(\hat{f}(x)) = C_\delta(x).$$

– (2.7) est vérifiée, alors on a :

$$(2.15) \quad \mathbb{E}(\hat{f}(x)) = C_\delta(x) + O(h^{b(x)}).$$

Lemme 2.2.2 *Sous les conditions (2.6), (2.8) et (2.9) on a :*

$$(2.16) \quad \hat{f}(x) - \mathbb{E}(\hat{f}(x)) = O\left(\sqrt{\frac{\log n}{nh^{\delta(x)}}}\right) \text{ p.co.}$$

Lemme 2.2.3 *Sous les conditions (2.6), (2.8), (2.9) et (2.11) on a :*

$$(2.17) \quad \lim_{n \rightarrow \infty} \mathbb{E}(\hat{g}(x)) = \rho(x)C_\delta(x).$$

Lemme 2.2.4 *Sous les conditions (2.6), (2.8), (2.9) et (2.10) on a :*

$$(2.18) \quad \hat{g}(x) - \mathbb{E}(\hat{g}(x)) = O\left(\sqrt{\frac{\log n}{nh^{\delta(x)}}}\right) \text{ p.co.}$$

2.3 Régression fonctionnelle par noyau

2.3.1 Présentation du modèle

Pour prédiction d'un $ARH(1)$, on cherche à estimer ρ , et les méthodes proposées rendent nécessaire l'estimation et l'inversion de l'opérateur Hilbertien de covariance défini ci-dessus. La difficulté de la démarche se situe dans cette inversion car l'opérateur de covariance est à priori non borné. Les différentes méthodes d'estimation à celui associé aux valeurs propres les plus fortes de l'opérateur. La méthode que nous présentons ici, proposée par Besse et Cardot [17], anticipe cette réduction de dimension de l'espace d'estimation.

2.3.2 Présentation de l'estimateur à noyau fonctionnel

On considère maintenant l'approche fonctionnelle, les observations sont issues d'un processus Hilbertien du second ordre $(Y_i)_{i \in \mathbb{Z}}$ Markovien. Afin de considérer des fonctions 'lisses', on se place sous l'hypothèse que l'espace de Hilbertien est l'espace de Sobolev W^2 des fonctions définies sur $[t_1, t_p]$.

En pratique, partant des observations discrètes $\{y_i(t_j) = x_{(i-1)p+j}; i = 1, \dots, n; j = 1, \dots, p\}$, on approxime les courbes y_i par les interpolations \tilde{y}_i par splines obtenues suivant

$$(2.19) \quad \tilde{y}_i = \arg \min \|D^2 \tilde{y}_i\|_{L^2}^2$$

sous la contrainte

$$\tilde{y}_i(t_j) = y_i(t_j), \quad j = 1, \dots, p, \text{ pour } i = 1, \dots, n$$

où D est l'opérateur de différenciation et $\|\cdot\|_{L^2}$ est la norme de $L^2([t_1, t_p])$. Dans ce cas, les fonctions splines employées sont des polynômes de degré 3 par morceaux.

Une méthode classique permettant d'approximer ce type de fonction par minimisation d'une norme de W^2 est l'interpolation par les fonctions splines. Naturellement, si les hypothèses de régularité précédentes ne s'appliquent pas à l'objet d'étude, il peut être plus judicieux d'employer d'autre type d'interpolation (telles que linéaire ou par ondelettes).

Bosq (1983,[7]) propose de réaliser la prédiction de des processus $ARH(1)$ via l'espérance conditionnelle $\rho(x) = \mathbb{E}(X_{n+1}|X_n = x)$. L'opérateur ρ ne peut être estimé aussi au moyen d'une régression non paramétrique par noyau.

On réalise alors l'approximation de ρ par l'estimateur à noyau

$$(2.20) \quad \hat{\rho}_{h_n}(y) = \sum_{i=1}^{n-1} w_i(y) \tilde{y}_{i+1}$$

avec

$$(2.21) \quad w_i(y) = \frac{K\left(\frac{\|\tilde{y}_i - y\|_{L^2}}{h_n}\right)}{\sum_{i=1}^{n-1} K\left(\frac{\|\tilde{y}_i - y\|_{L^2}}{h_n}\right)}$$

qui fournit la prédiction

$$\hat{y}_{n+1} = \hat{\rho}_{h_n}(y_n)$$

h_n étant une suite de nombres positifs. Il s'agit d'une version de l'estimateur de Nadaraya-Watson adaptée au cas où la variable explicative est à valeurs dans H . Son principe est le suivant : étant donné un élément y de H , on détermine $\hat{\rho}_{h_n}(y)$ en moyennant les \tilde{y}_{i+1} selon une pondération qui accorde plus de poids aux \tilde{y}_{i+1} pour lesquels les \tilde{y}_i sont "proches" de y :

- la semi-norme $\|\cdot\|$ mesure la proximité entre les courbes,
- la largeur de fenêtre h contrôle le nombre de termes retenus dans le calcul de la moyenne pondérée,

Remarque 2.3.1 *il est nécessaire de préciser pourquoi nous préférons utiliser une semi-norme plutôt qu'une norme. Le premier argument est d'ordre théorique : tout étude réalisée dans le cadre d'un espace vectoriel semi-normé s'applique immédiatement au cas d'un espace vectoriel normé. Le second argument est de nature heuristique. En effet, gardons à l'esprit que les observations d'une v.a.f. peuvent se présenter, par exemple, sous la forme de courbes. Comme nous le verrons par la suite, nous avons besoin d'un indice qui mesure la proximité entre ces différentes courbes. Or, l'utilisation de norme peut s'avérer catastrophique dans certaines situations, notamment dans le traitement de courbes spectrométriques (Ferraty et Vieu, 2001,[10]).*

- le noyau K est une fonction décroissante : plus \tilde{y}_i est proche de y c'est-à-dire plus la quantité $\|\tilde{y}_i - y\|$ est petite, plus le poids $w_i(y)$ est grand.

La valeur de h_n est obtenue par validation croisée

$$h_n = \arg \min CV(h)$$

avec

$$CV(h) = \sum_{k=n-m-r}^{n-m-1} \|\hat{\rho}_{h,n-m-r}(\tilde{y}_k) - \tilde{y}_{k+1}\|_{L^2}^2$$

où $\hat{\rho}_{h,n-m-r}$ est l'estimation obtenue à partir des observations $\{y_i; i = 1, \dots, n-r-m\}$. Les m dernières courbes sont gardées afin de tester la qualité de la prédiction.

Intéressons-nous maintenant à la convergence presque complète uniforme sur un compact S de l'espace semi-normé $(H, \|\cdot\|)$. Il est clair que l'obtention de résultats uniformes nécessite le renforcement de certaines hypothèses utilisées dans le paragraphe précédent. En particulier, on rend uniforme l'hypothèse (2.6) de la manière suivante :

$$(2.22) \quad \lim_{\alpha \rightarrow 0^+} \sup_{y \in S} \left\{ \frac{\mathbb{P}(Y \in \mathcal{B}(y, \alpha))}{\alpha^{\delta(y)}} - c(y) \right\} = 0 \quad \text{et} \quad \inf_{y \in S} c(y) > 0,$$

où δ est un réel strictement positif ne dépendant pas de y . Notons que dans ce cas, cette hypothèse peut s'interpréter en termes de dimension de Hausdorff de la loi de la v.a.f. Y (Bardet, 1997, p.28, [2]). On peut alors énoncer le résultat suivant.

Théorème 2.3.1 *Convergence presque complète uniforme.*

Sous les hypothèses (2.8), (2.10) et (2.22),

$$(2.23) \quad \rho \text{ est uniformment continu sur } S,$$

alors on a :

$$(2.24) \quad \lim_{\alpha \rightarrow 0^+} \sup_{y \in S} |\hat{\rho}_{h_n}(y) - \rho(y)| = 0, \quad p.co.$$

Démonstration. La démonstration de ce résultat est en pratique est en partie analogue à celle du Théorème 2.2.1. En effet, la décomposition (2.13) s'écrit maintenant de la façon suivante :

$$(2.25) \quad \hat{\rho}_{h_n}(y) - \rho(y) = \frac{1}{\hat{f}(y)} \{\hat{g}(y) - \rho(y)C'_\delta(y)\} - \frac{\rho(y)}{\hat{f}(y)} \{\hat{f}(y) - C'_\delta(y)\},$$

où $C'_\delta(y) = c(u)\delta \int_0^\xi K(v)v^{\delta-1}dv$ est une constante positive. En reprenant alors les calculs faits précédemment, on montre que :

$$(2.26) \quad \lim_{n \rightarrow \infty} \mathbb{E}(\hat{f}(y)) = C'_\delta(y) \text{ uniformment en } y,$$

ce qui implique que

$$(2.27) \quad \lim_{n \rightarrow \infty} \sup_{y \in S} |\mathbb{E}(\hat{f}(y)) - C'_\delta(y)| = 0.$$

Ainsi que la démonstration de ce théorème se repose sur le lemme suivant :

Lemme 2.3.1 *Sous les conditions du Théorème 2.2.1, on a :*

$$\lim_{n \rightarrow \infty} \sup_{y \in S} |\mathbb{E}(\hat{g}(y)) - r(y)C'_\delta(y)| = 0.$$

2.4 Prédiction du modèles autorégressifs Hilbertiens

comme déjà vus dans le 1^{ière} chapitre un processus $(Y_i, i \in \mathbb{Z})$ est un $ARH(1)$ s'il est défini par l'équation (1.1)

2.5 Ecritures matricielles

2.5.1 Spline de lissage

Rappelons la propriété suivante : soient des opérateurs $\Gamma = \mathbb{E}(Y_i \otimes Y_i)$ et $\Delta = \mathbb{E}(Y_i \otimes Y_{i+1})$ sont les opérateurs de covariance et de covariance retardée du processus :

$$\rho\Gamma = \Delta.$$

qui permet l'estimer ρ , en inversant Γ , à partir des versions empiriques de Γ et Δ

$$\hat{\Gamma} = \frac{1}{n} \sum_{i=1}^n y_i \otimes y_i \quad \text{et} \quad \hat{\Delta} = \frac{1}{n-1} \sum_{i=1}^{n-1} y_i \otimes y_{i+1}$$

Notons par $\langle \cdot, \cdot \rangle_H$ le produit scalaire de H et notons $x \otimes y$ l'opérateur de rang 1 tel que

$$\forall (x, y, z) \in H \times H \times H, \quad [x \otimes y](z) = \langle x, z \rangle_H y.$$

De même que pour le modèle précédent, les observations des $(Y_i)_{i \in \mathbb{Z}}$ étant discrètes, il faut approximer les courbes. Comme par ailleurs, l'inversion de Γ demande une projection dans un sous espace, Besse et Cardot (1996)[17] proposent l'approximation spline en contraignant les fonctions splines en terme de lissage (paramètre l) et de rang (paramètre q), ce qui revient à projeter dans un sous-espace de dimension q en assurant une approximation optimale par fonctions splines. Pour cela, on effectue la minimisation suivante

$$(2.28) \quad \min_{\hat{y}_i \in H_q} \left[\frac{1}{n} \sum_{i=1}^n \left(\frac{1}{P} \sum_{j=1}^p (y_i(t_j) - \hat{y}_i(t_j)) \right)^2 + l \|D^2 \hat{y}_i\|_{L^2}^2 \right].$$

H_q est un sous-espace de H de dimension q à estimer.

La première étape d'estimation fonctionnelle consiste à associer à nos observations discrètes des fonctions dans un espace continu. Nous avons choisi ici de travailler dans l'espace de Sobolov W^2 avec une représentation linéaire ou par des fonctions splines choisies est une base de noyaux reproduisants [cf.(Wahba1990,[13])].

Soit Y la matrice des observations $y_i(t_j)$ de taille $(n \times p)$ de vecteurs lignes y_i . On note \mathbf{M} et \mathbf{N} les matrices associées respectivement à la projection de la norme de $L^2([t_1, t_p])$ et la semi-norme W^2 dans S_p , c'est-à-dire telles que

$$\|\tilde{y}_i\|_{L^2}^2 = \mathbf{y}_i' \mathbf{M} \mathbf{y}_i = \|\mathbf{y}_i\|_{\mathbf{M}}^2$$

$$\|D^2 \bar{y}_i\|_{L^2}^2 = \mathbf{y}_i' \mathbf{N} \mathbf{y}_i = \|\mathbf{y}_i\|_{\mathbf{N}}^2$$

où \bar{y}_i est l'approximation spline dans y_i dans S_p obtenues par une minimisation similaire à celle de la forme (2.28).

notons \mathbf{A}_l la matrice de lissage définie par

$$\mathbf{A}_l = (\mathbf{I} + l\mathbf{N})^{-1}$$

2.5.2 Régression fonctionnelle par noyau

Dans le cadre de l'interpolation spline (qui est un lissage spline avec $l = 0$), la formule (2.20) du prédicteur par régression fonctionnelle s'écrit

$$(2.29) \quad \rho_{h_n}(\mathbf{y}) = \frac{\sum_{i=1}^{n-1} k((\mathbf{y}_i - \mathbf{y})' \mathbf{M} (\mathbf{y}_i - \mathbf{y}) / h_n)}{\sum_{i=1}^{n-1} k((\mathbf{y}_i - \mathbf{y})' \mathbf{M} (\mathbf{y}_i - \mathbf{y}) / h_n)}$$

et sa prédiction est

$$\hat{y}_{n+1} = \hat{\rho}_{h_n}(\mathbf{y}_n)$$

2.5.3 Estimation d'un ARH(1)

Soit \mathbf{U} la matrice des trajectoires centrées de vecteurs lignes u_i ($u_i = y_i - \frac{1}{n} \sum_{j=1}^n y_j$), et S la matrice de covariance lissée donnée par

$$S = \frac{1}{n} \mathbf{A}(l)^{1/2} \mathbf{U}' \mathbf{U} \mathbf{A}(l)^{1/2}$$

dont les éléments propres classés par ordre décroissant des valeurs propres sont notés $(\lambda_1, v_1), \dots, (\lambda_p, v_p)$. Soit \mathbf{V}_q la matrice orthogonale contenant les q vecteurs propres aux q premières valeurs propres de S . Alors

$$\hat{y}_i = \mathbf{A}(l)^{1/2} \mathbf{V}_q \mathbf{V}_q' \mathbf{A}(l)^{1/2} \mathbf{y}_i \quad , \quad i = 1, \dots, n$$

Enfin, on estime les opérateurs de covariance par

$$\hat{\Gamma}_{q,l} = \frac{1}{n} \sum_{i=1}^n \hat{y}_i \hat{y}_i' \mathbf{M},$$

$$\hat{\Delta}_{q,l} = \frac{1}{n-1} \sum_{i=1}^{n-1} \hat{y}_i \hat{y}_{i+1}' \mathbf{M}.$$

d'où

$$\hat{\rho}_{q,l} = \hat{\Delta}_{q,l} \hat{\Gamma}_{q,l}^{-1}$$

et la prévision s'écrit

$$\hat{y}_{i+1} = \hat{\rho}_{q,l} \mathbf{y}_n + \mathbf{A}_l \bar{\mathbf{y}}.$$

2.5.4 ARH(1) Local

Dans bien des situations l'hypothèse de stationnarité du processus fonctionnel $(Y_i)_{i \in \mathbb{Z}}$ est une hypothèse trop forte. C'est par exemple le cas pour la série *El Niño* où on observe certains événements exceptionnels. Besse et coll ont proposé des approximations locales des opérateurs de covariance permettant de s'affranchir de cette hypothèse. Cette approche revient à effectuer une approximation linéaire locale d'un modèle autorégressif Hilbertien non linéaire. Ils définissent pour cela des estimateurs à noyau des opérateurs de covariance permettant de tenir compte de la proximité de l'année en cours avec les années passées dans la procédure d'estimation. Considérons un noyau réel gaussien K_h :

$$(2.30) \quad K(x) = (2\pi) \exp\left(-\frac{x^2}{2h}\right)$$

et définissons les estimateurs locaux des opérateurs de covariance :

$$(2.31) \quad \hat{\Delta}_h(y) = \frac{\sum_{i=1}^{n-1} K_h(\|y_i - y\|_M) y_i y_{i+1}' M}{\sum_{i=1}^{n-1} K_h(\|y_i - y\|_M)}$$

$$(2.32) \quad \hat{\Gamma}_h(y) = \frac{\sum_{i=1}^n K_h(\|y_i - y\|_M) y_i y_{i+1}' M}{\sum_{i=1}^n K_h(\|y_i - y\|_M)}$$

Définissons la matrice de poids $W_h(y)$ d'éléments diagonaux :

$$(2.33) \quad w_{i,h}(y) = \frac{K_h(y_i - y)' M (y_i - y)}{\sum_{l=1}^n K_h(y_l - y)' M (y_l - y)}, \quad i = 1, \dots, n.$$

Alors, les opérateurs de covariance s'expriment simplement

$$\hat{\Delta}_h(y) = \sum_{i=1}^n w_{i,h}(y) y_i y_{i+1}' M \quad \text{et} \quad \hat{\Gamma}_h(y) = Y' W_h(y) Y M$$

L'estimateur de ρ est alors défini en effectuant l'analyse spectrale de la matrice $\hat{\Gamma}_h$. L'inverse de $\hat{\Gamma}_h(y)$ de rang q s'écrit :

$$(\hat{P}_q \hat{\Gamma}_h \hat{P}_q)^{-1} = \sum_{l=1}^q \frac{1}{\hat{\lambda}_l} \hat{v}_l \hat{v}_l' M$$

où $(\hat{v}_l, l = 1, \dots, q)$ sont les vecteurs propres de $\hat{\Gamma}_h$ pour la métrique N associés aux plus grandes valeurs propres $(\hat{\lambda}_l, l = 1, \dots, q)$. En inversant l'équation liant les opérateurs ρ, Δ et Γ on définit ensuite un estimateur local de l'opérateur de corrélation ρ par :

$$(2.34) \quad \hat{\rho}_{q,h} = \left(\sum_{l=1}^q \hat{v}_l \hat{v}_l' \right) M \hat{\Delta}_h \left(\sum_{l=1}^q \frac{1}{\hat{\lambda}_l} \hat{v}_l \hat{v}_l' \right) M.$$

Une variante de cette méthode consiste à effectuer au préalable un lissage spline des trajectoires plutôt qu'une interpolation lorsque celles-ci sont bruitées. Elle présente l'inconvénient de dépendre de trois paramètres de lissage dont il faut ensuite déterminer les valeurs.

2.5.5 Interpolation linéaire

Pumo (1992) effectue une interpolation linéaire des trajectoires entre les points de discrétisation puis estime, en utilisant (1.2) et (1.3), les opérateurs de covariance. Un estimateur de la moyenne est ainsi obtenu en interpolant la moyenne empirique. D'autre part, si l'on note $\{e_1, \dots, e_p\}$ le système de fonctions linéaires utilisé pour l'interpolation et \mathbf{E} la matrice de dimension $p \times p$ vérifiant :

$$\int_0^1 \hat{f}_1(t) \hat{f}_2(t) dt = y_1' \mathbf{E} y_2,$$

alors l'opérateur de covariance admet pour représentation matricielle :

$$\hat{\Gamma} = X' D X \mathbf{E}$$

Cette solution ne comporte pas de paramètre de lissage, et par conséquent, moins flexible. Dans ces applications par la suite elle fournit des moins bonnes prévisions.

Chapitre 3

Simulation

Les expériences ont été réalisées à l'aide du logiciel **R**. Nous avons utilisé le package `far` qui permet d'obtenir la prévision par la méthode de noyau.

Nous simulons un processus $ARH(1)$ et son prédicteur noyau, nous utilisons la dernière observation pour calculer l'erreur de prévision puis on applique la méthode du noyau sur des données réelles plus précisément sur ENSO et la température d'Alger.

La simulation, en général, demande beaucoup d'attention, car les erreurs peuvent causer des problèmes, du modèle, soit dans la réalisation du processus, soit dans l'estimation des paramètres du modèle. Ces erreurs sont liées d'une part avec le fait qu'on réalise en réalité des nombres pseudo-aléatoires et d'autre part avec les erreurs de calculs.

Concernant le nombre k_n des vecteurs propres à utiliser le choix de h_n , nous utilisons une fonction `R` qui permet de faire la validation croisée.

3.1 Présentation des données

- La simulation d'un processus à termes dépendants demande indispensablement un modèle. Pour construire un $ARH(1)$ on pourrait partir directement de la formule

$$(3.1) \quad X_i = \sum_{j=0}^{\infty} \rho^j \varepsilon_{i-j}$$

où ρ est un opérateur symétrique de norme plus petite que un et (ε_i) un H -bruit blanc.

- Soit W_u un processus de Wiener standard sur $[0, \infty)$ de trajectoire continue et (ε_i) les fonctions aléatoires à valeurs sur $[0, 1]$ définies par :

$$\varepsilon_i(w) = \{v \rightarrow W_{i+v} - W_i, v \in [0, 1]\}$$

où $i \in N$. En considérant un deuxième processus de Wiener standard indépendant du premier on peut étendre $(\varepsilon_i, i \in N)$ à Z . Le fait qu'un processus de Wiener soit à accroissements indépendant permet d'établir que $(\varepsilon_i, i \in Z)$ est un H -bruit blanc. Il suffit alors d'utiliser (3.1).

- Nous avons préféré d'utiliser la relation de récurrence pour simuler un $ARC(1)$, i.e.

$$X_i = \rho(X_{i+1}) + \varepsilon_i.$$

- En utilisant ce modèle, pour simuler $n + 1$ observations d'un $ARC(1)$ il nous faudra un X_0 pour commencer et une suite $\varepsilon_{i=1, n+1}$.

3.1.1 Simulation d'un H-bruit blanc

Nous donnons ici un rappel du Théorème de Karhunen-Loève et une application de ce Théorème pour construire un processus de Wiener.

Un processus stochastique L^2 est une famille de variables aléatoires réels (pour nous ce sera suffisant de considérer les processus stochastiques réels) $\xi_t, t \in T$ telle que $E(|\xi_t|^2) < \infty$ pour tout $t \in T$.

Soit $(\xi_t, a \leq t \leq b)$, a, b finis, un processus stochastique L^2 de moyenne nulle et de covariance K , continue. Soit $(e_j, j = 1, 2, \dots)$ une base orthonormale de l'espace engendré par les fonctions propres correspondantes aux valeurs propres non nulles de l'opérateur intégral associé à K

$$(Af)(s) = \int_a^b K(s, t)f(t)dt$$

Dans ces conditions nous avons le Théorème de Karhunen-Loève qui donne

$$X(t) = \sum_{j=0}^{\infty} \lambda_j e_j(t), \quad t \in [a, b]$$

où $Z_j = \int_a^b X(t)e_j(t)dt$, sont des variables aléatoires orthogonales de moyenne nulle et $E[|Z_j|^2] = \lambda_j$. cette série converge en L^2 , uniformément par rapport à t . Si le processus est gaussien alors pour tout r , (Z_1, \dots, Z_r) est une variable aléatoire r -dimensionnelle gussienne.

Dans le cas d'un processus de Wiener, ie gaussien et tel que

$$K(s, t) = \min(s, t)$$

on trouve comme vecteurs et valeurs propres de l'opérateur intégral A respectivement

$$(3.2) \quad e_i(t) = \sqrt{2} \sin \left[\left(i - \frac{1}{2} \right) \pi t \right]$$

$$\mu_i = \frac{1}{\left(i - \frac{1}{2} \right)^2 \pi^2}.$$

En prenant $Z_j^* = Z_j / \sqrt{\mu_j}$, où sont donnés par l'expansion de Karhunen-Loève on obtient

$$X(t) = \sqrt{2} \sum_{j=1}^{\infty} Z_j^* \frac{\sin \left[\left(j - \frac{1}{2} \right) \pi t \right]}{\left(j - \frac{1}{2} \right) \pi}$$

où les (Z_j) sont des variables aléatoires i.i.d. normales réduites. Posons pour tout $\omega \in \Omega$ et $i \in N$

$$\begin{aligned} \varepsilon_i : [0, 1] \times \Omega &\rightarrow \mathbb{R} \\ (v, \omega) &\rightarrow W_{i+v}(\omega) - W_i(\omega) \end{aligned}$$

3.1.2 Définition d'un opérateur ρ

Maintenant passons à la définition de ρ . Puisque nous resterons dans le sous-espace engendré par les vecteurs e_i il suffit de définir $\rho[e_i]$. Nous prenons

$$\rho[e_i(s)] = \beta_i e_i(s)$$

Puisque $\Gamma_\epsilon(e_i) = \mu_i e_i$, Γ_ϵ et ρ sont commutables ainsi que Γ_{X_0} et ρ , nous avons

$$\Gamma_{X_0} = \Gamma_{X_0} \rho^2 + \Gamma_\epsilon;$$

et

$$\Gamma_{X_0}(e_i) = \frac{\mu_i}{1 - \beta_i^2} e_i.$$

Donc les éléments propres de Γ_{X_0} sont respectivement e_i et $\mu_i / (1 - \beta_i^2)$.

3.1.3 Une valeur initiale de X_0

Sans trouver une justification convenable on a pris la valeur $X_0 = \epsilon_0$. Pour la simulation d'un $AR(p)$, on pourrait prendre $X_0 = 0$ pour commencer et puis simuler une série d'observations X_0 jusqu'à ce qu'un équilibre soit arrivé.

Soient $(X_i^j)_{j=1,\dots,m}$ les coordonnées de X_i par rapport (e_i) pour $i = 0, 1, \dots, n$.

On a

$$(3.3) \quad X_{i+1}^j = \beta^j X_i^j + \epsilon_{i+1}^j, \quad j = 1, \dots, M.$$

qui facilite les calculs. En fait d'après (3) il suffit de calculer les coordonnées de X_0 et de $(\epsilon_i), i = 1, \dots, n+1$ par rapport aux (e_i) .

Remarquons enfin que les vecteurs propres de l'opérateur de la covariance du processus (ϵ_i) sont (e_i) et les valeurs propres (μ_i) définies ci-dessus.

3.2 Les fonctions utilisés dans R

► La fonction `BaseK2BaseC`

– Description

Etant donnée les coordonnées dans la base de Karhunen-Loève On calcule les coordonnées dans la base canonique

– Usage

`BaseK2BaseC(x, nb)`

– Arguments

`x` : une matrice contenant les coordonnées dans la base de Karhunen-Loève, une observation par colonne

`nb` : la dimension de la base canonique. Par défaut c'est la dimension de la base de Karhunen-Loève (*i.e.* nombre de ligne de `x`).

– Détails

L'expansion de Karhunen-Loève c'est la somme d'une infinité de termes. Mais ici la somme est tronquée par un nombre fini de termes.

– Valeur

Objet avec `nb` points de discrétisation et `n` observations similaire a `x`.

► Programme de la fonction BaseK2BaseC

```
BaseK2BaseC-fonction (x,nb=nrow(x))
{
  if(!is.matrix(x))
  x <- as.matrix(x)
  n <- nrow(x)
  cst1 <- ((1:n) - 0.5) * pi
  prod1 <- outer((1:nb)/nb, cst1, *)
  res <- sin(prod1)%*(x/cst1)*sqrt2
  res <- as.fdata(res, dates 1:ncol(x))
  return (res)
}
```

► La fonction simul.wiener– Description

Simulation du processus de Wiener.

– Arguments

`m` : entier. Nombre des points discrétisation.

`n` : entier. Nombre d'observations.

`m2` : entier. Length de the Karhunen-Loève expansion (2m par default).

– Détails

Cette fonction utilise l'expansion de Karhunen-Loève de processus de Wiener connu pour simuler les observations d'un tel procédé.

L'option `m2` est utilisé en interne pour définir la longueur de l'expansion. Cette expansion a besoin d'être plus grande que le nombre de points de discrétisation, mais une valeur trop importante peuvent ralentir la production. La valeur par défaut est `2m`.

– Valeur

Un objet `fdata` contenant une variable ("`var`") qui est un processus de Wiener de longueur `n` avec

m discrétisation des points.

Exemples 3.2.1

```
noise <- simul.wiener(m=64,n=100,m2=512)
summary(noise)
par(mfrow=c(2,1))
plot(noise,date=1)
plot(select.fdata(noise,date=1:5),whole=TRUE,separator=TRUE)
```

► La fonction fdata

– Description

Objet de la classe `fdata` et ses méthodes.

– Usage

```
as.fdata(object,...)
as.fdata.matrix(object,..., col, p, dates, name)
as.fdata.list(object,..., dates, name)
```

– Arguments

`objet` Une matrice ou une liste.
`col` Vecteur donnant les noms des variables à inclure dans l'objet `fdata`.
`p` Une valeur réelle donnant le nombre de points de discrétisation choisi.
`Dates` Un vecteur de caractères contenant les dates des observations.
`name` Un vecteur de caractères contenant les noms des variables.
`...` Des arguments supplémentaires.

– Détails

Objets `fdata` sont principalement utilisés pour modéliser des données fonctionnelles dans le but de calculer un modèle autorégressif fonctionnelle par les fonctions `loin` et `kerfon`.

Un `fdata` est composé d'une ou plusieurs variables. Chacune d'entre elle est une série temporelle fonctionnelle. Pour être plus précis, chaque variable a une des données fonctionnelles par élément des dates (explicitement donné ou implicitement déduite). Ainsi, le nombre d'observations fonctionnelles sont des données communes.

– Valeur

Un objet de `fdata` classe.

► La fonction `kerfon`

– Description

Modélisation de `fdata` utilisant le noyau fonctionnel.

– Usage

`kerfon(data, x, r, hmin, hmax, na.rm=TRUE)`.

– Arguments

A *Données `fdata` objet.*
x *Le nom de la variable étudiée.*
r *Nombre d'observations utilisées pour valider le modèle.*
hmin *Valeur minimale de la bande passante.*
hmax *Valeur maximale de la largeur de bande.*
na.rm *logique. Est – ce que la N.A. besoin d'être enlevé.*

– Détails

Cette fonction construit un modèle de noyau fonctionnel et effectue l'estimation de cette bande passante.

Cet estimateur est défini par : (2.29).

La fonction `kerfon` utilise la validation croisée à une valeur de h_n . Les paramètres `hmin` et `hmax` sont utilisés, lorsqu'ils sont fournis, pour contrôler les valeurs admissibles de h_n . Pour choisir la valeur de h_n , vous devez fournir la même valeur pour les deux `hmin` et `hmax`. Lors de la validation croisée, considérant que l'objet `fdata` x contient n observations, la fonction utilise les $(n - r)$ premières observations, et de calculer l'erreur quadratique moyenne sur les dernières observations r .

– Valeur

Un objet `kerfon`. L'objet est une liste avec les éléments suivants :

call *L'appel de la fonction.*
h *La bande passante (trois valeurs : optimale, minimum, maximum).*
x *Le nom de la variable choisie.*
xdata *Les valeurs passées pour `x`.*
ydata *Les valeurs associées pour `x`.*

Exemples 3.2.2

```
#Simulation de procesus de Wiener
data1 <- simul.wiener(m=10,n=400,base=base.simul.kerfon(20,5)),
base.exo=base.simul.far(20,5),
d.a=matrix(c(0.5,0),nrow=1,ncol=2),
alpha.conj=matrix(c(0.2,0),nrow=1,ncol=2),
d.rho=diag(c(0.45,0.90,0.34,0.45)),
alpha=diag(c(0.5,0.23,0.018)),
d.rho.exo=diag(c(0.45,0.90,0.34,0.45)),
(cst1=0.0)
#Validation croisé
model1 <- kerfon(data=data1, x='X', r=10, na.rm=TRUE)
print(model1)
```

► La fonction predict.kerfon

– Description

Le calcul de la prédiction basée sur un modèle de noyau fonctionnel.

– Usage

```
predict.kerfon(object, ..., newdata=NULL, label, na.rm=TRUE, positive=FALSE)
```

– Arguments

<code>object</code>	<i>La suite de l'objet <code>kerfon</code> de la fonction de <code>kerfon</code>.</i>
<code>newdata</code>	<i>Un objet <code>fdata</code> utilisée dans le modèle de <code>kerfon</code> pour calculer la prdition, ou à <code>NULL</code>.</i>
<code>prdict</code>	<i>Prévision un pas en avant avec les données dans l'objet.</i>
<code>label</code>	<i>Etiqueter un vecteur de caractère donnant les dates associées aux observations prévues.</i>
<code>na.rm</code>	<i>Logique. Est – ce que la N.A. besoin d'tre enlevé.</i>
<code>positive</code>	<i>Logique. Le résultat doit tre forcée à des valeurs positives.</i>
<code>...</code>	<i>Des arguments supplémentaires.</i>

– Détails

Cette fonction calcule à un pas en avant pour la prédiction d'un modèle `kerfon`.

– Valeur

Un objet `fdata`

Exemples 3.2.3

```

#Simulation de processus de Wiener
data1 <- simul.wiener(m=10,n=400,base=base.simul.far(20,5)),
base.exo=base.simul.far(20,5),
d.a=matrix(c(0.5,0),nrow=1,ncol=2),
alpha.conj=matrix(c(0.2,0),nrow=1,ncol=2),
d.rho=diag(c(0.45,0.90,0.34,0.45)),
alpha=diag(c(0.5,0.23,0.018)),
d.rho.exo=diag(c(0.45,0.90,0.34,0.45)),
(cst1=0.0)
#Validation croisé
model1 <- kerfon(data=data1, x='X', r=10, na.rm=TRUE)
print(model1)
#La prévision des valeurs
pred1 <- predict(model1,newdata=select.fdata(data1,date=1:399))
#Les valeurs réelles
real1 <- select.fdata(data1,date=2:400)
errors0 <- persist1[[1]]-real1[[1]]
errors1 <- pred1[[1]]-real1[[1]]

```

3.3 Application**3.3.1 Sur des données simulés**

Nous testons les résultats précédents concernant le noyau sur des données simulées pour cela nous exécutons le programme suivant :

- **Programme :**

```

data1=simul.far.wiener(m=64, n=128, d.rho=diag(c(0.45, 0.9, 0.34, 0.45)), cst1=0.05,m2=NULL)
plot(data1,date=128)
DATA2=kerfon(data1)
print(DATA2)
data3=predict.kerfon(DATA2, newdata=data1,na.rm=TRUE, positive=FALSE)
par(new=T)
plot(data3,date=128,col='red')
legend(0.1,1,c('simul', 'prédicteur'),lty=1:2,col=1:2,pch=c(-1,1))

```



FIGURE 3.1 – Observation et prévision noyau d’un modèle simulé à temps continu.

Dans ce modèle on a simulé 128 trajectoires d’un processus ARH(1) avec 64 points de discrétisation on a utilisé 127 trajectoires pour estimer le prédicteur noyau, la figure (3.1) représenté la dernière trajectoire ainsi que celle de prédicteur noyau.

3.3.2 Sur des donnés réels

1. ENSO : (El Nino Southern Oscillation) est un phénomène climatique majeur d'interaction entre l'atmosphère et l'océan Pacifique dans la zone tropicale. El Nino (EN) se caractérise par des variations importantes de la température, particulièrement en décembre, très corrélées des oscillations de la pression atmosphérique (SO). EN est donc observé à travers l'évolution de températures moyennes mensuelles mesurées au large du Chili à la surface de l'océan tandis que SO est mesurée par la pression atmosphérique au niveau de la mer à Tahiti. Ces séries sont connues depuis 1982 et considérées dans cette étude ces observation jusqu'à l'année 2013.

- Programme :

```
f=file.choose() ;
g=read.table(f,skip=1) ;
g[[5]] ;tinin=g[[5]][1 :372]
tt=matrix(tinin,nrow=12,ncol=31)
ff=as.fdata(tt)
fun=kerfon(ff)
prev=predict.kerfon(fun,newdata=ff)
plot(as.fdata(prev[[1]]),date=31,col=4)
par(new=T)
gg=as.fdata(matrix(g[[5]][1 :384],nrow=12,ncol=32))
OB=matrix(g[[5]][1 :384],nrow=12,ncol=32)
plot(gg,date=32)
```

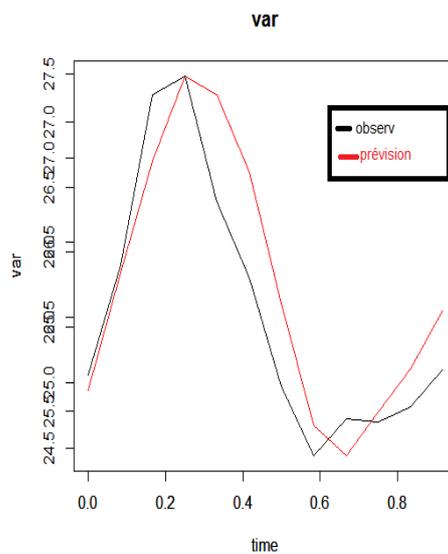


FIGURE 3.2 – Observation et prévision noyau d'un modèle réel à temps continu "ENSO" (année 2013).

```
err=sqrt(sum(OB[,32]-(prev[[1]][,31]))^2)
er
[1] 8.396426
```

2. Température d'Alger : On applique maintenant cette méthode de prévision sur les données de la température moyenne d'Alger. Nous utilisons les données de 1957 à 2003 et nous comparons les résultats de l'année 2004 à ceux obtenus par le prédicteur du noyau en calculant l'erreur quadratique minimale.

• **Programme :**

```
f=file.choose();g=read.table(f,skip=1)
names(g)=NULL
gg=g[1:47,]
ggg=t(gg)
f=as.fdata(ggg)
fun=kerfon(f)
prev=predict.kerfon(fun,newdata=f)
d= t(g)
p=d[,48]
p1=as.fdata(p, name="Température d'Alger")
```

```
plot(as.fdata(prev[[1]],name=""),date=47,col="red")
par(new=T)
plot(p1)
```

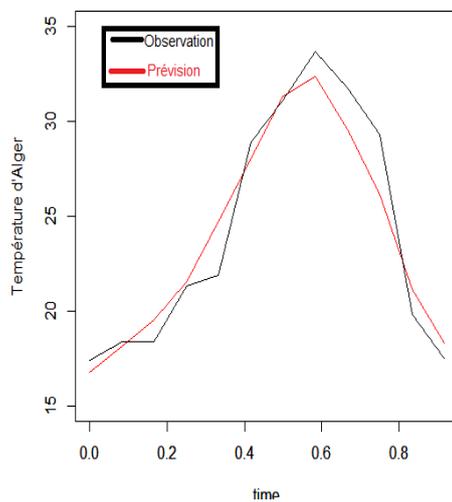


FIGURE 3.3 – Observation et prévision noyau d'un modèle réel à temps continu "Alger" (année 2004).

```
err=sqrt(sum(p-(prev[[1]][,47]))^2)
err
[1] 5.410018
```

Bibliographie

- [1] Approximation spline de la prévision d'un processus fonctionnel autorégressif d'ordre 1.
- [2] Bardet, J.M. 1997. *Tests d'autosimilarité des processus gaussiens. Dimension fractale et dimension de corrélation*. Thèse 3^{ème} cycle, Paris-Sud.
- [3] Bogachev, V.I. 1999. *Gaussian measures*. Math surveys and monographs, **62**, Amer. Math. Soc.
- [4] Benhenni, K., Ferraty, F., Rachdi, M. and Vieu, P. 2007. Local smoothing regression with functional data. *Comput. Statist.*, **22**.
- [5] D.BODQ. 1991. *Non-Parametric Estimation and Prediction for Continuous Time processes*, 509-529. Roussas, G.
- [6] D.Bosq, 2000. *Linear Processes in Function Spaces : Theory and applications*.
- [7] D.BOSQ. 1983. *Non Parametric Prediction in Stationary Processes*, 69-84.
- [8] Damon Julien, August 29, 2013. *Package "far"*.
- [9] D.BOSQ. 1998. *Non Parametric Statistics for Stochastics Processes, Estimation and Prediction*, 2^{ième} édition.
- [10] F.Ferraty, et P.Vieu, (2001). *Functional Nonparametric Model : a New Tool for Spectrometric Data*. *Soumis pour publication*.
- [11] E.Nadaraya, 1964. On estimating regression. *Theory Prob. Appl. Theory Prob. Appl.*, **10**, 186-196.
- [12] Frédéric Ferraty et Philippe Vieu, 2003. *Modèles Non-Paramétriques de Régression*.
- [13] G. WAHBA. 1990. *Spline Models for Observational Data SLAM*.
- [14] Tukey, J.W. 1961. Curve as parameters, and touch estimation. *Proceedings of the 4th Symposium on Mathematics, Statistics and Probability*, 681-694, Berkeley, CA, USA.
- [15] P.Besse, H.Cardot, and F.Ferraty, 1997. Simultaneous nonparametric regressions of unbalanced longitudinal data. *Comput. Statist. Data Anal.*, **24**.
- [16] P.Besse, H.Cardot, and F.Ferraty, 1997. Simultaneous nonparametric regressions of unbalanced longitudinal data. *Comput. Statist. Data Anal.*, **24**.

-
- [17] P.BESS et H.CARDOT.1996 Approximation spline de la prévision d'un processus fonctionnelle autorégressif d'ordre 1. *Revue Canadienne de Statistique*, **24**, 467-487. Lecture Notes in Statistics, **149**, Springer.
- [18] PHILIPPE, BESSE, H.CARDOT. *Autoregressive Forecasting of Some Functional Climatic Variations*. 2000.
- [19] Pumo, 1992. *Estimation et prévision de processus fonctionnels autorégressifs : application aux processus à temps continu*. PhD thesis, Université Paris VI.
- [20] T.MOURID. 1995. *Contribution à la statistique des processus autorégressifs à temps continu*.