

Remerciements

Tout d'abord, je tiens à remercier **Allah**, Le Tout Puissant, de m'avoir donné la santé, la volonté et la patience pour mener à terme ma formation de Master.

Je remercie Madame **MOKHARI FATIHA** qui a su me guider durant la rédaction de ce mémoire, sa rigueur et ses conseils nombreux et avisés ont stimulé en moi le sens de la recherche et l'ardeur au travail.

Je remercie tout particulièrement Monsieur **T. GUENDOUI** d'avoir présidé ce jury de mémoire.

Je remercie également Madame **F. BENZIADI** et Monsieur **F. MADANI**, d'avoir accepté de juger ce travail qu'ils veillent trouver ici l'expression de ma différence et de ma profonde gratitude.

Je remercie du fond de mon cœur, mes parents qui m'ont soutenus, encouragés et motivés tout au long de mes études.

Enfin, je ne pourrais terminer ces remerciements sans une pensée à l'ensemble de mes enseignants qui sont à l'origine de tout mon savoir.

Dédicaces

A l'aide du dieu et son aide et grâce à la prière de mes parents, j'ai terminé ce modeste travail que je le dédie en premier lieu à :

Mes très chers parents qui m'ont élevé et donné le meilleur de leur même, pour me permettre une situation professionnelle et sociale honorable dans le future.

Je remercie :

Mon frère : **ABDELHAMID.**

Mes chères sœurs : **NACIRA, HALIMA.**

Mes chères amies : **FALAH LATIFA, ROUANE KHADIJA, HAMADA ISRAA, BENARBIA ABDELAZIZ.**

Table des matières

| | | |
|----------|--|-----------|
| 1 | PROCESSUS AUTOREGRESSIVES | 7 |
| 1.1 | Processus autorégressifs Banachiques d'ordre un | 7 |
| 1.1.1 | Définition et existence d'un processus autorégressif Banachique d'ordre un | 7 |
| 1.1.2 | Représentation ARB(1) de processus réels à temps continu | 9 |
| 1.2 | Processus autorégressifs Hilbertiens d'ordre un | 12 |
| 1.2.1 | Définition d'un ARH(1) | 14 |
| 1.2.2 | Estimation des opérateurs de covariance Γ et Δ | 15 |
| 1.2.3 | Estimation des éléments propres de Γ | 17 |
| 1.2.4 | Estimation de ρ et prévision de X_{n+1} | 18 |
| 2 | UN MODÈLE AUTOREGRESSIF D'ORDRE UN DANS $C_{[0,1]}$ | 20 |
| 2.1 | Définitions et premiers résultats | 20 |
| 2.1.1 | Un résultat pour les fonctions Lipschiziennes d'ordre α | 22 |
| 2.2 | Estimation de ρ dans $C_{[0,1]}$ connaissant les vecteurs propres de l'opérateur Γ | 24 |
| 2.3 | Estimation de ρ dans le cas général | 26 |
| 3 | SIMULATION ET ESTIMATION D'UN ARC(1) | 28 |
| 3.1 | Simulation d'un ARC(1) | 28 |
| 3.1.1 | Simulation d'un H-bruit blanc | 29 |
| 3.1.2 | Définition d'un opérateur ρ | 30 |
| 3.1.3 | Une valeur initiale de X_0 | 31 |
| 3.2 | Estimation et prévision du processus $ARC(1)$ | 31 |
| 3.2.1 | Les vecteurs propres de Γ_{X_0} sont connus | 31 |
| 3.2.2 | Les vecteurs propres de Γ_{X_0} sont inconnus | 34 |

| | | |
|-------|---|----|
| 3.3 | Applications | 36 |
| 3.3.1 | Fonctions \mathbf{R} | 36 |
| 3.3.2 | Applications sur des données simulées | 45 |
| 3.3.3 | Applications sur des données réelles | 46 |

Introduction

Considérons un processus à temps continu observé sur des intervalles de même longueur δ qui peuvent être adjacents ou disjoints, alors chaque observation définit une variable aléatoire dans un espace fonctionnel. La prévision globale de l'évolution future du processus sur un intervalle de longueur δ est un problème qui se pose naturellement dans les applications.

Si le processus est observé sur l'intervalle $[0, T]$ on peut se ramener au schéma précédent en découpant $[0, T]$ en intervalles de longueur δ (non nécessairement disjoints). Nous nous intéressons à certains problèmes relatifs au processus à temps continu, que nous traitons à l'aide du processus à valeurs dans des espaces fonctionnels.

Deux types d'espaces fonctionnels nous semblent intéressants dans les application : l'espace de Hilbert $L^2_{[0,\delta]}$ et l'espace de Banach $C_{[0,\delta]}$.

Dans le premier chapitre, nous définissons d'un $ARB(1)$. Par ailleurs, nous montrons que ce processus vérifie la relation $X_i = \rho X_{i-1} + \varepsilon_i$ où ρ est un opérateur borné continu et $(\varepsilon_i, i \in Z)$ est un B-bruit blanc. Nous établissons ensuite une relation entre les opérateurs de la covariance du processus X_i et ρ .

Nous revenons ensuite aux processus $ARH(1)$. Nous citons les résultats les plus importants, dus à D.Bosq, concernant ces processus. Ces résultats consistent en l'estimation de ρ et le calcul de la prévision de X_{n+1} .

L'étude dans le second chapitre est consacré au processus autorégressif à valeurs dans l'espace des fonctions continues sur $[0, 1]$, noté $ARC(1)$. On définit dans ce cas un opérateur ρ de la forme intégrale. En supposant connus les vecteur propres de l'opérateur de la covariance Γ de X_0 , en étudie la convergence en probabilité du prédicteur construit par **Pumo** dans sa thèse [21].

Dans le cas général, et en considérant le sous-espace des fonctions de Lipschitz. En supposant que le processus est géométriquement α - mélangeant. On trouve encore un prédicteur presque surement convergeant construit à partir des vecteurs propres de l'opérateur de la covariance empirique Γ_n .

Enfin dans le dernier chapitre nous traitons les applications numériques. Nous simulons un $ARC(1)$ et testons les résultats obtenus par Pumo. On applique après ces résultats sur des données réelles telles que la température de Nottingham et du phénomène EL NINO. A la fin on compare la méthode $FAR(1)$ de ce travail avec d'autres méthodes de prévision telle que la méthode noyau.

Chapitre 1

PROCESSUS AUTOREGRESSIVES

1.1 Processus autorégressifs Banachiques d'ordre un

Dans ce chapitre nous donnons la définition d'un $ARB(1)$, puis une relation entre l'opérateur de la covariance Γ de X_0 , l'opérateur de la covariance croisée Δ^* de X_1 , X_0 et l'opérateur d'autocorrélation ρ . Nous indiquons ensuite deux exemples de processus réels à temps continu qui peuvent se mettre sous la forme d'un $ARB(1)$.

1.1.1 Définition et existence d'un processus autorégressif Banachique d'ordre un

Soit (B, β_B) un espace de Banach séparable muni de la tribu borélienne et de la norme $\|\cdot\|$. Considérons une suite $(\varepsilon_t, t \in Z)$ de v.a à valeurs dans B , définies sur l'espace probabilisé (Ω, A, P) indépendantes et de même loi telles que $0 < \sigma^2 = E\|\varepsilon_t\|^2 < +\infty$ et $E(\varepsilon_t) = 0$ où l'espérance est prise au sens de Bochner. Nous dirons qu'une telle suite est un B - bruit blanc. Soit ρ un opérateur linéaire borné de B dans B et on note $\|\rho\|_L$ la norme d'opérateurs linéaires bornés et $m \in B$.

On dira qu'une suite de v.a. $(X_t, t \in Z)$ définie sur (Ω, A, P) et à valeurs dans (B, β_B) est un processus autorégressif d'ordre 1 noté $ARB(1)$, si

$$(1.1) \quad X_t - m = \rho(X_{t-1} - m) + \varepsilon_t, \quad t \in Z$$

Lemme 1.1.1 *Si il existe $j_0 \geq 1$ tel que $\|\rho^{j_0}\| < 1$ alors la série*

$$(1.2) \quad \sum_{j \geq 1} \rho^j \varepsilon_{t-j}, \quad t \in Z$$

converge dans L_B^2 et presque sûrement.

On dira que (X_t) est un processus autorégressif d'ordre 1 à valeurs dans B d'opérateur d'autocorrélation ρ , d'innovation ε_t et de moyenne m .

Soit X une v.a. centrée à valeurs dans B et P_X la loi de X . La covariance r_X est définie par

$$r_X(x^*, y^*) = \int_B x^*(x) y^*(x) dP_x$$

où $x^*, y^* \in B^*$, dual topologique de B ; c'est une forme bilinéaire, symétrique, continue et positive sur B^* .

Il existe un opérateur linéaire

$$R_X : B^* \rightarrow B^{**}$$

associé à r_X défini pour tout $y^* \in B^*$ par

$$\langle x^*, R_X y^* \rangle = r_X(x^*, y^*)$$

Cet opérateur est positif, symétrique et continu. Si de plus $\mathbb{E}\|X\|^2 < +\infty$ alors

$$R_X(x^*) = \mathbb{E}(\langle X, x^* \rangle X)$$

donc $R_X : B^* \rightarrow B$. Il est nucléaire i.e. il admet la représentation $R_X(x^*) = \sum_k \langle x^*, x_k^{**} \rangle y_k$ où $y_k \in B$ et $\sum_k \|x_k^{**}\| \|y_k\| < +\infty$. De plus il existe une suite $(x_n) \subset B$ t.q. $\sum_k \|x_n\|^2 < \infty$

et pour laquelle on a

$$R_X(x^*) = \sum_n \langle x_n, x^* \rangle x_n$$

Si Y est une autre v.a. à valeurs dans B de moyenne nulle la covariance croisée de X et Y est définie par

$$r_{X,Y}(x^*y^*) = \mathbb{E}(x^*(X)y^*(Y))$$

et l'opérateur de la covariance croisée de X et Y , $R_{X,Y}$

$$r_{X,Y}(x^*y^*) = \langle x^*, R_{X,Y}(y^*) \rangle$$

Si X et Y sont des v.a. du second ordre alors

$$R_{X,Y}(y^*) = \mathbb{E}(y^*(Y)X)$$

Il est aussi nucléaire et on a $R_{X,X}(x^*) = R_X(x^*)$, pour $x^* \in B^*$.

Considérons le modèle (1.1) noté $ARB(1)$ et supposons que $m = 0$. L'indépendance de X_{t-1} et ε_t entraîne $\mathbb{E}(x^*(X_0)\varepsilon_1) = 0$. La suite $(X_t, t \in \mathbb{Z})$ étant stationnaire R_{X_t} et R_{X_{t+1}, X_t} ne dépendent pas de t .

Lemme 1.1.2 *Si (X_t) est un $ARB(1)$ associé à ρ et ε_t*

$$(1.3) \quad \begin{aligned} \mathbb{E}X_0 &= 0 \\ R_{X_1, X_0} &= \rho R_{X_0} \end{aligned}$$

1.1.2 Représentation $ARB(1)$ de processus réels à temps continu

Nous indiquons une classe de processus réels à temps continu admettant une représentation $ARB(1)$. Cette classe contient notamment le processus d'Ornstein-Uhlenbeck.

1.1.2.1 Le processus d'Ornstein-Uhlenbeck

On considère le processus réel

$$\xi_t = \int_{-\infty}^t \exp(-c(t-u)) dW(u), \quad t \in \mathbb{R}$$

où W est un processus de Wiener et c une constante strictement positive.

Pour construire un $ARB(1)$ à partir de (ξ_t) on peut choisir $B = C_{[0,1]}$ et poser

$$X_n(t) = \xi_{n+t}, \quad 0 \leq t \leq 1, \quad n \in \mathbb{Z}$$

La version choisie de ξ_t étant supposée à trajectoires continues, on définit ainsi des v.a X_n à valeurs dans $C = C_{[0,1]}$.

D'une part,

$$\begin{aligned} \mathbb{E}(\xi_{n+t}/\xi_s, s \leq n) &= \mathbb{E}(\int_{-\infty}^{n+t} e^{-c(t-u)} dW(u) / \xi_s, s \leq n) \\ &= \mathbb{E}(\int_{-\infty}^n e^{-c(t-u)} dW(u) + E(\int_n^{n+t} e^{-c(t-u)} dW(u) / \xi_s, s \leq n) \\ &= \mathbb{E}(\int_{-\infty}^n e^{-c(t-u)} dW(u) / \xi_s, s \leq n) \\ &= \int_{-\infty}^n e^{-c(t-u)} dW(u) \\ &= e^{-ct} \xi_n \quad 0 \leq t \leq 1 \end{aligned}$$

Ce qui amène à poser :

$$\rho f(t) = e^{-ct} f(1), \quad 0 \leq t \leq 1, \quad f \in C,$$

et

$$\varepsilon_n(t) = \int_n^{n+t} e^{-c(n+t-u)} dw(u) = \int_0^t e^{-c(t+v)} dw(v)$$

Alors

$$\begin{aligned}\varepsilon_n(t) &= \int_{\infty}^{n+t} e^{-c(n+t-u)} dw(u) - e^{-ct} \int_{\infty}^n e^{-c(n-v)} dw(v) \\ &= \xi_{n+t} - (\rho X_{n-1})(t) \quad 0 \leq t \leq 1, \quad n \in \mathbb{Z}\end{aligned}$$

D'autre part, comme $(W(u))$ est à accroissements indépendants, (ε_n) un bruit blanc et

$$\|\rho^n\| = \sup_{\|f\|=1} \|\rho^n f\| = e^{-c(n-1)}$$

alors $\sum_{n \geq 0} \|\rho^n\| < \infty$ et ainsi (X_n) est un $ARB(1)$.

1.1.2.3 Une classe d'ARB(1) à valeurs dans $L^2_{[0,1]}$

Soit $(Z_t, t \in \mathbb{R})$ un processus réel du second ordre, centré et à accroissements indépendants et strictement stationnaires. On suppose que nous avons une version de (Z_t) à trajectoires localement de carrés intégrables. Alors en posant :

$$\varepsilon_n(t) = Z_{n+1} - Z_n, \quad 0 \leq t \leq 1, \quad n \in \mathbb{Z}$$

on définit un bruit blanc dans $L^2_{[0,1]} = H$.

Soit ρ un opérateur linéaire sur H , intégral de noyau K défini par

$$(\rho f)(t) = \int_0^1 \mathbf{K}(s, t) f(s) ds, \quad 0 \leq t \leq 1, \quad f \in L^2_{[0,1]}$$

où

$$0 \leq \int_0^1 \int_0^1 \mathbf{K}^2(s, t) dt ds < 1$$

Alors $\|\rho\| < 1$ et on définit un $ARH(1)$ en posant :

$$X_n(t) = \sum_{i \geq 0} \rho^i(\varepsilon_{n-i})(t), \quad 0 \leq t \leq 1$$

(Z_t) peut être un processus de Wiener où un processus de Poisson centré.

1.1.2.4 Processus avec saisonnalité

Considérons un processus réel de la forme

$$\eta_t = m(t) + \xi_t, \quad t \in R$$

où (ξ_t) est un processus centré, à trajectoires continues et admettant une représentation $ARB(1)$ où $B = C_{[0,1]}$.

On suppose que m est une fonction continue, non aléatoire, de période h et non constante. Dans ces conditions η_t admet une représentation $ARB(1)$ avec

$$X_n(t) = \xi_{n+t}, \quad 0 \leq t \leq 1, \quad n \in Z$$

et

$$\mathbb{E}X_n = m$$

X_n est donc stationnaire alors que (η_t) ne l'est pas.

Conclusion : De nombreux processus usuels admettent une représentation autorégressive dans un espace de Banach bien choisi.

1.2 Processus autorégressifs Hilbertiens d'ordre un

le but principal de l'étude d'un processus autorégressif Hilbertien d'ordre un, $ARH(1)$, est la prévision d'un processus à temps continu. On cherche pour cela un estimateur de l'opérateur ρ .

Nous donnerons ici la définition d'un $ARH(1)$ tel qu'on utilisera dans ce chapitre avec

deux hypothèses supplémentaires par rapport à la définition d'un $ARB(1)$.

Mais avant de commencer donnons quelques définitions utiles :

opérateurs compact

Nous considérons un espace de Hilbert séparable H avec le produit intérieur $\langle \cdot, \cdot \rangle$ qui génère la norme $\|\cdot\|$, et notons \mathcal{L} l'espace des opérateurs linéaires bornés sur H avec la norme

$$\|\rho\|_{\mathcal{L}} = \sup \|\rho(x)\|, \|x\| \leq 1$$

on dit que $(\rho \in \mathcal{L})$ est un opérateur compact si il existe deux bases orthonormées $\{v_j\}$ et $\{f_j\}$, et une suite λ_j convergent vers zéro, telle que :

$$(1.4) \quad \rho(x) = \sum_{j=1}^{\infty} \lambda_j \langle x, v_j \rangle f_j, \quad x \in H$$

opérateurs de Hilbert-Schmidt

Un opérateur compact admettant la représentation (1.4) est dit être un opérateur de Hilbert-Schmidt si $\sum_{j=1}^{\infty} \lambda_j^2 < \infty$. L'espace S des opérateurs de Hilbert-Schmidt est séparable avec le produit scalaire :

$$(1.5) \quad \langle \rho_1, \rho_2 \rangle_S = \sum_{i=1}^{\infty} \langle \rho_1(e_i), \rho_2(e_i) \rangle$$

Où e_i est une base orthonormée arbitraire. On peut montrer que

$$\|\rho\|_S^2 = \sum_{j \geq 1} \lambda_j^2 \text{ et :}$$

$$(1.6) \quad \|\rho\|_{\mathcal{L}} \leq \|\rho\|_S$$

Un opérateur $\rho \in \mathcal{L}$ est dit symétrique si :

$$(1.7) \quad \langle \rho(x), y \rangle = \langle x, \rho(y) \rangle, \quad x, y \in H$$

et définie positive si :

$$(1.8) \quad \langle \rho(x), x \rangle \geq 0, \quad x \in H$$

Un opérateur ρ définie positif de Hilbert-Schmidt symétrique admet la décomposition :

$$(1.9) \quad \rho(x) = \sum_{j=1}^{\infty} \lambda_j \langle x, v_j \rangle v_j, \quad x \in H$$

avec v_j les fonctions propres de ρ , c'est à dire $\rho(v_j) = \lambda_j v_j$.

1.2.1 Définition d'un ARH(1)

Définition 1.2.1 *Considérons un espace de Hilbert réel et séparable muni de la norme $\|\cdot\|$. Un ARH(1) est une suite de variables aléatoires à valeurs dans H telle que :*

$$(1.10) \quad X_i = \rho X_{i-1} + \varepsilon_i \quad i = 0, \pm 1, \pm 2$$

où (ε_i) est un H -bruit blanc et ρ est un opérateur compact symétrique sur H tel que $\|\rho^{j_0}\|_L < 1$ pour $j_0 \geq 1$.

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

$$\begin{aligned} C_{x_0}(x) &= E(\langle X_0, x \rangle X_0) \\ C_{X_0 X_1}(x) &= E(\langle X_0, x \rangle X_1) \end{aligned}$$

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 :*

$$\begin{aligned} EX_0 &= 0 \\ C_{X_0} &= \rho C_{X_0} \rho + C_\varepsilon \\ &= \sum_{j=0}^{\infty} \rho^j C_\varepsilon \rho^j \\ C_{X_0 X_{-h}} &= C_{X_0} \rho^h, \quad h = 1, 2, \dots \\ C_{X_{-h}, X_0} &= \rho^h C_{X_0}, \quad h = 1, 2, \dots \end{aligned}$$

Dans la suite on note

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

Les deux dernières équations de la proposition précédente s'écrivent donc

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

Comme cela a été signalé dans le livre de Bosq, 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.

1.2.2 Estimation des opérateurs de covariance Γ et Δ

Rappelons encore une fois que $\Gamma(x) = E(\langle X_0, x \rangle X_0)$ et $\Delta(x) = E(\langle X_0, x \rangle X_1)$. Ce sont des opérateurs de Hilbert-Schmidt. En fait on a un résultat plus général.

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

$$\begin{aligned} \sum_j \|C_{XY}(e_j)\|^2 &= \sum_j \sum_l \langle C_{XY}(e_j), e_l \rangle^2 \\ &= \sum_j \sum_l (E\langle X, e_j \rangle \langle Y, e_l \rangle)^2 \\ &\leq \sum_j \sum_l (E\langle X, e_j \rangle^2 E\langle Y, e_l \rangle^2) \\ (1.11) \qquad &\leq E\|X\|^2 E\|Y\|^2 < \infty \end{aligned}$$

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

$$\|\Gamma\|_S = \left(\sum_{jl} |E(\langle X_0, e_j \rangle \langle X_0, e_l \rangle)|^2 \right)^{\frac{1}{2}}$$

$$\|\Delta\|_S = \left(\sum_{jl} |E(\langle X_0, e_j \rangle \langle X_1, e_l \rangle)| \right)^{\frac{1}{2}}$$

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

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

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

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

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 $E\|X_0\|^4 < \infty$. Alors*

$$(1.12) \quad E\|\Gamma_n - \Gamma\|_S^2 \leq \frac{A}{n}$$

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

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

Proposition 1.2.3 *Pour $n \geq 2$*

$$(1.13) \quad E\|\Delta_n - \Delta\|_S^2 = E\|\Delta_n^*\|_S^2 \leq \frac{B}{n-1}$$

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

De plus on a :

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

1.2.3 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*

$$(1.14) \quad |\lambda_{jn} - \lambda_j| \leq \|\Gamma_n - \Gamma\|,$$

et si $\lambda_1 > \lambda_2$

$$(1.15) \quad \|\varphi_{1n} - \varphi'_1\| \leq a_1 \|\Gamma_n - \Gamma\|,$$

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

$$(1.16) \quad \|\varphi_{jn} - \varphi'_j\| \leq a_j \|\Gamma_n - \Gamma\|,$$

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

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

Proposition 1.2.4 *Nous avons*

$$(1.17) \quad E \left(\sup_{j \geq 1} |\lambda_{jn} - \lambda_j|^2 \right) \leq \frac{A}{n}$$

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

et sous les hypothèses de la proposition (1.2.2) la vitesse de convergence est de l'ordre $O(\exp(-cn^\gamma)); 0 < \gamma < \frac{1}{4}$

Proposition 1.2.5 *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 la Prop (1.2.1) est $\|\varphi_{jn} - \varphi'_j\|^2 \rightarrow 0$ p.s et la vitesse de convergence est $O(\exp(-cn^\gamma))$, où $0 < \gamma < \frac{1}{4}$.

Proposition 1.2.6 Si $\|X_0\| \leq d$ p.s, (X_i) 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 $n > n_\gamma$

$$(1.18) \quad P(\|\Gamma_n - \Gamma\|_S > \varepsilon) \leq 2n^\gamma \exp\left(-\frac{\varepsilon n^\gamma}{10d^2 \varepsilon \sqrt{2}}\right) + \frac{2ad^2 r}{\varepsilon^2(1-\sqrt{r})^2} \exp\left(-\log \frac{1}{r} n^\gamma\right)$$

et

$$(1.19) \quad P(\|\Delta_n - \Delta\|_S > \varepsilon) \leq 2n^\gamma \exp\left(-\frac{\varepsilon(n-1)^\gamma}{10(d^2 + \|\Delta\|_S) \varepsilon \sqrt{2}}\right) + \frac{2ad^2 r}{\varepsilon^2(1-\sqrt{r})^2} \exp\left(-\log \frac{1}{r} (n-1)^\gamma\right)$$

1.2.4 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 d'un ARH(1) connaissant les vecteurs propres de Γ_{X_0}

Soit Π^{k_n} le projecteur orthogonal sur le sous-espace vectoriel engendré par les k_n premiers vecteurs propres $(\phi_j, j = 1, \dots, k_n)$ de Γ_{X_0} et $Y_{in} = \Pi^{k_n}(X_i), i = 1, \dots, n$.

Les opérateurs de la covariance de Y_{0n} et de la covariance croisée de Y_{0n}, Y_{1n} sont :

$$G_n = \Pi^{k_n} \Gamma \Pi^{k_n}, \quad H_n = \Pi^{k_n} \Delta \Pi^{k_n}, \quad H_n^* = \Pi^{k_n} \Delta^* \Pi^{k_n}$$

et les opérateurs respectifs de la covariance empirique :

$$C_n = \Pi^{k_n} \Gamma_n \Pi^{k_n}, \quad D_n = \Pi^{k_n} \Delta_n \Pi^{k_n}, \quad D_n^* = \Pi^{k_n} \Delta_n^* \Pi^{k_n}$$

Pour construire un estimateur de ρ nous faisons les hypothèses suivantes :

- H_1 $\lambda_j > 0$ pour tout j
- H_2 Le déterminant de $(\sum_{i=1}^n \langle X_i, \phi_k \rangle \langle X_i, \phi_l \rangle)_{1 \leq k, l \leq k_n}$ ne s'annule pas p.s

Ces deux hypothèses assurent l'existence de l'inverse de G_n et C_n . Partant de $\Delta^* = \Gamma\rho$, un estimateur de ρ serait

$$\rho_n = C_n^{-1} D_n^*$$

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 1.2.7 *Sous H_1, H_2 et si k_n et $n\lambda_{k_n}^4$ tendent vers l'infini alors on a :*

$$\|\rho_n(X_n) - \rho(X_n)\| \rightarrow 0, \quad n \rightarrow \infty. \quad p.s.$$

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 :

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

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.

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

$$\|\hat{\rho}_n(X_n) - \rho(X_n)\| \rightarrow 0, \quad n \rightarrow \infty. \quad p.s.$$

Chapitre 2

UN MODÈLE AUTOREGRESSIF D'ORDRE UN DANS $C_{[0,1]}$

2.1 Définitions et premiers résultats

Dans cette partie on suppose que l'opérateur aléatoire ρ est un opérateur intégral défini par

$$(2.1) \quad \rho f(t) = \int_0^1 r(s, t) f(s) ds$$

où les v.a r continues symétriques en (s, t) et $\sup_{s, t \in [0, 1]} |r(s, t)| < 1$ p.s.

Soit (X_n) un processus autorégressif $ARC(1)$ défini par (1.1) (chapitre 1). Nous disposons des observations X_1, \dots, X_n et nous définissons $c, c_n, d, d_n, d^*, d_n^*$ par :

$$\begin{aligned} c(s, t) &= E(X_0(s)X_0(t)), & c_n(s, t) &= \frac{1}{n} \sum_{i=1}^n X_i(s)X_i(t) \\ d(s, t) &= E(X_0(s)X_1(t)), & d_n(s, t) &= \frac{1}{n-1} \sum_{i=1}^{n-1} X_i(s)X_{i+1}(t) \\ d^*(s, t) &= E(X_1(s)X_0(t)), & d_n^*(s, t) &= \frac{1}{n-1} \sum_{i=1}^{n-1} X_{i+1}(s)X_i(t) \end{aligned}$$

où $s, t \in [0, 1]$. Toutes ces quantités existent puisque $E\|X_t\|^2 < +\infty$.

Soient $\Gamma, \Gamma_n, \Delta, \Delta_n, \Delta^*, \Delta_n^*$ les opérateurs intégraux de noyaux respectifs $c, c_n, d, d_n, d^*, d_n^*$.

Nous avons par exemple

$$(2.2) \quad \Delta^*(f)(\cdot) = \int_0^1 d^*(y, \cdot) dy, \quad f \in C_{[0,1]}$$

Lemme 2.1.1 *On a les égalités suivantes :*

- a) $d^*(s, t) = \int_0^1 r(s, y)c(y, t)dy$
- b) $r_{i,j}\lambda_j = d_{ij}^*$
- c) $\Delta^* = \Gamma\rho$

Preuve.

- a) Soient $s, t \in [0, 1]$. On a

$$X_i(s)X_{i-1}(t) = \int_0^1 r(s, y)X_{i-1}(y)X_{i-1}(t)dy + \varepsilon_i X_{i-1}(t)$$

En prenant l'espérance de chaque côté et en utilisant le fait que $X_{i-1}(t)$ et ε_i sont indépendants on a le résultat.

- b) Partons de a) et calculons les coefficients de Fourier

$$\int_0^1 \int_0^1 d^*(s, t)\phi_i(s)\phi_j(t)dsdt = \int_0^1 \int_0^1 \left(\int_0^1 r(s, y)c(y, t)dy \right) \phi_i(s)\phi_j(t)dsdt.$$

Les trajectoires de $X_i(t, w)$ étant continues, c l'est aussi et on peut changer l'ordre de l'intégration du côté droit,

$$\begin{aligned} d_{ij}^* &= \int_0^1 \int_0^1 r(s, y)\phi_i(s) \left(\int_0^1 c(y, t)\phi_j(t)dt \right) dsdy \\ &= \int_0^1 \int_0^1 r(s, y)\phi_i(s)\lambda_j\phi_j(y)dsdy \\ &= \lambda_j r_{ij} \end{aligned}$$

- c) Cela vient de l'interprétation de Γ et Δ , mais on va le vérifier ici en utilisant une méthode directe. Soit $f \in C_{[0,1]}$ et $s \in [0, 1]$

$$\Delta^*(f)(s) = \int_0^1 d^*(t, s)f(t)dt$$

$$\begin{aligned}
&= \int_0^1 \left(\int_0^1 r(t, y) c(y, s) dy \right) f(t) dt \\
&= \int_0^1 c(y, s) \left(\int_0^1 r(t, y) f(t) dt \right) dy \\
&= \int_0^1 c(y, s) \rho f(y) dy \\
&= \Gamma[\rho(f)](s)
\end{aligned}$$

2.1.1 Un résultat pour les fonctions Lipschiziennes d'ordre α

Supposons maintenant que les $(X_t, t \in Z)$ sont Lipschiziennes d'ordre α i.e.

$$|X_t(x) - X_t(y)| \leq L(w)|x - y|^\alpha, \quad \alpha > 0, \quad t \in Z$$

Avec cette hypothèse on a $\forall x \in [0, 1]$

$$|X_t(x)| \leq |X_t(x) - X_t(0)| + |X_t(0)| \leq L(w) + |X_t(0)|$$

On suppose aussi que le processus $(X_t, t \in Z)$ est géométriquement α mélangeant et les coefficients de mélangeance vérifient l'inégalité $\alpha(n) \leq bq^n$ où $b > 0$, $0 < q < 1$.

Si on considère le processus centré $(Y_t, t \in Z)$ géométriquement α mélangeant, de coefficients de mélangeance $\alpha(n) \leq bq^n$ où $b > 0$, $0 < q < 1$, vérifiant les deux hypothèses suivantes :

$$\exists m, M \text{ tq } 0 < m \leq EY_t^2 \leq M, \quad t \in Z$$

$$E|Y_t|^k \leq M^{k-2} k! EY_t^2, \quad k \geq 3, \quad t \in Z$$

alors

$$(2.3) \quad P \left(\left| \frac{S_n}{n} \right| \right) \leq (2\sqrt{n} + \eta_1) \exp(-\eta_2 \frac{\varepsilon^2}{5 + \varepsilon} \sqrt{n}) + \eta_3 (1 + \frac{1}{\varepsilon}) n \exp(-\log \frac{1}{q} \sqrt{n})$$

où $S_n = \sum_1^n Y_t$, $\eta_1 = 3 + 1/M$, $\eta_2 = 1/10M$, $\eta_3 = 11[\sqrt{2M} \frac{5}{2}(1 + \frac{4}{5\sqrt{m}})]^{2/5}$.

Si le processus est stationnaire les conditions se réduisent à

$$0 < EY_t^2 \leq M < +\infty$$

$$E|Y|^k \leq M^{k-2} k! EY_t^2, \quad k \geq 3$$

Nous nous intéresserons à deux processus réels, $(|X_t(0)|)$ et $X_t(x) - X_t(y)$, $x, y \in [0, 1]$.

Faisons les hypothèses suivantes :

$$H \left\{ \begin{array}{l} 0 < m(0) \leq E(X_0^2(0)) \leq M(0); \\ E|X_0^k| \leq M^{k-2}(0) k! EY_0^2, \quad k \geq 3; \\ E|X_0^k| \leq M^{k-2}(0) k! EY_0^2, \quad k \geq 3; \\ \forall x, y \in [0, 1] E|X_0(x)X_0(y)|^k \leq M^{k-2}(x, y) k! E(X_0^2(x)X_0^2(y)), \\ \quad k \geq 3 \end{array} \right.$$

H_1 : $\lambda_j > 0$, pour tout j

H_1^* : $\lambda_j = ar^j$; $a > 0$, $0 < r < 1$

H_3 : $P(\lambda_{1n} > \lambda_{2n} > \dots > \lambda_{nn} > 0) = 1$

Remarquons que la première relation est vérifiée automatiquement parce que X_0 est du second ordre. D'autre part on a $E\|X_0\|_L^4 < \infty$ et donc $E(X_0^2(x)X_0^2(y)) < \infty$ pour tout x et y fixe. D'où les deux conditions à vérifiées sont la deuxième et la quatrième qui s'appellent les conditions de Cramer.

Pour les deux types de suites centrées on est sous les conditions du résultat annoncé.

Ajoutons enfin que pour $\varepsilon < 1$, (2.3) devient

$$(2.4) \quad P\left(\left|\frac{S_n}{n}\right|\right) \leq (2\sqrt{n} + \eta_1) \exp(-\eta_2 \varepsilon^2 \sqrt{n}) + \eta_3 \frac{1}{\varepsilon} n \exp(-\log \frac{1}{q} \sqrt{n})$$

où $\eta_2 = \eta_2/6$, $\eta_3 = 2\eta_3$.

Proposition 2.1.1 *Si les fonctions aléatoires X_t sont Lipschitziennes d'ordre α , si le processus $(X_t, t \in \mathbb{Z})$ géométriquement α mélangeant, et si les conditions H sont vérifiées,*

alors pour $\varepsilon \leq 1$, pour tout $n > 0$ et $0 < \beta < 1/4$

$$(2.5) \quad \begin{aligned} P(\|c_n - c\|_c > \varepsilon) &\leq (2\sqrt{n} + \eta_1) [\exp(-\eta_2 \varepsilon^2 \sqrt{n}) + \exp(-\eta_2 \varepsilon^2 16^{-2} n^{1/2-2\beta})] \\ &+ \eta_3 \frac{1}{\varepsilon} (1 + n^\beta) n \exp(-\log \frac{1}{q} \sqrt{n}) + \eta_5 \exp(-n^\beta) \end{aligned}$$

où $\eta_5 = E(e^{L(w)})$. Un tel résultat est vrai aussi pour $\|d_n^* - d^*\|_c$. On suppose donc que $E(e^{L(w)}) < \infty$.

2.2 Estimation de ρ dans $C_{[0,1]}$ connaissant les vecteurs propres de l'opérateur Γ

Soient $(\lambda_j, \phi_j; j \in N^*)$ les éléments propres de l'opérateur intégral Γ et $(\lambda_{jn}; j \in N^*)$ les vecteurs propres de Γ_n . On suppose que les (ϕ_j) sont connus et on se propose d'estimer les λ_j à partir des λ_{jn} . Soient d_{ijn}^* les coefficients de Fourier de d_n^* par rapport aux $(\phi_i \otimes \phi_j, i, j \in N^*)$

$$(2.6) \quad d_{ijn}^* = \int_0^1 \int_0^1 d_n^*(s, t) \phi_i(s) \phi_j(t) ds dt$$

Lemme 2.2.1 *Soit B la constante définie dans le deuxième chapitre. On a*

$$(2.7) \quad E(\sup_{i,j \geq 0} |d_{ijn}^* - d_{ij}^*|)^2 \leq \frac{2B}{n}$$

Proposition 2.2.1 *Soient M_i tel que $\|\phi_i\|_c \leq M_i$. Sous H_1, H_3*

$$P(R_n > \varepsilon) \leq \frac{1}{n} \sum_{i,j=1}^{k_n} \left(\frac{4A}{\lambda_{kn}^2} + C_0 M_i^2 M_j^2 \frac{k_n^4}{\varepsilon^2 \lambda_{kn}^4} \right)$$

où $R_n = \sup_{s,t \in [0,1]} \left| \sum_{i,j=1}^{k_n} (r_{ijn} - r_{ij}) \phi_i(s) \phi_j(t) \right|$ et $C_0 = 16(2B\lambda_1^2 + A\|\Delta\|_L^2)$.

Remarque : En notant $M(k_n) = \sup_{1 \leq i \leq k_n} M_i$, on obtient

$$(2.8) \quad P(R_n > \varepsilon) \leq \frac{1}{n} \left(\frac{4A}{\lambda_{k_n}^2} + C_0 M^4(k_n) \frac{k_n^4}{\varepsilon^2 \lambda_{k_n}^4} \right) k_n^2.$$

Soient

$$r_n(s, t) = \sum_{i, j = k_n + 1}^{+\infty} r_{ij} \phi_i(s) \phi_j(t)$$

$$\|r_n\|_c = \sup_{s, t \in [0, 1]} |r_n(s, t)|$$

Supposons que la série $\sum r_{ij} \phi_i \phi_j$ converge absolument et uniformément i.e.

$$(2.9) \quad r(s, t) = \sum_{i, j = 1}^{+\infty} r_{ij} \phi_i(s) \phi_j(t)$$

et $\|r_n\|_c \rightarrow 0$ si $k_n \rightarrow +\infty$.

Lemme 2.2.2 *Sous H_1, H_3 , si k_n tend vers l'infini et $n^{-1} \lambda_{k_n}^{-4} M^4(k_n) k_n^6 \rightarrow 0$ alors*

$$\sup_{s, t \in [0, 1]} |r_{n, k_n}(s, t) - r(s, t)| \xrightarrow{p} 0$$

Proposition 2.2.2 *Sous les conditions du lemme (2.2.2)*

$$\|\tilde{X}_{n+1} - \check{X}_{n+1}\|_c \xrightarrow{p} 0.$$

On le compare avec le meilleur estimateur probabiliste

$$\check{X}_{n+1} = \int_0^1 r(y, \cdot) X_n(y) dy$$

Enfin il faut remarquer que au lieu du modèle (2.1), où ρ est tel que $\|\rho\|_c < 1$, on considère $\|\rho\|_L < 1$ et ε_t est un L^2 bruit blanc alors on trouve le résultat analogue suivant.

Proposition 2.2.3 Soit $N_t = \|X_t\|_c$. Si les conditions de la Proposition (2.2.2) sont satisfaites et si au lieu de $n^{-1}\lambda_{k_n}^{-4}M^4(k_n)k_n^6 \rightarrow 0$ je suppose $P(n^{-1}N_n^2\lambda_{k_n}^{-4}M^4(k_n)k_n^6 \rightarrow 0) = 1$ alors

$$\|\tilde{X}_{n+1} - \check{X}_{n+1}\|_c \xrightarrow{p} 0.$$

2.3 Estimation de ρ dans le cas général

Soient $(\lambda_{jn}, \phi_{jn}; j \in N^*)$ les éléments propres de l'opérateur intégral Γ_n . Supposons pendant toute cette section le processus est géométriquement α mélangeant, X_t Lipschitziennes d'ordre α , avec $E(e^{L(w)}) < \infty$ et borné. Supposons enfin que les conditions H de la Proposition (2.1.1) sont satisfaites. Ces hypothèses assurent en fait les conditions de la proposition (2.1.1).

Proposition 2.3.1 Sous H_1 et H_3 et si $k_n = o(\log n)$ et pour $n > n_{\gamma_1}$ on a

$$\sup_{1 \leq j \leq k_n} \|\phi_{jn} - \phi'_j\|_c \rightarrow 0. \text{ p.s}$$

et

$$P\left(\sup_{1 \leq i \leq k_n} \|\phi_{jn} - \phi'_j\|_c > \varepsilon\right) \leq A_1 \exp(-A_2 n^{\gamma_1}) + A_3 \exp(-A_4 \varepsilon^2 n^{1/6-\eta}) + A_5 \frac{1}{\varepsilon^2} \exp(-A_6 n^{1/6-\eta})$$

(2.10)

où $A_i, i = 1, 6$, sont des constantes positives et $\eta > 0$, $1/6 < \gamma_1 < 1/4$ peuvent être choisis convenablement.

Proposition 2.3.2 Soient M'_i t.q $\|\phi_{in}\|_c \leq M'_i$. Sous H_1^* H_3 et si $k_n = o(\log n)$ et $M'(k_n) = O(\log n)$

$$P(R_{1,n} > \varepsilon) \leq c_1 \exp(-c_2 n^{1/7})$$

on en tire alors la conclusion que $R_{1,n} \rightarrow 0$ p.s.

Proposition 2.3.3 *Sous les conditions de la Proposition (2.3.2) et si $M(k_n) = O(\log n)$*

$$P(R_{2,n} > \varepsilon) \leq c_3 \exp(-c_4 n^{1/7})$$

pour $n > n^{\gamma_1}$ et on a alors que $R_{2,n} \rightarrow 0$ p.s.

Proposition 2.3.4 *Sous les conditions de la Proposition (2.3.3)*

$$P(R_n > \varepsilon) \leq c_5 \exp(-c_6 n^{1/7})$$

et donc $R_n \rightarrow 0$ p.s.

Proposition 2.3.5 *Sous les conditions de la Proposition (2.3.3)*

$$\|\widehat{r}_{n,k_n} - r\|_c \rightarrow 0 \text{ p.s.}$$

Théorème 2.3.1 *Soit X_t , $t \in Z$ un processus ARC(1) géométriquement α mélangeant et borné vérifiant les hypothèses H_1^* et H_3 . Supposons que les trajectoires X_t sont Lipschitziennes d'ordre α , et vérifient les hypothèses H (voir section 2.1); que la série de Fourier de r converge uniformément et absolument.*

si $M(k_n) = O(\log n)$, $M'(k_n) = O(\log n)$ et si $k_n = o(\log n)$ alors pour $n > n^{\gamma_1}$

$$P(\|\tilde{X}_{n+1} - \check{X}_{n+1}\| > \varepsilon) \leq B_1 \exp(-B_2 n^{1/7})$$

d'où

$$\|\tilde{X}_{n+1} - \check{X}_{n+1}\|_c \rightarrow 0. \text{ p.s.}$$

Chapitre 3

SIMULATION ET ESTIMATION D'UN ARC(1)

3.1 Simulation d'un ARC(1)

La simulation d'un processus à termes dépendants demande indispensablement un modèle. Pour construire un $ARC(1)$ 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 ρ un opérateur intégral symétrique de $C_{[0,1]}$ et tel que $\sup_{x,y} |r(x,y)| < 1$

$$(\rho\phi)(x) = \int_0^1 r(y,x)\phi(y)dy, \quad x \in [0,1], \quad \phi \in H.$$

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 indépendants permet d'établir que $(\varepsilon_i, i \in \mathbb{Z})$ 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.

$$(3.2) \quad X_i = \rho(X_{i+1}) + \varepsilon_i.$$

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

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

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 gaussienne.

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

$$(3.3) \quad e_i(t) = \sqrt{2} \sin \left[\left(i - \frac{1}{2} \right) \pi t \right]$$

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

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

$$(3.4) \quad X(t) = \sqrt{2} \sum_{j=1}^{\infty} Z_j^* \frac{\sin \left[\left(j - \frac{1}{2} \right) \pi t \right]}{\left(j - \frac{1}{2} \right) \pi}$$

où les (Z_j^*) sont des variables aléatoires i.i.d. normales réduites. Il suffit donc d'itérer n fois un mouvement Brownien sur l'intervalle $[0, 1]$.

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 $X_0 = \epsilon_0$. De toute façon comme [38], 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

$$(3.5) \quad X_{i+1}^j = \beta_j X_i^j + \epsilon_{i+1}^j, \quad j = 1, \dots, M$$

qui facilite les calculs. En fait d'après (3.3) 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 Estimation et prévision du processus $ARC(1)$

On discutera les deux cas : le cas où les vecteurs propres de l'opérateur Γ_{X_0} sont connus et le cas où ces vecteurs ne sont pas connus. Bien sûr en pratique le premier cas se présente rarement mais on pourrait par exemple diviser les observations en deux groupes différents ; du premier groupe on estime les vecteurs propres, qui ensuite peuvent être considérés comme de "vrais" vecteurs propres de Γ_{X_0} .

3.2.1 Les vecteurs propres de Γ_{X_0} sont connus

Sachant que les vecteurs propres de Γ_{X_0} sont orthonormaux dans l'espace de Hilbert considéré, calculons les projections des observations $X_i : Y_{in} = \Pi_{k_n}(X_i)$ sur les k_n premiers

vecteurs propres. On peut calculer les opérateurs de covariance de Y_{in} :

$$\begin{aligned} C_n &= \Pi^{k_n} \Gamma_n \Pi^{k_n} \\ D_n &= \Pi^{k_n} \Delta_n^* \Pi^{k_n} \end{aligned}$$

où

$$\begin{aligned} \Gamma_n(f) &= \frac{1}{n} \sum_{i=1}^n \langle X_i, f \rangle X_i \\ \Delta_n^*(f) &= \frac{1}{n-1} \sum_{i=1}^{n-1} \langle X_{i+1}, f \rangle X_i \end{aligned}$$

On construit à partir de C_n et D_n^* un estimateur de ρ , noté ρ_n :

$$\rho_n = C_n^{-1} D_n^*$$

D'après la proposition (2.4.5)

$$\|\rho_n(X_n) - \rho(X_n)\|_2 \xrightarrow{p} 0$$

et d'après la proposition (2.2.4)

$$\|\rho_n(X_n) - \rho(X_n)\|_c \xrightarrow{p} 0, \quad \text{si } n \rightarrow \infty.$$

Supposons que les observations sont des éléments d'un sous-espace vectoriel engendré par les vecteurs propres de Γ_{X_0} . Nous trouvons la matrice M_{C_n} corespondante à la transformation linéaire C_n ; pour cela il faut calculer $C_n(\phi_j)$. Soit $j \leq k_n$

$$\begin{aligned} \Pi^{k_n} \Gamma_n \Pi^{k_n}(\phi_j) &= \frac{1}{n} \sum_{l=1}^n \left(\langle X_l, \phi_j \rangle \sum_{k=1}^{k_n} \langle X_l, \phi_k \rangle \phi_k \right) \\ &= \frac{1}{n} \sum_{l=1}^n \sum_{k=1}^{k_n} \langle X_l, \phi_j \rangle \langle X_l, \phi_k \rangle \phi_k \\ &= \sum_{k=1}^{k_n} \left(\frac{1}{n} \sum_{l=1}^n \langle X_l, \phi_j \rangle \langle X_l, \phi_k \rangle \right) \phi_k \end{aligned}$$

d'où l'élément général de la matrice M_{C_n} de dimension $k_n \times k_n$ est

$$m_{ij}^\Gamma = \frac{1}{n} \sum_{l=1}^n \langle X_l, \phi_j \rangle \langle X_l, \phi_i \rangle$$

De même pour le terme général ($m_{ij}^{\Delta^*}$) de la matrice $M_{D_n^*}$ correspondante de D_n^* nous avons pour $j \leq k_n$

$$\begin{aligned} \Pi^{k_n} \Delta_n \Pi^{k_n}(\phi_j) &= \frac{1}{n-1} \sum_{l=1}^{n-1} \langle X_{l+1}, \phi_j \rangle \sum_{k=1}^{k_n} \langle X_l, \phi_k \rangle \phi_k \\ &= \frac{1}{n-1} \sum_{l=1}^{n-1} \sum_{k=1}^{k_n} \langle X_{l+1}, \phi_j \rangle \langle X_l, \phi_k \rangle \phi_k \\ &= \sum_{k=1}^{k_n} \left(\frac{1}{n-1} \sum_{l=1}^{n-1} \langle X_{l+1}, \phi_j \rangle \langle X_l, \phi_k \rangle \right) \phi_k \end{aligned}$$

et donc le terme général de $M_{D_n^*}$ est

$$m_{ij}^{\Delta^*} = \frac{1}{n-1} \sum_{l=1}^{n-1} \langle X_{l+1}, \phi_j \rangle \langle X_l, \phi_i \rangle$$

D'où la matrice correspondante de la transformation ρ est :

$$\tilde{P}_{k_n} = M_{C_n}^{-1} M_{D_n^*}$$

de terme général (P_{ij}). On aura alors

$$\tilde{X}_{n+1} = C_n^{-1} D_n^*(X_n)$$

Soient $(X_n^i, i = 1, \dots, k_n)$ les coordonnées de l'observation X_n par rapport aux vecteurs $(\phi_1, \dots, \phi_{k_n})$, autrement dit Y_{nn} , on trouve finalement le prédicteur \tilde{X}_{n+1} de X_{n+1} donné par

$$\tilde{X}_{n+1} = \sum_{k=1}^{k_n} \left(\sum_{i=1}^{k_n} P_{ki} X_n^i \right) \phi_k$$

3.2.2 Les vecteurs propres de Γ_{X_0} sont inconnus

Soit (e_1, e_2, \dots) une base orthonormale convenablement choisie pour représenter les observations (X_1, \dots, X_n) . Nous avons

$$X_i = \sum_{k=1}^{\infty} \langle X_i, e_k \rangle e_k, \quad i = 1, 2, \dots, n$$

On ne considérera qu'un nombre fini de termes de cette somme. En fait dans les applications on connaît toujours un nombre fini de données pour chaque observation (X_i) et deuxièmement dans les calculs numériques on sera obligé de travailler sur les matrices d'ordre fini.

Soit m un entier suffisamment grand. Deux cas peuvent se présenter en pratique. On connaît $m \times n$ valeurs $\langle X_i, e_i \rangle$ où e_i est un système de vecteurs orthonormaux et alors on observe en fait

$$(3.6) \quad Y_i = \sum_{k=1}^m \langle X_i, e_k \rangle e_k$$

On connaît $m \times n$ valeurs X_i^j et on considère une approximation par interpolation linéaire. On choisit un système convenable pour avoir la représentation (3.5).

Dans la suite on travaillera la forme (3.5) et dans le sous-espace engendré par les fonctions (e_1, \dots, e_m) .

Soient Γ_n^a, Δ_n^a les opérateurs de la covariance empirique de Y_0 et Y_0, Y_1 respectivement. La matrice corespondante de Γ_n^a est $M_{\Gamma_n^a}$ de terme général

$$\begin{aligned} m_{lk}^a &= \langle \Gamma_n^a(e_k), e_l \rangle \\ &= \frac{1}{n} \sum_{i=1}^n \langle X_i, e_k \rangle \langle X_i, e_l \rangle \end{aligned}$$

$1 \leq l, k \leq m$. Soit $v_i = (v_i^1, \dots, v_i^m)'$ un vecteur propre de la matrice $M_{C_n^a}$ et μ_i la valeur propre correspondante. Soit ψ_i la fonction

$$(3.7) \quad \psi_i = v_i^1 e_1 + \dots + v_i^m e_m$$

En sachant la correspondance entre l'opérateur Γ_n^a et la matrice $M_{\Gamma_n^a}$ nous concluons que ψ_i est un vecteur propre de Γ_n^a . En plus si $\|\psi_i\|_{l_2} = 1$.

Soient alors (ψ_1, \dots, ψ_m) les vecteurs propres de Γ_n^a , (μ_1, \dots, μ_m) les valeurs propres correspondantes en ordre décroissant. En supposant estimer les éléments propres de Γ_{X_0} par (ψ_i, μ_i) nous reprendrons les mêmes calculs faits dans le cas connu pour estimer ρ et trouver une prévision de X_{n+1} .

Ici le système ψ_1, \dots, ψ_m est particulier et puisque $\Gamma_n^a(\psi_i) = \mu_i \psi_i$ alors

$$(\Gamma_n^a)^{-1}(\psi_i) = \frac{1}{\mu_i} \psi_i.$$

Remarquons enfin que les coordonnées de X_i par rapport aux vecteurs propres ψ_j peuvent être calculées facilement connaissant les coordonnées de X_i par rapport aux vecteurs (e_j) et la relation (3.6). On a

$$\langle X_i, \psi_k \rangle = \sum_{j=1}^m v_k^j \langle X_i, e_j \rangle.$$

En particulier on a

$$(3.8) \quad X_n^k = \langle X_n, \psi_k \rangle = \sum_{j=1}^m v_k^j \langle X_n, e_j \rangle$$

qui entraîne

$$X_n = \sum_{j=1}^m \left(\sum_{l=1}^m \langle X_n, e_l \rangle v_l^j \right) \psi_j.$$

Pour les calculs numériques on a utilisé le système orthonormal (avec m fixé)

$$e_i(t) = \begin{cases} \sqrt{m}, & t \in (\frac{k-1}{m}) \\ 0, & t \in [0, 1] \setminus (\frac{k-1}{m}) \end{cases}$$

Pour avoir une base de même style il fallait considérer celle de Haar mais comme nous verrons dans les applications le choix de la base ne semble pas très important.

Soient (x_i^1, \dots, x_i^m) m données de la i^{me} observation. Les coordonnées de cette observation par rapport au système (e_i) choisi ci-dessus sont

$$(X_i^1, \dots, X_i^m)' = \left(\frac{1}{\sqrt{m}}x_i^1, \dots, \frac{1}{\sqrt{m}}x_i^m \right)$$

D'où on peut calculer les matrices $M_{\Gamma_n^a}$ et ses éléments propres. La suite est expliquée comme dans le cas où vecteurs propres sont connus.

3.3 Applications

Avant de passer aux applications numériques, donnons quelques fonctions **R** utiles pour réaliser les simulations.

3.3.1 Fonctions R

► La fonction BaseK2BaseC

★ Description

Etant donnée les coordonnées dans la base de Karhunen-Loève, en 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 simulaires a x .

★ Programme de la fonction BaseK2BaseC

```
BaseK2BaseC <- function(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) * sqrt(2)
  res <- as.fdata(res, dates1 : ncol(x))
  return(res)
}
```

► **La fonction simul.wiener**

★ Description

Simulation d'un processus de Wiener.

★ Usage

simul.wiener($m = 64$, $n = 128$, $m2 = \text{NULL}$)

★ Arguments

m Entier. Nombre de points de discrétisation.

n Entier. Nombre d'observations.

$m2$ Entier. Longueur de l'expansion de Karhunen-Loève ($2m$ par défaut).

★ Détails

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

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 procédure. La valeur par défaut de $m2$ est $2m$.

★ Valeur

Un objet fdata fonctionnel contenant une variable ("var") qui est un processus de Wiener de longueur n avec m points de discrétisation.

★ Exemple

```
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)
```

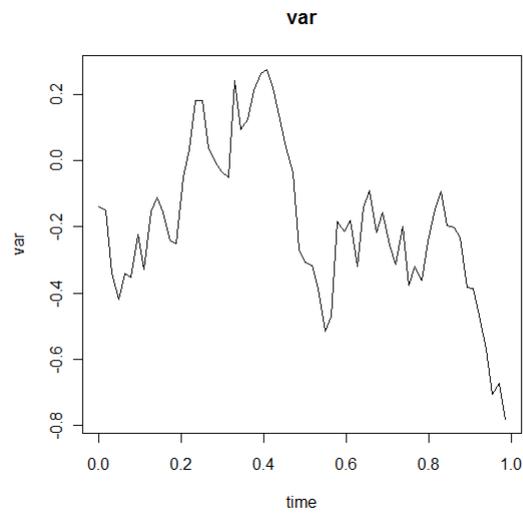


FIGURE 3.1 – Simulation d'un bruit blanc wienerien

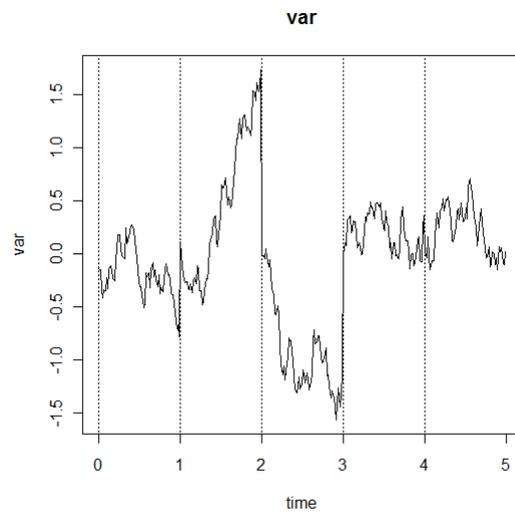


FIGURE 3.2 – Simulation de cinq trajectoires d'un processus wienerien

► **La fonction `simul.far.wiener`**

★ Description

Simulation d'un processus $ARC(1)$ en utilisant le bruit Wienerien.

★ Usage

`simul.far.wiener(m = 64, n = 128,`
`d.rho = diag(c(0.45, 0.9, 0.34, 0.45)), cst1 = 0.05, m2 = NULL)`

★ Arguments

m Entier. Nombre de points de discrétisation.

n Entier. Nombre d'observations.

d.rho. Matrice numérique. L'expression du premier bloc de l'opérateur linéaire dans la base de Karhunen-Loève.

cst1 Coefficient numérique de la perturbation de l'opérateur linéaire.

m2 Entier. Longueur de l'expansion de Karhunen-Loève ($2m$ par défaut).

★ Détails

Cette fonction simule un processus $ARC(1)$ avec un bruit Wienerien. En utilise l'expansion de Karhunen-Loève du bruit. Le processus $ARC(1)$, définie par son opérateur linéaire est calculé dans la base de Karhunen-Loève puis projeté dans la base naturelle. Les paramètres donnés de (*d.rho* et *cst1*) sont exprimés dans la base de Karhunen-Loève.

L'opérateur linéaire, exprimé dans la base de Karhunen-Loève est de la forme *d.rho*

$$\begin{pmatrix} d.rho & 0 \\ 0 & eps.rho \end{pmatrix}$$

où *d.rho* est une matrice. Les deux 0 sont deux blocs de 0, et *eps.rho* est la matrice

diagonale ayant pour diagonale les termes suivants :

$$\varepsilon_{k+1}, \varepsilon_{k+2}, \dots, \varepsilon_{m2}$$

où

$$\varepsilon_t = \frac{cst1}{i^2} + \frac{1 - cst1}{e^i}$$

et k c'est la longueur de la diagonale de $d.rho$.

★ Valeur

Fonction représentant n observations d'un processus $FAR(1)$ avec m points de discrétisation.

★ Exemple

```
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)  
par(mfrow=c(2,1))  
plot(data1,date=1)  
plot(select.fdata(data1,date=1 :5),whole=TRUE,separator=TRUE)
```

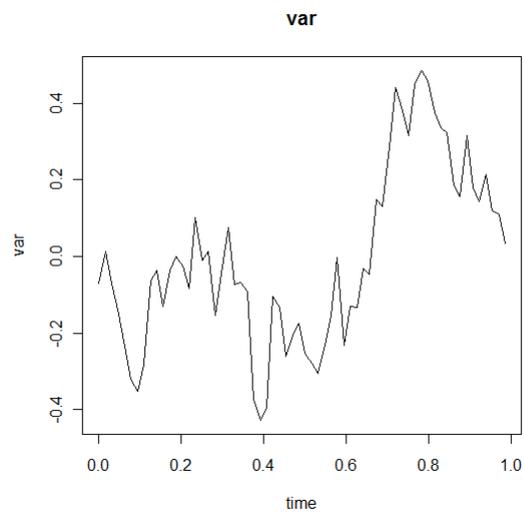


FIGURE 3.3 – Simulation d'une trajectoire d'un modèle FAR(1)

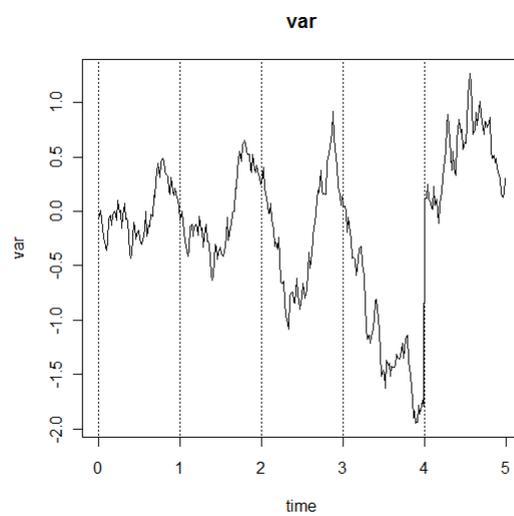


FIGURE 3.4 – Simulation de cinq trajectoires d'un modèle FAR(1)

► **La fonction far**

★ Description

Estimations des paramètres du processus $FAR(1)$ (moyenne et opérateur d'auto-corrélation).

★ Usage

`far(data, kn, center =TRUE, na.rm =TRUE, joined =TRUE)`

★ Arguments

data Objet de données fdata.

kn Un vecteur donnant les valeurs des différents noeuds (dimension de plug-in dans l'algorithme). Si elle n'est pas fournie, la valeur par défaut est un.

center logique. Est-ce que l'observation doivent être centrés.

na.rm logique. Est-ce que une valeur les N.A. besoin d'être enlevés.

★ Détails

un modèle autorégressif fonctionnelle d'ordre 1 ($FAR(1)$) est définie par l'équation suivante :

$$T_n = \rho(T_{n-1}) + \varepsilon_n, n \in Z$$

où T_n et ε_n prennent leurs valeurs dans un espace fonctionnel (par exemple un hilbertienne d'ordre un), et ρ est un opérateur linéaire. ε_n est un bruit blanc fort. Considérons, maintenant le vecteur d'observation, par exemple :

$$(T_{1,n}, \dots, T_{i,n}, \dots, T_{m,n})$$

où chaque $T_{i,n}$ est un éléments dans un espace fonctionnel

★ Valeur

Estimation de opérateur ρ . Cette estimation est effectuée dans un sous-espace de dimension k_n . Une façon de choisir cette dimension est d'utiliser d'abord la fonction de `far.cv` (validation croisée).

► **La fonction `predict.far`**★ Description

Prévision en utilisant $FAR(1)$

★ Usage

`predict.far(object, ..., newdata =NULL, label, na.rm =TRUE)`

★ Arguments

object Objet de type FAR

newdata Une matrice de données (une colonne pour chaque observation) utilisée pour prédire le modèle $FAR(1)$.

label Un vecteur de caractère donnant les dates aux observations prévues.

na.rm Logique. Est-ce que la N.A. à besoin d'être enlevé..

... Des arguments supplémentaires.

★ Détails

Cette fonction calcule à un pas en avant la prédiction du modèle.

★ Valeur

Un objet `fdata`.

3.3.2 Applications sur des données simulées

Programme

```
data22 = simul.far.wiener(m = 64, n = 129,  
d.rho = diag(c(0.45, 0.9, 0.34, 0.45)), cst1 = 0.05, m2 = NULL)  
data1 = select.fdata(data22, date = 1 : 128)  
model3.cv = far.cv(data = data1, ncv = 64, center = FALSE, na.rm = FALSE, joined =  
TRUE)  
k3 = model3.cv$minL2[1]  
DATA2 = far(data1, kn = k3, center = TRUE, na.rm = TRUE, joined = T)  
data3 = predict(DATA2, newdata = data1)  
errors = sqrt(sum((data3[[1]][, 128] - data22[[1]][, 129])2))  
plot(data22, date = 128)  
par(new = T)  
plot(data3, date = 128, col = "red")  
legend(0.1, 1, c("observation", "prévision"), lty = 1 : 2, col = 1 : 2, pch = c(-1, 1))
```

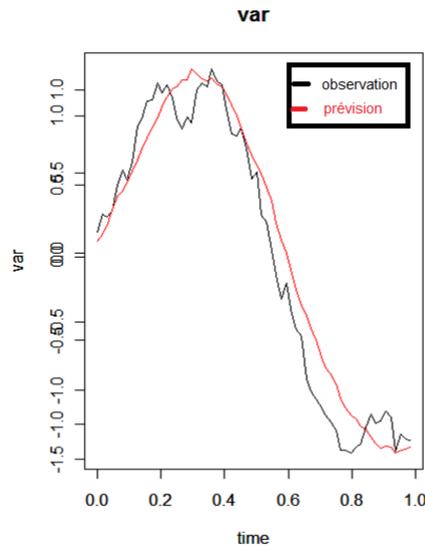


FIGURE 3.5 – Simulation d'une trajectoire FAR(1) et de la prévision

3.3.3 Applications sur des données réelles

► Température de Nottingham

Programme

```

m = matrix(nottem, nrow = 12, ncol = 20)
m1 = m[, 1 : 19]
mm = as.fdata(m1)
model1.cv <- far.cv(data = mm, ncv = 12, center = FALSE, na.rm = FALSE, joined =
TRUE)
k1 = model1.cv$minL2[1]
fun = far(mm, kn = k1)
prev = predict.far(fun, newdata = mm)
errors1 = sqrt(sum((prev[[1]][, 19] - m[, 20])^2))
m3 = as.fdata(m[, 20])
plot(m3, date = 20)
par(new = T)

```

```
plot(prev, date = 20, col = "red")  
legend(0.1, 1, c("observation", "prèvision"), lty = 1 : 2, col = 1 : 2, pch = c(-1, 1))
```

► EL NINO

El Nino est un phénomène climatique dont l'origine est assez mal connue. Contrairement à La Niña, il se traduit pas une hausse de la température à la surface de l'eau (10 mètres environ) à l'est de l'océan Pacifique, autour de l'équateur.

Le nom de El Nino veut dire « *l'enfant* » en espagnol - fait référence à Jésus, car ce phénomène atteint son apogée à l'époque de Noël. Ce sont les pêcheurs péruviens qui avaient attribué ce nom au courant marin froid correspondant.

Ce phénomène a été découvert dans les années 1920 par Sir Gilbert Walker, un physicien anglais ayant mis en évidence l'oscillation australe qui est liée à El Nino.

En temps normal, dans l'océan Pacifique, autour de l'équateur, les alizés soufflent d'est (Amérique) en ouest (Australie et Asie). Ceci a pour effet de pousser les eaux chaudes de surfaces vers l'ouest, où il se produit conséquemment de fortes précipitations (à cause de la chaleur et l'humidité) et une remontée des eaux froides le long des côtes américaines.

Programme

Pour simuler EL NINO et sa prévision

```
f = file.choose(); g = read.table(f, skip = 1);
g[[5]]; tinin = g[[5]][1 : 372]
tt = matrix(tinin, nrow = 12, ncol = 31)
f = as.fdata(tt)
model2.cv <- far.cv(data = f, ncv = 12, center = FALSE, na.rm = FALSE, joined =
TRUE)
k2 = model2.cv$minL2[1]
fun = far(f, kn = k2)
prev = predict.far(fun, newdata = f)
g1 = as.fdata(matrix(g[[5]][1 : 384], nrow = 12, ncol = 32))
gg = matrix(g[[5]][1 : 384], nrow = 12, ncol = 32)
errors2 = sqrt(sum((prev[[1]][, 31] - gg[, 20])2))
print(errors2)
plot(as.fdata(prev[[1]]), date = 31, col = "red")
par(new = T)
plot(g1, date = 32)
legend(0.1, 1, c("observation", "prèvision"), lty = 1 : 2, col = 1 : 2, pch = c(-1, 1))
```

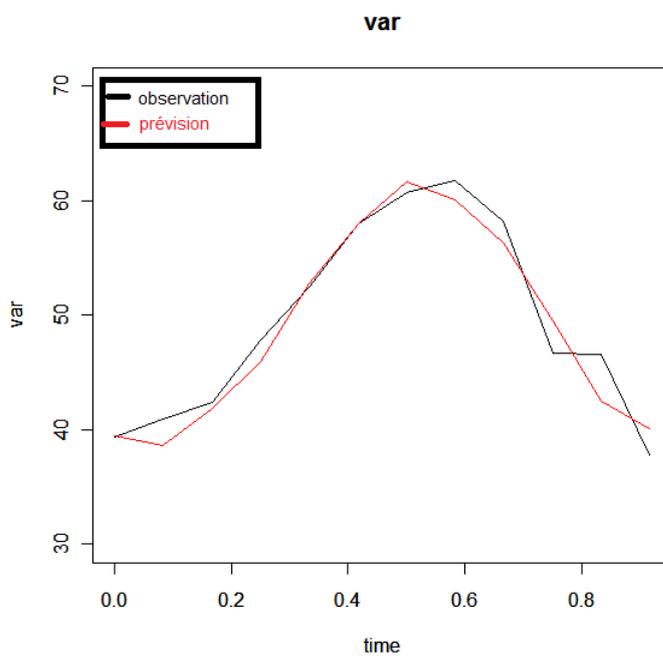


FIGURE 3.6 – Température de Nottingham

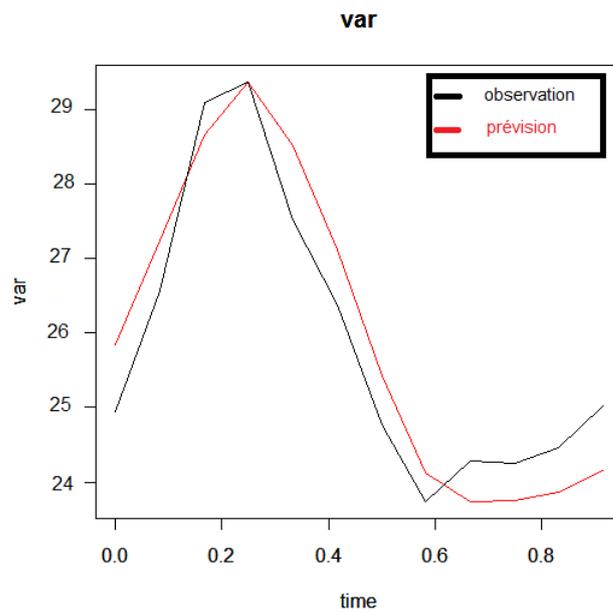


FIGURE 3.7 – Température de la surface (EL NINO)

Comparaison avec d'autres méthodes de prévision

Programme par la méthode noyau fonctionnel

```

fun1 = kerfon(f)
prev1 = predict.kerfon(fun, newdata = f)
ggg = as.fdata(matrix(g[[5]][1 : 384], nrow = 12, ncol = 32))
OB = matrix(g[[5]][1 : 384], nrow = 12, ncol = 32)
errors3 = sqrt(sum(OB[, 32] - (prev1[[1]][, 31]))^2)

```

Programme par la méthode lissage exponentiel (H.W.m)

```

x = ts(tinin, start = 1982, end = 2012, freq = 12)
xlisse = HoltWinters(x, seasonal = "mul")
p = predict(xlisse, n.ahead = 12)
tinini = g[[5]][373 : 384]
xp = ts(tinini, start = 2013, frequency = 12)
errur4 = sqrt(sum((p - xp)^2))

plot(g1, date = 32, ylim = c(24, 29))
par(new = T)
plot(as.fdata(prev1[[1]]), date = 31, col = "green", ylim = c(24, 29))
par(new = T)
plot(as.fdata(prev[[1]]), date = 31, col = "red", ylim = c(24, 29))
par(new = T)
plot(as.fdata(p), col = "blue", ylim = c(24, 29))

```

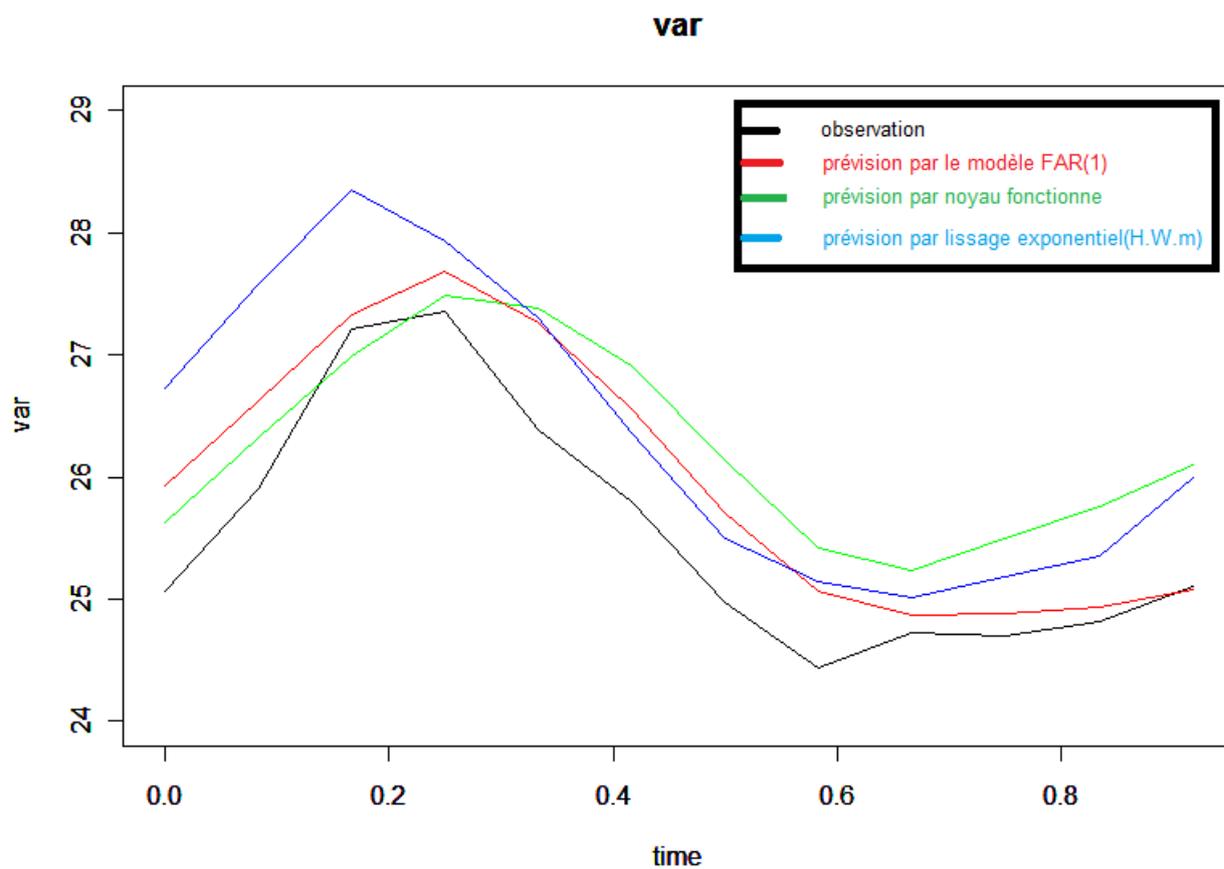


FIGURE 3.8 – Comparaison avec d'autres méthodes de prévision

| Méthode | Erreur de prévision |
|---------------------|---------------------|
| FAR(1) | 1.644827 |
| Noyou fonctionnel | 8.396426 |
| Lissage exponentiel | 3.163454 |

Tableau1

D'après le tableau 1 on s'aperçoit que la prévision par la méthode $FAR(1)$ a donné le meilleur résultat au sens d'erreur quadratique minimal.

Bibliographie

- [1] **AKHIEZER N.I. GLAZMAN I.M.** 1961 Theorie of Linear Operators in Hilbert space **Frederic Ungar Publ. Comp, NewYork**
- [2] **ANTONIADIS A. SAPATINAS T.** Wavelet methods for continuous- time prediction using Hilbert-valued autoregressive processes. *J. Multivariate Anal.*,87, 133-158.
- [3] **ARONSZAJN N.** La théorie générale des noyaux reproduisant et ses applications. *Proc. CAMBRIDGE Philo. Soc*, vol 39,pp. 133-153 .1944.
- [4] **ASH R, GARDNER M.** 1975 Topics in stochastic Processes **ACADZMIC.**
- [5] **BENEDETTIJ.** 1977 On the Nonparametric Estimation of Regression Functions **J. R. Statist. Soc. B, 39, 248 – 253.**
- [6] **BESSE. P, CARDOT. H, STEPHENSON D.** 2003 Autoregressive Forecasting of Some Functional Climatic Variations. *Scandinavian Journal of Statistics* 27 (4), 673.687.
- [7] **BILLINGSLEY.** 1989 Convergence of probability measures. **John Wiley, New York.**
- [8] **BOSQ D.** 1989 Propriétés de l'opérateur d'un processus vectoriel. **John Wiley, New York**
- [9] **BOSQ D.** 1989 Sur les processus autorégressifs dans un espace de Hilbert. **Publ LSTA, 97.**
- [10] **BOSQ D.** 1990 Estimation et prévision d'un processus Hilbertien. **Publ LSTA, 107.**
- [11] **BOSQ D.** 1990 Estimation et prévision d'un ARH(1). (le cas général).**Publ. LSTA, 108.**

-
- [12] **BOSQ D.** 1991 Mean and covariance operator of autoregressive processes in Banach spaces. *Stat. Inference Stoch. Process.*
- [13] **BOSQ D.** 1998 Nonparametric statistics for stochastic processes, volume 110 of *Lecture Notes in Statistics*. Springer-Verlag, New York, second edition. Estimation and prediction.
- [14] **BOSQ D.** 2000 Linear processes in function spaces, volume 149 of *Lecture Notes in Statistics*. Springer-Verlag, New York. Theory and applications.
- [15] **BOSQ D.** 2002 Estimation of mean and covariance operator of autoregressive processes in Banach spaces. *Stat. Inference Stoch. Process.*, 5(3), 287-306.
- [16] **MOKHTARI. F., MOURID. T.** 2003 Prediction of continuous time autoregressive processes via the Reproducing Kernel spaces . *Statistical Inference for Stochastic Processes* 6 : 247-266.
- [17] **MOURID. T.** 1995 Contribution à la statistique des processus autorégression à temps continu. Thèse de Doctorat Es Sciences. Paris 6.
- [18] **MOURID. T.** 2004 Processus autorégressifs hilbertiens à coefficients aléatoires. *Ann.I.S.U.P.* 48, no. 3, 79.85. 60B12
- [19] **Pumo, B.** 1995 Les processus autorégressifs à valeurs dans $C_{[0,1]}$. Estimation de processus discrétisées. *C. R. Acad. Sci. Paris Sér. I Math.*, 320(4), 497-500.
- [20] **Pumo, B.** 1999 Prediction of continuous time processes by $C[0,1]$ -valued autoregressive processes. *Stat. Inference Stoch. Process.*, 3, 1-13.
- [21] **Pumo, B.** 1993 Prédiction d'un processus autorégressif à temps continu. Thèse de doctorat. Université Paris 6.