



**HAL**  
open science

## Estimation et prévision des processus autorégressifs à valeurs dans $D[0,1]$

Layal El Hajj

► **To cite this version:**

Layal El Hajj. Estimation et prévision des processus autorégressifs à valeurs dans  $D[0,1]$ . Annales de l'ISUP, 2014, 58 (3), pp.61-76. hal-03611399

**HAL Id: hal-03611399**

**<https://hal.science/hal-03611399v1>**

Submitted on 17 Mar 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## ESTIMATION ET PRÉVISION DES PROCESSUS AUTORÉGRESSIFS À VALEURS DANS $D[0, 1]$

LAYAL EL HAJJ,

*L.S.T.A., Université Pierre et Marie Curie*

*Résumé.* Afin de prédire un processus à temps continu sur un intervalle, nous introduisons le processus autorégressif d'ordre un à valeurs dans l'espace  $D[0, 1]$  des fonctions càdlàg. Sous certaines hypothèses de régularité, nous montrons la convergence en probabilité du prédictor. Nous illustrons sa performance par une étude de simulation.

*Abstract.* In order to predict a continuous time process on an entire time interval, we introduce the autoregressive process of first order with values in the space  $D[0, 1]$  of cadlag functions. Under some regularity conditions, we show the convergence in probability of the predictor. We illustrate his performance by a simulation study.

**1. Introduction.** Souvent, pour des raisons de simplicité, les modèles paramétriques et non paramétriques construits pour étudier les processus à temps continu ne prennent pas en compte les sauts qui peuvent se produire. Cependant, dans plusieurs domaines, des variations brutales dans les trajectoires des processus ne peuvent pas être simplement négligées. C'est pourquoi, récemment, beaucoup d'attention a été portée sur les processus à sauts, autrement dit, les processus dont les trajectoires sont discontinues. La prise en compte des sauts a conduit au développement de nombreux modèles basés sur des processus discontinus. L'introduction de tels modèles va permettre non seulement de construire des modèles plus réalistes, mais aussi de mieux prédire. Il existe de nombreux exemples de processus à trajectoires discontinues pour les applications, tel que le processus de Poisson qui est un exemple classique de ces processus à sauts.

Dans notre travail, nous nous concentrons sur l'étude de ces processus à sauts. L'approche adoptée consiste à interpréter les trajectoires des processus comme des réalisations de variables fonctionnelles, en d'autres termes, de variables à valeurs dans un espace de dimension infinie. Plus précisément, soit  $\xi = (\xi_t, t \in \mathbb{R})$  un processus réel observé sur un intervalle  $[0, T]$ . Nous nous intéressons principalement à la prédiction de  $\xi$  sur l'intervalle  $[T, T + h]$ ,  $h > 0$ . En supposant que  $T = nh$ , l'approche introduite par Bosq [2] consiste alors à découper le processus  $\xi = (\xi_t, t \in \mathbb{R})$  en  $n$  trajectoires successives de la manière suivante :

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

où  $X = (X_n, n \in \mathbb{Z})$  est un processus à temps discret à valeurs dans un espace fonctionnel convenable. Cette représentation motive l'étude des suites de variables aléatoires à valeurs dans un tel espace. Ainsi le problème est ramené à la prédiction de la variable  $X_{n+1}$ . Pour ce faire, Bosq [2] propose de modéliser ce processus à l'aide des processus autorégressifs Hilbertiens. Cette modélisation a prouvé son efficacité par rapport aux techniques usuelles de prévision sur un certain nombre d'exemples ([3], chapitre 9). De nombreux travaux théoriques et pratiques sont ensuite réalisés tel que Pumo [9] qui s'intéresse au cas des autorégressifs d'ordre un à valeurs dans l'espace de Banach  $C[0, 1]$  des fonctions continues. Nous pouvons également citer, entre autres, Mourid [6] qui considère les autorégressifs d'ordre supérieur, ainsi que Mas [5] qui s'intéresse aux propriétés asymptotiques des autorégressifs Hilbertiens.

Cependant, les trajectoires du processus que nous considérons dans ce travail sont discontinues. Ceci nécessite un espace fonctionnel plus large que ceux considérés dans la littérature. C'est pourquoi on introduit l'espace  $D = D[0, 1]$ , des fonctions continues à droite et ayant des limites à gauche (càdlàg), qui est plus général et suffisamment riche pour ces classes de processus. Nous renvoyons à Billingsley [1] et Taylor [10] pour plus de détails sur cet espace.

L'objectif est alors la prévision d'un processus à temps continu en lui associant un processus autorégressif d'ordre un à valeurs dans  $D$  défini par

$$X_n = \rho(X_{n-1}) + \varepsilon_n, \quad n \in \mathbb{Z},$$

où  $(X_n, n \in \mathbb{Z})$  sont des variables aléatoires à valeurs dans  $D$ , la suite  $(\varepsilon_n, n \in \mathbb{Z})$  est un bruit blanc à valeurs  $D$  et  $\rho$  est un opérateur intégral à noyau donné par

$$(1.1) \quad \rho(x)(t) = \int_0^1 r(s, t)x(s)ds, \quad x \in D, \quad 0 \leq s, t \leq 1,$$

tel que  $\|\rho\|_{\mathcal{L}} = \sup_{\|x\| \leq 1} \|\rho(x)\| < 1$  et  $r$  une fonction symétrique définie sur  $[0, 1]^2$  dont  $\sup_{(s, t) \in [0, 1]^2} |r(s, t)| < 1$ . Le meilleur prédicteur de  $X_{n+1}$  connaissant  $X_1, \dots, X_n$  est

$$X_{n+1}^* = \rho(X_n).$$

Donc la prévision de  $X_{n+1}$  peut s'effectuer en estimant  $\rho$ , ce qui ramène à estimer le noyau  $r$ . Nous montrons, sous certaines conditions, la convergence en probabilité des estimateurs et du prédicteur. L'objectif principal de ce travail est la généralisation des résultats d'estimation et de prédiction des processus autorégressifs à valeurs dans un espace de Banach séparable établis par Bosq [3] et Pumo [9] pour les processus autorégressifs à valeurs dans l'espace  $D$ .

Le plan de l'article est comme suit. La Section 2 contient les notations et définitions. Les résultats de la convergence d'estimateur de l'opérateur  $\rho$  et du prédicteur associé sont présentés dans la Section 3. Dans la Section 4, nous étudions la performance du prédicteur sur des données simulées.

**2. Notations et définitions.** Soit  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probabilisé. Nous considérons que l'espace  $D$  est muni de la topologie de la convergence uniforme, où la norme est définie par

$$\|x\| = \sup_{0 \leq t \leq 1} |x(t)|, \quad x \in D.$$

$D$  est alors un espace de Banach non séparable, ce qui entraîne des problèmes de mesurabilité. Pour surmonter ces difficultés, nous considérons la tribu  $\mathcal{D}$  engendrée par les boules ouvertes. D'après Billingsley [1], celle-ci coïncide avec la tribu de Kolmogorov définie comme étant la plus petite tribu rendant les applications  $\Pi_t : x \rightarrow x(t), x \in D$  mesurables pour tout  $t \in [0, 1]$  (voir [1] pour plus de détails sur la mesurabilité). De plus, si  $\mathbb{E}\|X\| < \infty$  alors l'espérance est définie par  $(\mathbb{E}X)(t) = \mathbb{E}[X(t)]$  pour tout  $t \in [0, 1]$  et  $\mathbb{E}X \in D$  (voir [10]).

Soit  $\rho : D \rightarrow D$  un opérateur linéaire borné mesurable par rapport à  $\mathcal{D}$  (voir [8]) et  $\varepsilon = (\varepsilon_n, n \in \mathbb{Z})$  un bruit blanc fort à valeurs dans  $D$  défini par une suite de variables aléatoires i.i.d. telles que

$$\mathbb{E}[\varepsilon_n(s)] = 0, \quad 0 \leq s \leq 1 \quad \text{et} \quad \mathbb{E}\|\varepsilon_n\|^2 < \infty.$$

Soient  $c_0$  et  $c_1$  les fonctions de covariance et covariance croisée définies respectivement, pour tout  $0 \leq s, t \leq 1$ , par

$$c_0(s, t) = \mathbb{E}[X_0(s)X_0(t)] \quad \text{et} \quad c_1(s, t) = \mathbb{E}[X_1(s)X_0(t)].$$

Leurs estimateurs naturels sont les fonctions de covariance empiriques définies pour tout  $0 \leq s, t \leq 1$ , par

$$c_{0,n}(s, t) = \frac{1}{n} \sum_{i=1}^n X_i(s)X_i(t) \quad \text{et} \quad c_{1,n}(s, t) = \frac{1}{n-1} \sum_{i=1}^{n-1} X_i(s)X_{i-1}(t), \quad n \in \mathbb{Z}.$$

Pour toute fonction  $f \in D[0, 1]$ , l'opérateur intégral de noyau  $c_0$  est donné par

$$C_0(f)(x) = \int_0^1 c_0(x, y)f(y)dy.$$

D'une manière similaire on peut écrire les opérateurs de covariance  $C_1, C_{0,n}$  et  $C_{1,n}$  en utilisant respectivement les noyaux  $c_1, c_{0,n}$  et  $c_{1,n}$ . L'estimation des éléments propres de  $C_0$  jouent un rôle important dans la construction d'un estimateur consistant de  $\rho$ . Soit  $(\lambda_j, \phi_j, j = 1, 2, \dots)$  les éléments propres de  $C_0$ . Leurs estimateurs naturels sont  $(\lambda_{jn}, \phi_{jn}, j = 1, 2, \dots)$ , les éléments propres de  $C_{0,n}$ .

D'après (1.1), estimer l'opérateur intégral  $\rho$  revient à estimer le noyau  $r$ . L'estimateur de  $r$  est donné par

$$\begin{aligned}\hat{r}_{n,k_n}(s,t) &= \sum_{i,j=1}^{k_n} \hat{r}_{ijn}(\phi_{in} \otimes \phi_{jn}) \\ &= \sum_{i,j=1}^{k_n} \hat{r}_{ijn} \phi_{in}(s) \phi_{jn}(t), \quad 0 \leq s, t \leq 1,\end{aligned}$$

où  $k_n$  est un entier positif tel que  $k_n < n$  et  $\hat{r}_{ijn}$  est l'estimateur de  $r_{ij}$  avec

$$r_{ij} = \int_0^1 \int_0^1 r(s,t) \phi_i(s) \phi_j(t) ds dt.$$

L'estimateur de  $\rho(X_n)$  correspondant est

$$\hat{\rho}_{n,k_n}(X_n) = \int_0^1 \hat{r}_{n,k_n}(\cdot, y) X_n(y) dy.$$

**3. Estimation et prévision.** Nous supposons dans la suite que la fonction de covariance  $c_0$  est continue. Soit

$$a_j = \begin{cases} 2\sqrt{2}(\lambda_1 - \lambda_2)^{-1}, & j = 1; \\ 2\sqrt{2} \max[(\lambda_{j-1} - \lambda_j)^{-1}, (\lambda_j - \lambda_{j+1})^{-1}], & j \geq 2. \end{cases}$$

Tout d'abord nous regroupons les hypothèses dont nous avons besoin pour établir nos résultats :

H.1 Pour tout  $n \in \mathbb{Z}$ ,  $\mathbb{E}\|X_n\|^4 = \mathbb{E}\|X_0\|^4 < \infty$ ;

H.2 Pour tout  $n \in \mathbb{Z}$  et  $0 \leq s, t \leq 1$ , nous avons

(a)  $\mathbb{E}[X_n(t) - X_n(s)]^2 \leq M_1 |t - s|^{\alpha_1}$ ,  $\alpha_1 > 1/2$ ,

(b)  $\mathbb{E}[X_n(u) - X_n(t)]^2 [X_n(t) - X_n(s)]^2 \leq M_2 |u - s|^{\alpha_2}$ ,  $\alpha_2 > 1/2$ ,

(c)  $\mathbb{E}[X_n(t) - X_n(s)]^4 \leq M_3 |t - s|^{\alpha_3}$ ,  $\alpha_3 \geq 1$ ,

où les  $M_i$ ,  $i = 1, 2, 3$  sont des constantes positives;

H.3  $\lambda_1 > \lambda_2 > \dots > 0$  et  $\lambda_{1n} > \lambda_{2n} > \dots > \lambda_{k_n n} > 0$  p.s.;

H.4 Pour tout  $n \in \mathbb{Z}$ , nous avons

$$\lim_{n \rightarrow +\infty} \frac{n^{\beta\mu} \lambda_{k_n}^4}{(\log_2 n)^2} = +\infty \quad \text{et} \quad \lim_{n \rightarrow +\infty} \frac{n^{\beta\mu} \lambda_{k_n}^2}{(\sum_{j=1}^{k_n} a_j)^2 (\log_2 n)^2} = +\infty,$$

avec  $\mu > 0$ ,  $0 < \beta < 1/2$  et  $\log_2 n = \log(\log n)$ .

H.5 Pour tout  $0 \leq s, t \leq 1$ ,  $r(s,t) = \sum_{i,j=1}^{+\infty} r_{ij} \phi_i(s) \phi_j(t)$  converge absolument et uniformément.

Commentaires : L'hypothèse *H.1* assure une régularité suffisante des trajectoires du processus *ARD(1)*. L'hypothèse *H.2*, exprimée en fonction des moments des accroissements du processus, est nécessaire pour montrer la convergence de la fonction de covariance empirique. Ce type de conditions est fréquemment utilisé pour établir une convergence dans l'espace *D*, par exemple voir [1] et [7]. Pour montrer la convergence du prédicteur vers la meilleure prévision possible, il est nécessaire de contrôler la décroissance vers zéro des valeurs propres de l'opérateur de covariance, ce qui justifie les conditions *H.3* et *H.4*. Enfin l'hypothèse *H.5* assure que *r* peut être exprimée par les vecteurs propres. Le même genre de conditions a été introduit dans [3] et [9] et permet d'obtenir des vitesses de convergence.

REMARQUE. Si nous avons  $k_n = o(\log n)$  et  $\lambda_j = ar^j$  avec  $a > 0, 0 < r < 1$ , alors la condition *H.4* est vérifiée.

A titre d'exemple d'un processus *ARD(1)* qui vérifie la condition *H.2*, nous considérons le suivant.

EXEMPLE. Soit  $(X_n)$  un processus autorégressif où  $\rho$  est un opérateur à noyau défini par (1.1) et  $(\varepsilon_n)$  est un bruit blanc à valeurs dans *D* donné par

$$\varepsilon_n(t) = N_{n+t} - N_n - \lambda t, \quad 0 \leq s \leq 1, \quad n \in \mathbb{Z},$$

avec *N* un processus de Poisson bilatéral d'intensité  $\lambda$ . Nous observons que

$$\mathbb{E}[X_n(s) - X_n(t)]^2 \leq 2\mathbb{E}\left[\int_0^1 (r(u,s) - r(u,t))X_{n-1}(u)du\right]^2 + 2\mathbb{E}[\varepsilon_n(t) - \varepsilon_n(s)]^2.$$

D'après Cauchy-Schwarz,

$$\mathbb{E}\left[\int_0^1 (r(u,s) - r(u,t))X_{n-1}(u)du\right]^2 \leq \int_0^1 (r(u,s) - r(u,t))^2 du \quad \mathbb{E}\left[\int_0^1 X_{n-1}^2(u)du\right].$$

Supposons que le noyau *r* vérifie la condition

$$(r(u,s) - r(u,t)) \leq |t - s|^{1/2}.$$

D'autre part, la stationnarité et l'indépendance des accroissements du processus de Poisson implique que pour  $0 \leq s < t \leq 1$ ,

$$\mathbb{E}[\varepsilon_n(t) - \varepsilon_n(s)]^2 = \lambda(t - s),$$

alors  $(X_n)$  vérifie la condition *H.2(a)* pour  $\alpha_1 = 1$ . Nous procédons d'une manière similaire pour vérifier *H.2(c)* en tenant compte que pour  $0 \leq s < t < u \leq 1$ , nous avons

$$\mathbb{E}[\varepsilon_n(t) - \varepsilon_n(s)]^4 \leq (\lambda + 3\lambda^2)(t - s).$$

De plus, d'après Cauchy-Schwartz, nous avons

$$\mathbb{E}[|X_n(u) - X_n(t)|^2 |X_n(t) - X_n(s)|^2] \leq (\mathbb{E}[|X_n(u) - X_n(t)|^4])^{1/2} (\mathbb{E}[|X_n(t) - X_n(s)|^4])^{1/2}.$$

En traitant chaque terme d'une manière similaire à H.2(c), la condition H.2(b) est vérifiée.  $\square$

Dans un premier temps, nous montrons la convergence de la fonction de covariance empirique.

PROPOSITION 3.1. *Sous les hypothèses H.1 et H.2, nous avons pour tout  $0 \leq s, t \leq 1$*

$$\lim_{n \rightarrow \infty} \mathbb{E} \left( \sup_{0 \leq s, t \leq 1} |c_{0,n}(s, t) - c_0(s, t)| \right)^2 = 0.$$

Plus précisément,

$$\mathbb{E} \left( \sup_{0 \leq s, t \leq 1} |c_{0,n}(s, t) - c_0(s, t)| \right)^2 = \mathcal{O} \left( \frac{(\log_2(n))^2}{n^{\beta\mu}} \right), \quad 0 < \beta < 1/2,$$

avec  $\mu = \min\{\frac{2\alpha_2-1}{2}, \frac{\alpha_3-1}{2}, \alpha_1\} > 0$  et  $\log_2(n) = \log(\log n)$ .

Un résultat similaire est obtenu pour  $c_{1,n}$ . Le théorème suivant montre une convergence en probabilité de  $\hat{r}_{n,k_n}$ .

THÉORÈME 3.1. *Sous les hypothèses H.1-H.4, nous avons pour tout  $\varepsilon > 0$ ,*

$$\lim_{n \rightarrow \infty} \mathbb{P} \left( \sup_{(s,t) \in [0,1]^2} |\hat{r}_{n,k_n}(s, t) - r(s, t)| > \varepsilon \right) = 0.$$

COROLLAIRE 3.1. *Sous les hypothèses H.1-H.5 et si  $(X_n)$  est borné, nous avons pour tout  $\varepsilon > 0$ ,*

$$(3.1) \quad \lim_{n \rightarrow \infty} \mathbb{P} (\|\hat{\rho}_{n,k_n}(X_n) - \rho(X_n)\| > \varepsilon) = 0.$$

**4. Simulations.** Pour simuler des trajectoires d'un processus  $ARD(1)$  et montrer la performance du prédicteur proposé, nous appliquons les techniques de simulation proposées par Damons et Guillas [4] et implémentées dans le logiciel R dans le package `far`. Leur méthode permet de généraliser celle de Pumo [9] et de simuler un bruit blanc non gaussien. L'opérateur  $\rho$  est exprimé dans une base de vecteurs propres de l'opérateur de covariance  $C_0$ . Nous simulons  $n = 80$  trajectoires successives d'un processus  $ARD(1)$  où le bruit blanc est composé des accroissements d'un processus de Poisson. Chaque trajectoire est calculée en  $m = 50$  points de discrétisation de  $[0, 1]$ . Pour des raisons de lisibilité, nous illustrons  $n = 15$  trajectoires successives dans la Figure 1.

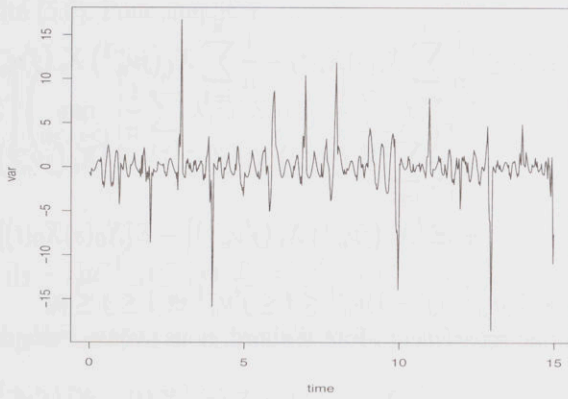


FIGURE 1. Simulation d'un  $ARD(1)$ .

La Figure 2 représente une trajectoire du processus  $ARD(1)$  (en trait continu) et son prédicteur estimé à partir des observations précédentes (en pointillés). La dimension du sous-espace sur lequel on projette nos observations est choisie par validation croisée. La valeur retenue est  $k_n = 4$ .

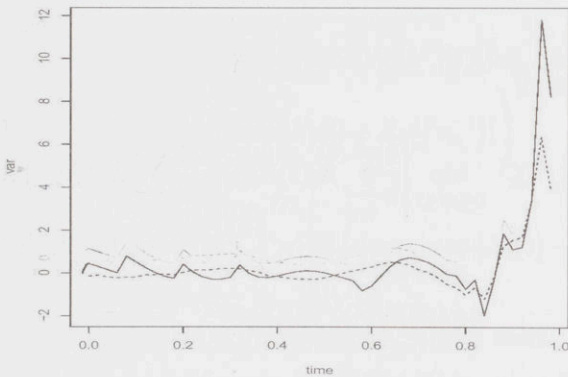


FIGURE 2. Trajectoire d'un processus  $ARD(1)$  et son prédicteur.



**5. Preuves.** *Preuve de la Proposition 3.1 :* Nous avons

$$\begin{aligned} |c_{0,n}(s,t) - c_0(s,t)| &\leq \left| \frac{1}{n} \sum_{i=1}^n X_i(s)X_i(t) - \frac{1}{n} \sum_{i=1}^n X_i(j\nu_n^{-1})X_i(j'\nu_n^{-1}) \right| \\ &+ \left| \frac{1}{n} \sum_{i=1}^n X_i(j\nu_n^{-1})X_i(j'\nu_n^{-1}) - \mathbb{E}[X_i(j\nu_n^{-1})X_i(j'\nu_n^{-1})] \right| \\ &+ \left| \mathbb{E}[X_i(j\nu_n^{-1})X_i(j'\nu_n^{-1})] - \mathbb{E}[X_0(s)X_0(t)] \right|, \end{aligned}$$

pour  $(j-1)\nu_n^{-1} \leq s \leq j\nu_n^{-1}$ ,  $(j'-1)\nu_n^{-1} \leq t \leq j'\nu_n^{-1}$  et  $1 \leq j \leq \nu_n$ .

Traisons chaque terme séparément. Tout d'abord, nous avons

$$(5.1) \quad \begin{aligned} X_i(s)X_i(t) - X_i(j\nu_n^{-1})X_i(j'\nu_n^{-1}) &= X_i(s)[X_i(t) - X_i(j'\nu_n^{-1})] \\ &+ X_i(j'\nu_n^{-1})[X_i(s) - X_i(j\nu_n^{-1})]. \end{aligned}$$

Tenant compte de la décomposition (5.1) et de la condition *H.2(a)*, et d'après Cauchy-Schwartz, nous obtenons

$$\left| \mathbb{E}[X_i(j\nu_n^{-1})X_i(j'\nu_n^{-1})] - \mathbb{E}[X_0(s)X_0(t)] \right| \leq 2(\mathbb{E}\|X_0\|^2)^{1/2} M_1 \nu_n^{-\alpha_1/2}.$$

D'autre part,

$$\left| \frac{1}{n} \sum_{i=1}^n X_i(j\nu_n^{-1})X_i(j'\nu_n^{-1}) - \mathbb{E}[X_i(j\nu_n^{-1})X_i(j'\nu_n^{-1})] \right| = |c_{0,n}(j\nu_n^{-1}, j'\nu_n^{-1}) - c_0(j\nu_n^{-1}, j'\nu_n^{-1})|$$

Notons  $\|x\|_u = \sup_{0 \leq s, t \leq 1} |x(s,t)|$ . Alors nous avons

$$\begin{aligned} \|c_{0,n} - c_0\|_u &\leq \sup_{0 \leq s, t \leq 1} \left| \frac{1}{n} \sum_{i=1}^n X_i(s)X_i(t) - \frac{1}{n} \sum_{i=1}^n X_i(j\nu_n^{-1})X_i(j'\nu_n^{-1}) \right| \\ &+ \max_{1 \leq j, j' \leq \nu_n} |c_{0,n}(j\nu_n^{-1}, j'\nu_n^{-1}) - c_0(j\nu_n^{-1}, j'\nu_n^{-1})| \\ &+ 2(\mathbb{E}\|X_0\|^2)^{1/2} M_1 \nu_n^{-\alpha_1/2}. \end{aligned}$$

Alors d'après le Théorème 6.10 de Bosq [3], nous obtenons

$$(5.2) \quad \begin{aligned} \mathbb{E}[\|c_n - c_0\|_u^2] &\leq 2\mathbb{E} \left[ \left( \sup_{0 \leq s, t \leq 1} \left| \frac{1}{n} \sum_{i=1}^n X_i(s)X_i(t) - \frac{1}{n} \sum_{i=1}^n X_i(j\nu_n^{-1})X_i(j'\nu_n^{-1}) \right| \right)^2 \right] \\ &+ 4M_0 \frac{\nu_n^2}{n} + 8\mathbb{E}\|X_0\|^2 M_1^2 \nu_n^{-\alpha_1}, \end{aligned}$$

où  $M_0$  est une constante. Il suffit maintenant de considérer le premier terme du membre droit de l'inégalité (5.2). Pour simplifier, on note

$$A_n = \mathbb{E} \left[ \left( \sup_{0 \leq s, t \leq 1} \left| \frac{1}{n} \sum_{i=1}^n X_i(s) X_i(t) - \frac{1}{n} \sum_{i=1}^n X_i(j\nu_n^{-1}) X_i(j'\nu_n^{-1}) \right| \right)^2 \right].$$

Posons

$$J_n = [(j-1)\nu_n^{-1}, j\nu_n^{-1}] \text{ et } J'_n = [(j'-1)\nu_n^{-1}, j'\nu_n^{-1}], \quad 1 \leq j, j' \leq \nu_n.$$

D'après la décomposition (5.1) et en utilisant l'inégalité de Jensen, nous obtenons

$$\begin{aligned} A_n &\leq \frac{1}{n} \sum_{i=1}^n \mathbb{E} \left[ \left( \sup_{0 \leq s \leq 1} |X_i(s)| \sup_{0 \leq t \leq 1} |X_i(t) - X_i(j'\nu_n^{-1})| \right)^2 \right] \\ &\quad + \frac{1}{n} \sum_{i=1}^n \mathbb{E} \left[ \left( \max_{1 \leq j \leq \nu_n} |X_i(j'\nu_n^{-1})| \sup_{0 \leq s \leq 1} |X_i(s) - X_i(j\nu_n^{-1})| \right)^2 \right] \\ &\leq \frac{1}{n} \sum_{i=1}^n \mathbb{E} \left[ \left( \|X_i\| \max_{1 \leq j' \leq \nu_n} \sup_{t \in J'_n} |X_i(t) - X_i(j'\nu_n^{-1})| \right)^2 \right] \\ &\quad + \frac{1}{n} \sum_{i=1}^n \mathbb{E} \left[ \left( \max_{1 \leq j' \leq \nu_n} |X_i(j'\nu_n^{-1})| \max_{1 \leq j \leq \nu_n} \sup_{s \in J_n} |X_i(s) - X_i(j\nu_n^{-1})| \right)^2 \right]. \end{aligned}$$

Par Cauchy-Schwarz,

$$\begin{aligned} A_n &\leq \frac{1}{n} \sum_{i=1}^n (\mathbb{E} \|X_i\|^4)^{1/2} \left[ \mathbb{E} \left( \max_{1 \leq j' \leq \nu_n} \sup_{t \in J'_n} |X_i(t) - X_i(j'\nu_n^{-1})| \right)^4 \right]^{1/2} \\ &\quad + \frac{1}{n} \sum_{i=1}^n \left( \mathbb{E} \left( \max_{1 \leq j' \leq \nu_n} |X_i(j'\nu_n^{-1})| \right)^4 \right)^{1/2} \left[ \mathbb{E} \left( \max_{1 \leq j \leq \nu_n} \sup_{s \in J_n} |X_i(s) - X_i(j\nu_n^{-1})| \right)^4 \right]^{1/2}. \end{aligned}$$

Or d'après le lemme de Paulauskas et Juknevičienė [7] nous avons

$$\begin{aligned} \mathbb{E} \left[ \sup_{t \in J'_n} |X_i(t) - X_i(j'\nu_n^{-1})| \right]^4 &\leq 4 \mathbb{E} [|X_i(j\nu_n^{-1}) - X_i((j-1)\nu_n^{-1})|]^4 \\ &\quad + 4 \mathbb{E} \left[ \sup_{t \in J'_n} \min \{ |X_i(t) - X_i((j'-1)\nu_n^{-1})|, |X_i(j'\nu_n^{-1}) - X_i(t)| \} \right]^4. \end{aligned}$$

D'après Billingsley [1], sous la condition *H.2(b)* nous avons

$$\mathbb{E} \left[ \sup_{t \in J'_n} \min \{ |X_i(t) - X_i((j' - 1)\nu_n^{-1})|, |X_i(j'\nu_n^{-1}) - X_i(t)| \} \right]^4 \leq \log_2(2n)^4 \nu_n^{-2\alpha_2}.$$

Par ailleurs, la condition *H.2(c)* implique

$$4\mathbb{E} [|X_i(j\nu_n^{-1}) - X_i((j - 1)\nu_n^{-1})|]^4 \leq 4M_3\nu_n^{-\alpha_3}.$$

En regroupant toutes les inégalités, on obtient

$$\begin{aligned} A_n &\leq 2 (\mathbb{E}\|X_0\|^4)^{1/2} \left[ \nu_n \left( 4 (\log_2(2n))^4 \nu_n^{-2\alpha_2} + 4M_3\nu_n^{-\alpha_3} \right) \right]^{1/2} \\ &\leq 2 (\mathbb{E}\|X_0\|^4)^{1/2} \left[ 4 (\log_2(2n))^4 \nu_n^{\frac{1-2\alpha_2}{2}} + 4M_3\nu_n^{\frac{1-\alpha_3}{2}} \right]. \end{aligned}$$

On déduit de l'ensemble des inégalités précédentes ci-dessus,

$$\begin{aligned} \mathbb{E}\|c_{0,n} - c_0\|_u^2 &\leq 8 (\mathbb{E}\|X_0\|^4)^{1/2} \left[ (\log_2(2n))^4 \nu_n^{\frac{1-2\alpha_2}{2}} + M_3\nu_n^{\frac{1-\alpha_3}{2}} \right] + 4M_0 \frac{\nu_n^2}{n} \\ &\quad + 8M_1^2 \mathbb{E}\|X_0\|^2 \nu_n^{-\alpha_1}, \end{aligned}$$

avec  $\alpha_1 > 1/2$ ,  $\alpha_2 > 1/2$  et  $\alpha_3 \geq 1$ . Alors

$$\begin{aligned} \mathbb{E}\|c_{0,n} - c_0\|_u^2 &\leq \mathcal{O} \left( (\log_2(n))^2 \nu_n^{\frac{1-2\alpha_2}{2}} \right) + \mathcal{O} \left( \nu_n^{\frac{1-\alpha_3}{2}} \right) + \mathcal{O} \left( \frac{\nu_n^2}{n} \right) + \mathcal{O} (\nu_n^{-\alpha_1}) \\ &\leq \mathcal{O} \left( (\log_2(n))^2 \nu_n^{-\mu} \right), \end{aligned}$$

avec  $\mu = \min \left\{ \frac{2\alpha_2-1}{2}, \frac{\alpha_3-1}{2}, \alpha_1 \right\} > 0$ . En choisissant  $\nu_n = n^\beta$ ,  $0 < \beta < 1/2$ , nous obtenons

$$\mathbb{E}\|c_{0,n} - c_0\|_u^2 = \mathcal{O} \left( \frac{(\log_2(n))^2}{n^{\beta\mu}} \right), \quad 0 < \beta < 1/2, \quad \mu > 0.$$

□

Soit

$$\phi'_j = \text{sgn} < \phi_{jn}, \phi_j > \phi_j,$$

avec  $\text{sgn}(u) = \mathbb{1}_{u \geq 0} - \mathbb{1}_{u < 0}$ ,  $u \in \mathbb{R}$ . Le lemme suivant montre une convergence en probabilité de  $\phi_{jn}$ .

LEMME 5.1. *Soit  $X_n$  un  $ARD(1)$ . Sous les hypothèses *H.1-H.5* nous avons pour tout  $\varepsilon > 0$ ,*

$$(5.3) \quad \lim_{n \rightarrow \infty} \mathbb{P} \left( \sup_{j=1, \dots, k_n} \|\phi_{jn} - \phi'_j\| > \varepsilon \right) = 0.$$

*Preuve* : Comme la fonction de covariance  $c_0$  est supposée continue, nous avons

$$\lambda_j \phi'_j(x) = \int_0^1 c_0(x, y) \phi'_j(y) dy,$$

et

$$\lambda_{jn} \phi_{jn}(x) = \int_0^1 c_{0,n}(x, y) \phi_{jn}(y) dy.$$

Alors nous obtenons

$$\begin{aligned} \lambda_{jn} [\phi_{jn}(x) - \phi'_j(x)] &= \int_0^1 [c_{0,n}(x, y) - c_0(x, y)] \phi_{jn}(y) dy \\ &\quad + \int_0^1 c_0(x, y) [\phi_{jn}(y) - \phi'_j(y)] dy + (\lambda_j - \lambda_{jn}) \phi'_j(x). \end{aligned}$$

D'après Cauchy-Schwartz,

$$\begin{aligned} \|\phi_{jn} - \phi'_j\| &\leq \frac{1}{\lambda_{jn}} [\|c_{0,n} - c_0\|_u + \|c_0\|_u \|\phi_{jn} - \phi'_j\|_{L^2} + |\lambda_j - \lambda_{jn}| \|\phi'_j\|] \\ (5.4) \quad &\leq \frac{1}{\lambda_{jn}} \left[ \|c_{0,n} - c_0\|_u + \|c_0\|_u \|\phi_{jn} - \phi'_j\|_{L^2} + |\lambda_j - \lambda_{jn}| \frac{\|c_0\|_u}{\lambda_j} \right]. \end{aligned}$$

Soit  $E_n = \left\{ |\lambda_{k_n} - \lambda_{k_n n}| \leq \frac{\lambda_{k_n}}{2} \right\}$ . Alors

$$\mathbb{P} \left( \sup_{j=1, \dots, k_n} \|\phi_{jn} - \phi'_j\| > \varepsilon \right) \leq \mathbb{P} \left( \sup_{j=1, \dots, k_n} \|\phi_{jn} - \phi'_j\| > \varepsilon, E_n \right) + \mathbb{P}(E_n^c).$$

D'après la décomposition (5.4), on trouve

$$\begin{aligned} \mathbb{P} \left( \sup_{j=1, \dots, k_n} \|\phi_{jn} - \phi'_j\| > \varepsilon, E_n \right) &\leq \mathbb{P} \left( \|c_{0,n} - c_0\|_u > \frac{\lambda_{k_n} \varepsilon}{6} \right) \\ &\quad + \mathbb{P} \left( \sup_{j=1, \dots, k_n} \|\phi_{jn} - \phi'_j\|_{L^2} > \frac{\lambda_{k_n} \varepsilon}{6 \|c_0\|_u} \right) \\ &\quad + \mathbb{P} \left( \sup_{j=1, \dots, k_n} |\lambda_j - \lambda_{jn}| > \frac{\lambda_{k_n}^2 \varepsilon}{6 \|c_0\|_u} \right). \end{aligned}$$

Or d'après l'inégalité de Chebyshev,

$$\mathbb{P} \left( \|c_{0,n} - c_0\|_u > \frac{\lambda_{k_n} \varepsilon}{6} \right) \leq \frac{36}{\lambda_{k_n}^2 \varepsilon^2} \mathbb{E} [\|c_{0,n} - c_0\|_u^2].$$

D'après [9], pour tout entier  $j \geq 1$  nous avons

$$|\lambda_{jn} - \lambda_j| \leq \|c_{0,n} - c_0\|_u,$$

alors

$$\begin{aligned} \mathbb{P} \left( \sup_{j=1, \dots, k_n} |\lambda_j - \lambda_{jn}| > \frac{\lambda_{k_n}^2 \varepsilon}{6\|c\|_u} \right) &\leq \mathbb{P} \left( \|c_{0,n} - c_0\|_u > \frac{\lambda_{k_n}^2 \varepsilon}{6\|c_0\|_u} \right) \\ &\leq \frac{36\|c_0\|_u^2}{\lambda_{k_n}^4 \varepsilon^2} \mathbb{E} [\|c_{0,n} - c_0\|_u^2]. \end{aligned}$$

On observe d'après le Lemme 4.3 de Bosq [3], que nous avons pour tout  $j \geq 1$ ,

$$\|\phi_{jn} - \phi'_j\|_{L^2} \leq a_j \|c_{0,n} - c_0\|_u.$$

Ce qui implique

$$\begin{aligned} \mathbb{P} \left( \sup_{j=1, \dots, k_n} \|\phi_{jn} - \phi'_j\|_{L^2} > \frac{\lambda_{k_n} \varepsilon}{6\|c_0\|_u} \right) &\leq \mathbb{P} \left( \sup_{j=1, \dots, k_n} a_j \|c_{0,n} - c_0\|_u > \frac{\lambda_{k_n} \varepsilon}{6\|c_0\|_u} \right) \\ &\leq \frac{36\|c_0\|_u^2}{\lambda_{k_n}^2 \varepsilon^2} \mathbb{E} \left[ \sup_{j=1, \dots, k_n} a_j \|c_{0,n} - c_0\|_u \right]^2 \\ &\leq \frac{36\|c_0\|_u^2}{\lambda_{k_n}^2 \varepsilon^2} \left( \sum_{j=1}^{k_n} a_j \right)^2 \mathbb{E} [\|c_{0,n} - c_0\|_u^2]. \end{aligned}$$

Par ailleurs, on a

$$\begin{aligned} \mathbb{P}(E_n^c) &\leq \mathbb{P} \left( \|c_{0,n} - c_0\|_u > \frac{\lambda_{k_n}}{2} \right) \\ &\leq \frac{4}{\lambda_{k_n}^2} \mathbb{E} [\|c_{0,n} - c_0\|_u^2]. \end{aligned}$$

Il suffit alors de combiner les inégalités ci-dessus avec la Proposition 3.1 et la condition  $H.4$  pour obtenir le résultat.  $\square$

*Preuve du Théorème 3.1 :* Nous reprenons la stratégie adoptée par Pumo [9] pour démontrer la convergence de l'estimateur d'autocorrélation. D'après [2] et [9] sous la condition  $H.3$  nous avons pour tout  $i, j \geq 1$ ,

$$r'_{ijn} = \frac{\delta'_{ijn}}{\lambda_{jn}},$$

où  $\delta'_{ijn}$  est la projection de la fonction de covariance croisée  $c_{1,n}$  sur l'espace vectoriel engendré par les fonctions  $(s, t) \rightarrow \phi'_i(s)\phi'_j(t)$  et défini par

$$\delta'_{ijn} = \int_0^1 \int_0^1 c_{1,n}(s, t) \phi'_i(s) \phi'_j(t) ds dt.$$

D'après Pumo [9], nous avons la décomposition suivante

$$\|\hat{r}_{n,k_n} - r\|_u \leq R_{1,n} + R_{2,n} + R_{3,n} + R_{4,n},$$

avec

$$\begin{aligned} R_{1,n} &= \left\| \sum_{i,j=1}^{k_n} \hat{r}_{ijn}(\phi_{in} \otimes \phi_{jn}) - \sum_{i,j=1}^{k_n} r'_{ijn}(\phi_{in} \otimes \phi_{jn}) \right\|_u \\ &\leq \frac{2}{\lambda_{k_n}^2} (\|c_{1,n}\|_u + \|c_{1,n} - c_1\|_u) (\|c_0\|_u + \|c_{0,n} - c_0\|_u) \sup_{j=1, \dots, k_n} \|\phi_{jn} - \phi'_j\|_u k_n^2, \\ R_{2,n} &= \left\| \sum_{i,j=1}^{k_n} r'_{ijn}(\phi_{in} \otimes \phi_{jn}) - \sum_{i,j=1}^{k_n} r'_{ijn}(\phi'_i \otimes \phi'_j) \right\|_u \\ &\leq \frac{(\|c_1\|_u + \|c_{1,n} - c_1\|_u)}{\lambda_{k_n n}} \left( \frac{\|c_0\|_u + \|c_{0,n} - c_0\|_u}{\lambda_{k_n n}} + \frac{\|c_0\|_u}{\lambda_{k_n}} \right) \sup_{j=1, \dots, k_n} \|\phi_{jn} - \phi'_j\|_u k_n^2, \\ R_{3,n} &= \left\| \sum_{i,j=1}^{k_n} r'_{ijn}(\phi'_i \otimes \phi'_j) - \sum_{i,j=1}^{k_n} r_{ij}(\phi_i \otimes \phi_j) \right\|_u \\ &\leq \left( \frac{\|c_{1,n} - c_1\|_u}{\lambda_{k_n n}} + \frac{\|c_1\|_u}{\lambda_{k_n n} \lambda_{k_n}} \sup_{j=1, \dots, k_n} |\lambda_{jn} - \lambda_j| \right) \frac{\|c_0\|_u^2}{\lambda_{k_n}^2} k_n^2, \\ R_{4,n} &= \left\| \sum_{i,j=1}^{k_n} r_{ij}(\phi_i \otimes \phi_j) - r \right\|_u. \end{aligned}$$

Alors pour tout  $\varepsilon > 0$ ,

$$\mathbb{P}(\|\hat{r}_{n,k_n} - r\|_u > \varepsilon) \leq \mathbb{P}(R_{1,n} > \varepsilon/4) + \mathbb{P}(R_{2,n} > \varepsilon/4) + \mathbb{P}(R_{3,n} > \varepsilon/4) + \mathbb{P}(R_{4,n} > \varepsilon/4).$$

On traite chaque terme séparément. Soit  $E_n = \left\{ |\lambda_{k_n} - \lambda_{k_n n}| \leq \frac{\lambda_{k_n}}{2} \right\}$ . Alors pour tout  $\varepsilon > 0$ ,

$$\mathbb{P}(R_{1,n} > \varepsilon/4) \leq \mathbb{P}(R_{1,n} > \varepsilon/4, E_n) + \mathbb{P}(R_{1,n} > \varepsilon/4, E_n^c).$$

Or

$$\mathbb{P}(R_{1,n} > \varepsilon/4, E_n^c) \leq \mathbb{P}(|\lambda_{k_n n} - \lambda_{k_n}| > \lambda_{k_n}/2).$$

D'autre part

$$\mathbb{P}(R_{1,n} > \varepsilon/4, E_n) \leq \mathbb{P}(R_{1,n} > \varepsilon/4, \lambda_{k_n n} > \lambda_{k_n}/2).$$

En adoptant les mêmes techniques de la décomposition de Pumo [1], et d'après la Proposition 3.1, nous avons

$$R_{1,n} \leq \frac{2}{\lambda_{k_n n}^2} (\|c_{1,n}\|_u + \|c_{1,n} - c_1\|_u) (\|c_0\|_u + \|c_{0,n} - c_0\|_u) \sup_{j=1, \dots, k_n} \|\phi_{jn} - \phi'_j\|_u k_n^2.$$

Sous la condition H.3, il s'ensuit que

$$\mathbb{P}\left(R_{1,n} > \frac{\varepsilon}{4}, \lambda_{k_n n} > \frac{\lambda_{k_n}}{2}\right) \leq \mathbb{P}\left(\frac{8}{\lambda_{k_n n}^2} V_n \sup_{j=1, \dots, k_n} \|\phi_{jn} - \phi'_j\|_u k_n^2 > \varepsilon/8\right),$$

avec

$$V_n = (\|c_1\|_u + \|c_{1,n} - c_1\|_u) (\|c_0\|_u + \|c_{0,n} - c_0\|_u).$$

D'après la Proposition 3.1 et le Lemme 5.1, nous obtenons la convergence en probabilité du terme  $R_{1,n}$ . Nous procédons de la même manière pour les autres inégalités en tenant compte que dans  $E_n$ , nous avons

$$R_{2,n} \leq \frac{2(\|c_1\|_u + \|c_{1,n} - c_1\|_u)(3\|c_0\|_u + 2\|c_{0,n} - c_0\|_u)}{\lambda_{k_n n}^2} \sup_{j=1, \dots, k_n} \|\phi_{jn} - \phi'_j\|_\infty k_n^2,$$

$$R_{3,n} \leq \left( \frac{2\|c_{1,n} - c_1\|_u}{\lambda_{k_n n}} + \frac{2\|c_1\|_u}{\lambda_{k_n n}^2 \lambda_{k_n}} \sup_{j=1, \dots, k_n} |\lambda_{jn} - \lambda_j| \right) \frac{\|c_0\|_u^2}{\lambda_{k_n n}^2} k_n^2.$$

D'autre part, sous la condition H.5, pour tout  $\varepsilon > 0$ ,

$$\lim_{n \rightarrow \infty} \mathbb{P}(R_{4,n} > \varepsilon/4) = 0.$$

□

*Preuve du Corollaire 3.1 :* Nous avons

$$\|\hat{\rho}_{n,k_n}(X_n) - \rho(X_n)\| \leq \|\hat{r}_{n,k_n} - r\|_u \|X_n\|.$$

Alors pour tout  $\varepsilon > 0$ , nous avons

$$\mathbb{P}(\|\hat{\rho}_{n,k_n}(X_n) - \rho(X_n)\| > \varepsilon) \leq \mathbb{P}\left(\|\hat{r}_{n,k_n} - r\|_u > \frac{\varepsilon}{\|X_n\|}\right)$$

Comme  $(X_n)$  est supposé bornée et d'après la Proposition 3.1 nous obtenons le résultat.

### Remerciements.

J'exprime ma gratitude au Professeur Denis Bosq pour ses orientations et ses suggestions constructives qui m'ont permis de mener à terme ce travail.

## Références.

- [1] BILLINGSLEY, P. (1999), *Convergence of Probability Measures*, 2nd ed. Wiley, New York.
- [2] BOSQ, D. (1991), *Modelization, non-parametric estimation and prediction for continuous time processes*, Nato Asi Series C, ed. Roussas, 335, 509-529.
- [3] BOSQ, D. (2000), *Linear Processes in Functions spaces. Theory and Applications*. Lecture Notes in Statistics 149. Springer.
- [4] DAMONS, J, GUILLAS, S (2010), *Modelization for functional Autoregressive processes Package : far du logiciel R, version 6.3.*
- [5] MAS, A. (1999), *Normalité asymptotique de l'estimateur empirique de l'opérateur d'autocorrélation d'un processus ARH(1)*, C.R Acad.Sci Paris Sér I Math., 329, 899-902.
- [6] MOURID, T. (1995), *Contribution à la statistique des processus autorégressifs à temps continu*. Thesis. university of Paris 6.
- [7] PAULASKAS, V., JUKNEVICIENE, D. (1988) *Rate of convergence in the central limit theorem in the space  $D[0, 1]$* , Litovsk. Mat. Sb. 28, no. 3, 507-519.
- [8] PESTMAN, W. R. *Mesurability of linear operators in the Skorokhod topology*, Bull. Belg. Math. Soc. Simon Stevin 2(4) (1995) 381-388.
- [9] PUMO, B. (1998) *Prediction of continuous time processes by  $C[0, 1]$ -valued autoregressive process*, Statistical Inference for Stochastic Processes 1, 297-309.
- [10] TAYLOR, R. L. (1978), *Stochastic convergence of weighted sums of random elements in linear spaces*, Lecture Notes in Mathematics, vol. 672, Springer, Berlin.

L.S.T.A., UNIVERSITÉ PIERRE ET MARIE CURIE  
 4 PLACE JUSSIEU, 75252 PARIS CEDEX 05, FRANCE  
 E-MAIL : layal.elhajj@gmail.com