



Statistique

Estimateurs du minimum de distance de Hellinger des processus linéaires à longue mémoire

Hellinger distance estimates of long memory linear processes

Armel Landry Bitty^a, Ouagnina Hili^b

^a Université d'Abobo-Adjamé 01, BP 8458, Abidjan 01, Côte d'Ivoire

^b Institut national polytechnique Félix-Houphouët – Boigny de Yamoussoukro, BP 1911 Yamoussoukro, Côte d'Ivoire

INFO ARTICLE

Historique de l'article :

Reçu le 6 août 2009

Accepté après révision le 18 février 2010

Présenté par Paul Deheuvels

RÉSUMÉ

On considère le processus linéaire $(X_t, t \in \mathbb{Z})$ à valeurs dans \mathbb{R} , défini de la manière suivante : $X_t = \sum_{i=0}^{\infty} a_i(\theta)\varepsilon_{t-i}$ où $(\varepsilon_t)_{t \in \mathbb{Z}}$ est une suite de variables aléatoires dans \mathbb{R} , indépendantes et identiquement distribuées, et $\theta \in \Theta$ avec $\Theta \subset \mathbb{R}^q$. X_t est supposé être un processus gaussien à longue mémoire. On se propose, dans cette note, d'estimer le paramètre θ par la méthode du minimum de distance de Hellinger. On établit, sous certaines conditions, des théorèmes limites de l'estimateur ainsi obtenu.

© 2010 Académie des sciences. Publié par Elsevier Masson SAS. Tous droits réservés.

ABSTRACT

We consider the real-valued linear process $(X_t, t \in \mathbb{Z})$ which is defined as: $X_t = \sum_{i=0}^{\infty} a_i(\theta)\varepsilon_{t-i}$ where $(\varepsilon_t)_{t \in \mathbb{Z}}$ is a sequence of real-valued random variables, independent and identically distributed, and $\theta \in \Theta$ with Θ a compact subset of \mathbb{R}^q . The process is assumed to be a Gaussian and long memory process. We propose, in this note, to estimate the parameter θ by the minimum Hellinger distance method. We establish, under some mild assumptions, the asymptotic properties of this estimates.

© 2010 Académie des sciences. Publié par Elsevier Masson SAS. Tous droits réservés.

1. Introduction

Nous considérons un processus gaussien $X_t = \sum_{i=0}^{\infty} a_i(\theta)\varepsilon_{t-i}$, centré et à longue mémoire de densité $f(\cdot, \theta)$. Nous supposons la suite $a_n(\theta) = n^{-\lambda}L(n)$ où $L(\cdot)$ est une fonction à variation lente à l'infini et λ le paramètre de longue mémoire. Le processus considéré étant gaussien, la présence de longue mémoire apparaît pour $1/2 < \lambda < 1$. Nous estimons le paramètre vectoriel θ basé sur un nombre fini d'observations X_1, \dots, X_n par la méthode du minimum de distance de Hellinger. Cette méthode a été introduite pour la première fois par Beran [1–3] dans le cas des observations indépendantes. Ensuite Hili [4–6] l'a développé dans le cas des processus de mélange fort. Nous construisons un estimateur $\hat{\theta}_n$ du paramètre θ . La distance de Hellinger se définit comme suit : $\|\hat{f}_n^{1/2}(\cdot) - f(\cdot, \theta)^{1/2}\|_2$ où $\|\cdot\|_2$ est la norme L_2 et $\hat{f}_n(\cdot)$ est un estimateur non paramétrique de la densité de X_t . Nous définissons alors $\hat{\theta}_n = \arg \min_{\theta \in \Theta} \|\hat{f}_n^{1/2}(\cdot) - f(\cdot, \theta)^{1/2}\|_2$. Dans le paragraphe 2, on définit les notations et les hypothèses utiles. Le paragraphe 3, qui constitue l'essentiel de cette note concerne les théorèmes-limites (convergence presque sûre et distribution limite) de l'estimateur.

Adresses e-mail : allbye1@yahoo.fr (A.L. Bitty), o_hili@yahoo.fr (O. Hili).

2. Notations et hypothèses

Soit $\mathcal{F} = \{f(\cdot, \theta)\}_{\theta \in \Theta}$ une famille de fonctions où Θ est un sous-ensemble compact de \mathbb{R}^q tel que pour tout $\theta \in \Theta$, $f(\cdot, \theta) : \mathbb{R} \rightarrow \mathbb{R}$ est une fonction positive et intégrable. Soit r un entier positif. $f^{(r)}(x, \theta)$ est la dérivée de $f(x, \theta)$ à l'ordre r par rapport à x . Supposons que $f(\cdot, \theta)$ vérifie les hypothèses suivantes.

(A₁) Pour tout $x \in \mathbb{R}$, $f(x, \theta) > 0$, $f(\cdot, \theta)$ et $f^{(r)}(\cdot, \theta)$ sont continues et pour tout $x \in \mathbb{R}$, $f^{1/2}(\cdot, \theta)$ est continûment différentiable au point $\theta \in \Theta$.

(A₂) L'ensemble de discontinuités de $f(\cdot, \theta)$ et $\frac{\partial}{\partial x} f^{1/2}(\cdot, \theta)$ est de mesure de Lebesgue nulle et $f(\cdot, \theta)$ et $f^{(r)}(\cdot, \theta)$ sont bornées sur \mathbb{R} .

(A₃) Pour $\theta, \mu \in \Theta$, $\theta \neq \mu$ implique que $\{x/f(x, \theta) \neq f(x, \mu)\}$ a une mesure de Lebesgue positive.

(A₄) (i) Le noyau K est une fonction positive telle que $\int_{\mathbb{R}} K(u) du = 1$, $\int_{\mathbb{R}} u^j K(u) du = 0$ pour $1 \leq j \leq s-1$ et $\int_{\mathbb{R}} |u^s| K(u) du < +\infty$, pour tout $s \geq 2$. (ii) K est une fonction deux fois absolument continue et K'' est bornée. (iii) K est une fonction symétrique autour de zéro et a un support compact. (iv) $\int_{\mathbb{R}} |K'(u)| du < +\infty$.

(A₅) La fenêtre $\{b_n\}$ vérifie les conditions, $b_n \rightarrow 0$, $nb_n \rightarrow \infty$, $(nb_n)^l \rightarrow \infty$ et $b_n = o(n^{\alpha/l})$ pour $-1 < \alpha < -9/13$ et $3 \leq l \leq 5$.

(A₆) Il existe une constante $\beta > 0$ telle que $\inf_{x \in \mathbb{R}} f(x, \theta) \geq \beta$.

Soit \mathfrak{F} l'ensemble des fonctions de densité par rapport à la mesure de Lebesgue sur \mathbb{R} . Définissons la fonctionnelle $T : \mathfrak{F} \rightarrow \Theta$. Soit $g \in \mathfrak{F}$. Définissons $B(g) = \{\theta \in \Theta : H_2(f_\theta, g) = \min_{\beta \in \Theta} H_2(f_\beta, g)\}$ où H_2 est la distance de Hellinger. Si $B(g)$ est réduit à un seul élément, alors $T(g)$ est la valeur de cet élément. Sinon, nous choisissons arbitrairement un unique élément de ces minimums, et nous l'appelons $T(g)$.

Nous prenons comme estimateur de $f(\cdot, \theta)$, la fonction suivante : $\hat{f}_n(x) = \frac{1}{nb_n} \sum_{j=1}^n K\left(\frac{x-X_j}{b_n}\right)$, pour tout $x \in \mathbb{R}$.

3. Théorèmes-limites

Les Théorèmes 1 et 2 utilisent les lemmes ci-après :

Lemme 1. Supposons que (A₃) est vérifiée. Si $f(\cdot, \theta)$ est continue sur \mathbb{R} , alors (i) Pour tout $g \in \mathfrak{F}$, $B(g) \neq \emptyset$. (ii) Si $B(g)$ est réduit à un seul élément, alors T est continue en g dans la topologie de Hellinger. (iii) $T(f(\cdot, \theta)) = \theta$ uniquement sur Θ .

Preuve. Voir le lemme 3.1 dans Hili [4]. \square

Lemme 2. Supposons que $R_\theta = f^{1/2}(\cdot, \theta)$, $\dot{R}_\theta = \frac{\partial f^{1/2}(\cdot, \theta)}{\partial \theta}$ et $\ddot{R}_\theta = \frac{\partial^2 f^{1/2}(\cdot, \theta)}{\partial^2 \theta}$ satisfont les hypothèses (A₁)–(A₄). $\{\hat{f}_n\}_{n \in \mathbb{N}}$ converge vers $f(\cdot, \theta)$ dans la topologie de Hellinger. Alors,

$$T(\hat{f}_n(\cdot)) = T(f(\cdot, \theta)) + \int_{\mathbb{R}} \rho(x, \theta) [\hat{f}_n^{1/2}(x) - f^{1/2}(x, \theta)] dx + v_n \int_{\mathbb{R}} \dot{R}_{T(f(\cdot, \theta))}(x) [\hat{f}_n^{1/2}(x) - f^{1/2}(x, \theta)] dx$$

où $\rho(x, \theta) = -[\int_{\mathbb{R}} \ddot{R}_{T(f(\cdot, \theta))}(x) f^{1/2}(x, \theta) dx]^{-1} \dot{R}_{T(f(\cdot, \theta))}(x)$ avec v_n une matrice ($q \times q$) dont ses composantes convergent vers zéro quand $n \rightarrow \infty$.

Preuve. Voir théorème 2 dans Beran [1]. \square

Lemme 3. Sous les hypothèses (A₄) et (A₅), si la densité $f(\cdot, \theta)$ des observations satisfait les hypothèses (A₁)–(A₃), alors $\hat{f}_n(\cdot)$ converge vers $f(\cdot, \theta)$ dans la topologie de Hellinger.

Indication de preuve. Nous avons $|\hat{f}_n(x) - f(x, \theta)| = |\hat{f}_n(x) - E(\hat{f}_n(x)) + E(\hat{f}_n(x)) - f(x, \theta)|$ ainsi, $|\hat{f}_n(x) - f(x, \theta)| \leq |\hat{f}_n(x) - E(\hat{f}_n(x))| + |E(\hat{f}_n(x)) - f(x, \theta)|$. Pour la preuve du second membre à droite de l'inégalité, un développement de Taylor-Lagrange à l'ordre 4 montre que $(E(\hat{f}_n(x)) - f(x, \theta)) \rightarrow 0$ p.s., lorsque $n \rightarrow \infty$. En effet

$$e_n(x) = E(\hat{f}_n(x)) - f(x, \theta) = \int_{\mathbb{R}} \left(-b_n u f'(x, \theta) + \frac{b_n^2 u^2}{2} f''(x, \theta) - \frac{b_n^3 u^3}{6} f^{(3)}(x, \theta) + \frac{b_n^4 u^4}{24} f^{(4)}(x, \theta) \right) K(u) du.$$

On obtient quand $n \rightarrow \infty$, $-1 < \alpha < -9/13$

$$\sup_{x \in \mathbb{R}} |e_n(x)| \leq \sup_{x \in \mathbb{R}} |f^{(4)}(x, \theta)| \frac{n^{4\alpha/l}}{24} \int_{\mathbb{R}} u^4 |K(u)| du = o(n^{4\alpha/l}), \quad \text{lorsque } n \rightarrow \infty.$$

De Wu et Mielniczuk [8] pour le premier membre à droite de l'inégalité, nous montrons que $\hat{f}_n(x) - E(\hat{f}_n(x)) \rightarrow 0$ p.s., lorsque $n \rightarrow \infty$. En utilisant l'inégalité de Tchebychev, on a, pour $\mu > 0$, $\text{Prob}(|\hat{f}_n(x) - E(\hat{f}_n(x))| > \mu) \leq \frac{1}{\mu^2} V(\hat{f}_n(x))$ où $V(\cdot)$ est la variance. Alors la convergence de $\sum_n \text{Prob}(|\hat{f}_n(x) - E(\hat{f}_n(x))| > \mu)$ dépend de la convergence de $\sum_n V(\hat{f}_n(x))$ qui est finie pour $3 \leq l \leq 5$, $-1 < \alpha < -9/13$ et $1/2 < \lambda < 1$. \square

Théorème 1 (Convergence presque sûre). *Supposons que (A₁)–(A₅) sont vérifiées. Alors, $\hat{\theta}_n = T(\hat{f}_n) \rightarrow \theta = T(f(\cdot, \theta))$ p.s., quand $n \rightarrow \infty$.*

Indication de preuve. Du Lemme 3 $\text{Sup}_{x \in \mathbb{R}} |\hat{f}_n(x) - f(x, \theta)| \rightarrow 0$ p.s. lorsque $n \rightarrow \infty$. Alors $\text{Prob}\{\lim_{n \rightarrow \infty} \hat{f}_n^{1/2}(x) = f^{1/2}(x, \theta)$ pour tout $x\} = 1$. Ainsi, $\int_{\mathbb{R}} \hat{f}_n(x) dx = \int_{\mathbb{R}} f(x, \theta) dx = 1$, alors $H_2(\hat{f}_n(x), f(x, \theta)) = \{\int_{\mathbb{R}} [\hat{f}_n^{1/2}(x) - f^{1/2}(x, \theta)]^2 dx\}^{1/2} \rightarrow 0$ p.s. lorsque $n \rightarrow \infty$. $T(\hat{f}_n(\cdot)) = \hat{\theta}_n$ et $T(f(\cdot, \theta)) = \theta$, du Lemme 1 on a le resultat. \square

Soit

$$\sigma_{n,r}^2 = C(\lambda, r)n^{2-r(2\lambda-1)}L^{2r}(n)[E(\varepsilon^2)]^r \tag{1}$$

où

$$C(\lambda, r) = [r![1 - r(\lambda - 1/2)][1 - r(2\lambda - 1)]]^{-1} \left[\int_0^\infty (x + x^2)^{-\lambda} dx \right]^r.$$

Posons

$$J_n = \begin{cases} (nb_n)^{l/2}, \\ \frac{n^l}{\sigma_{n,1}}, \\ \frac{(nb_n)^l}{b_n^l \sigma_{n,1}}, \\ \frac{n^l}{\sigma_{n,r}}. \end{cases}$$

Posons $R_\theta(\cdot) = f^{1/2}(\cdot, \theta)$, $\dot{R}_\theta(\cdot) = \partial f^{1/2}(\cdot, \theta) / \partial \theta$ et $W(\cdot, \theta)$ la fonction suivante :

$$W(x, \theta) = \left[\int_{\mathbb{R}} \dot{R}_\theta(x) \dot{R}_\theta^t(x) dx \right]^{-1} \dot{R}_\theta(x)$$

où $\dot{R}_\theta(x)$ est une quantité qui existe et t représente la transposée.

Condition 1. $J_n v_n \rightarrow 0, n \rightarrow \infty$.

Théorème 2 (Distribution asymptotique). *Supposons que (A₁)–(A₅) et la Condition 1 sont vérifiées. Si (i) $\int_{\mathbb{R}} \ddot{g}(x, \theta)g(x, \theta) dx$ est une $(q \times q)$ -matrice non singulière, (ii) $W(\cdot, \theta)$ admet un support compact alors,*

$$J_n[\hat{\theta}_n - \theta] \rightarrow^{\mathcal{L}} N\left(0; \int_{\mathbb{R}} Y(x, \theta)\Sigma^2(x)Y^t(x, \theta) dx\right) \text{ ou } J_n[\hat{\theta}_n - \theta] \Rightarrow \int_{\mathbb{R}} Y(x, \theta)U(x)Y^t(x, \theta) dx$$

où $\Sigma^2(x)$ et $U(x)$ prennent les différentes formes suivantes :

$$\Sigma^2(x) = \begin{cases} f(x, \theta) \int_{\mathbb{R}} K^2(u) du, \\ |f^{(1)}(x, \theta)|^2, \\ |f^{(r)}(x, \theta)k_{r-1}|^2, \\ \sigma^2(x, c), \end{cases}$$

$$U(x) = \begin{cases} (-1)^r Z_{r,\lambda} f^{(r)}(x, \theta), \\ \sum_{j=0}^{r-1} \left[\frac{C(\lambda, r-j)}{C^{r-j-1}(\lambda, 1)} \right]^{1/2} \times \frac{k_j}{C^{r-j-1}(\lambda, 1)} Z_{r-j,\lambda} f^{(r)}(x, \theta) \end{cases}$$

où

$$Z_{r,\lambda} = C(\lambda, r)^{-1/2} \int_S \left\{ \int_0^1 \prod_{i=1}^r [\max(v - u_i, 0)]^{-\lambda} dv \right\} dB(u_1) \cdots dB(u_r) \tag{2}$$

est l'intégrale multiple de Wiener-Itô. Et $Y(x, \theta) = W(x, \theta) f^{-1/2}(x, \theta)$.

Indication de preuve. Du Lemme 2,

$$J_n[\hat{\theta}_n - \theta] = J_n \int_{\mathbb{R}} W(x, \theta) [\hat{f}_n^{1/2}(x) - f^{1/2}(x, \theta)] dx + v_n J_n \int_{\mathbb{R}} \hat{R}(x, \theta) [\hat{f}_n^{1/2}(x) - f^{1/2}(x, \theta)] dx$$

avec $J_n v_n$ qui tend vers zéro quand n tend vers l'infini. Du théorème 4 de Beran [1] et le théorème 2.1 et la remarque 2.2 de Ho et Hsing [7], nous montrons que $J_n v_n \int_{\mathbb{R}} \hat{R}(x, \theta) [\hat{f}_n^{1/2}(x) - f^{1/2}(x, \theta)] dx \rightarrow_p 0$. Donc la distribution limite de $J_n[\hat{\theta}_n - \theta]$ dépend de la distribution limite de $L_n(x, \theta)$, avec

$$L_n(x, \theta) = J_n \int_{\mathbb{R}} W(x, \theta) [\hat{f}_n^{1/2}(x) - f^{1/2}(x, \theta)] dx.$$

Pour $a \geq 0$, $b > 0$, de l'égalité algébrique on a :

$$a^{1/2} - b^{1/2} = 2^{-1} b^{-1/2} (a - b) - [2b^{1/2} (a^{1/2} + b^{1/2})^2]^{-1} (a - b)^2$$

ainsi

$$\begin{aligned} L_n(x, \theta) &= 2^{-1} J_n \int_{\mathbb{R}} W(x, \theta) f(x, \theta)^{-1/2} [\hat{f}_n(x) - f(x, \theta)] dx \\ &\quad + 2^{-1} \left(J_n \int_{\mathbb{R}} W(x, \theta) f(x, \theta)^{-1/2} \frac{(\hat{f}_n(x) - f(x, \theta))^2}{(\hat{f}_n^{1/2}(x) + f(x, \theta)^{1/2})^2} dx \right) \\ &= D_n(x, \theta) + E_n(x, \theta) \end{aligned}$$

avec

$$D_n(x, \theta) = 2^{-1} J_n \int_{\mathbb{R}} W(x, \theta) f(x, \theta)^{-1/2} [\hat{f}_n(x) - f(x, \theta)] dx$$

et

$$E_n(x, \theta) = 2^{-1} J_n \int_{\mathbb{R}} W(x, \theta) f(x, \theta)^{-1/2} \frac{[\hat{f}_n(x) - f(x, \theta)]^2}{(\hat{f}_n^{1/2}(x) + f(x, \theta)^{1/2})^2} dx.$$

De l'hypothèse (A_6) , $\omega = \inf_x f(x, \theta) > 0$, on a

$$|E_n(x, \theta)| \leq 2^{-1} \omega^{-3/2} J_n \int_{\mathbb{R}} |W(x, \theta)| [\hat{f}_n(x) - f(x, \theta)]^2 dx.$$

Du Lemme 3 et du théorème 4 de Beran [1], on déduit que E_n tend vers zéro en probabilité quand $n \rightarrow \infty$.

On a $|\hat{f}_n(\cdot) - f(\cdot, \theta)| \rightarrow 0$, when $n \rightarrow \infty$.

Il suffit de montrer que la distribution limite de $J_n[\hat{\theta}_n - \theta]$ dépend de la distribution limite de $D_n(x, \theta)$:

$$D_n(x, \theta) = 2^{-1} J_n \int_{\mathbb{R}} W(x, \theta) f^{-1/2}(x, \theta) [\hat{f}_n(x) - f(x, \theta)] dx.$$

De la preuve du Lemme 3, on déduit que

$$D_n(x, \theta) \rightarrow^{\mathcal{L}} N\left(0; \int_{\mathbb{R}} Y(x, \theta) \Sigma^2(x) Y^t(x, \theta) dx\right) \quad \text{ou} \quad D_n(x, \theta) \Rightarrow \int_{\mathbb{R}} Y(x, \theta) U(x) Y^t(x, \theta) dx$$

avec $Y(x, \theta)$, $U(x, \theta)$ et $\Sigma^2(x, \theta)$ définient plus haut. \square

Références

- [1] R. Beran, Minimum Hellinger distance estimates for parametric models, *Ann. Statist.* 5 (1977) 445–463.
- [2] R. Beran, An efficient and robust adaptive estimator of location, *Ann. Statist.* 6 (1978) 292–313.
- [3] R. Beran, Efficient robust estimation for parametric models, *Z. Wahrsch. Verw. Gebiete* 55 (1981) 91–109.
- [4] O. Hili, On the estimation of nonlinear time series models, *Stochastics and Stochastics Reports* 52 (1995) 207–226.
- [5] O. Hili, On the estimation of β -ARCH models, *Statist. Probab. Lett.* 45 (1999) 285–293.
- [6] O. Hili, Hellinger distance estimation of general bilinear time series models, *Statist. Methodol.* 5 (2008) 119–128.
- [7] H.C. Ho, T. Hsing, On the asymptotic expansion of the empirical process of long memory moving averages, *Ann. Statist.* 24 (1996) 992–1024.
- [8] W.B. Wu, J. Mielniczuk, Kernel density estimation for linear processes, *Ann. Statist.* 30 (2002) 1441–1459.