C. R. Acad. Sci. Paris, Ser. I 350 (2012) 721–725

Contents lists available at SciVerse ScienceDirect

C. R. Acad. Sci. Paris, Ser. I

www.sciencedirect.com

Statistique

Estimation par le minimum de distance de Hellinger d'un processus ARFIMA

Minimum Hellinger distance estimation of an ARFIMA process

Amadou Kamagate [a,b], Ouagnina Hili [b]

[a] *Université d'Abobo-Adjamé, Abidjan, Côte d'Ivoire*

[b] *Laboratoire de mathématiques et des nouvelles technologies de l'information, Institut National Polytechnique Houphouët-Boigny de Yamoussoukro, BP 1911, Yamoussoukro, Côte d'Ivoire*

INFO ARTICLE

Historique de l'article :
Reçu le 21 février 2009
Accepté après révision le 19 juillet 2012
Disponible sur Internet le 3 août 2012

Présenté par le Comité de rédaction

RÉSUMÉ

Dans cette Note, nous définissons l'estimateur du minimum de distance de Hellinger d'un processus ARFIMA (AutoRegressive Fractionally Integrated Moving Average). L'estimateur minimise la distance de Hellinger entre la fonction de densité de probabilité de l'innovation du processus et une fonction aléatoire paramétrée. Sous certaines conditions, nous établissons les propriétés asymptotiques de cet estimateur.

ABSTRACT

In this Note, we determine the minimum Hellinger distance estimate of an ARFIMA (AutoRegressive Fractionally Integrated Moving Average) process. The estimate minimizes the Hellinger distance between the probability density function of the innovation of the process and a parameterized random function. Under some assumptions, we establish the asymptotic properties of this estimate.

Abridged English version

The purpose of this Note is to give an estimate of λ defined in Eq. (1) using the Minimum Hellinger Distance (MHD) method. The asymptotic normality of the estimate requires the stationarity of the model (1) and the existence moments of the stationary distribution.

Denote by $f_\lambda(.)$ the density of the white noise ε_t and by $\widehat{f}_{\lambda,n}$ the following random function:

$$\widehat{f}_{\lambda,n}(x) = \frac{1}{nb_n}\sum_{t=1}^{n} K\left(\frac{x-\widehat{\varepsilon}_t}{b_n}\right),$$

where $\widehat{\varepsilon}_t = \sum_{j=0}^{n}\alpha_j(\lambda)y_{t-j}$ is an estimate of ε_t for $t = 1,\dots,n$, (b_n) is a sequence of bandwidths and K is a kernel function on $\mathbb{R}$.

The algorithm of estimation consists in minimizing the Hellinger distance between $f_\lambda(.)$ and the random function $\widehat{f}_{\lambda,n}(.)$ in (2) (see the French version) with respect to the parameter λ.

In our work, we consider here and after that d_0 is the true value of the long memory parameter, where $0 < d_0 < 1/2$.

Adresses e-mail : a.kamagate@yahoo.fr (A. Kamagate), o_hili@yahoo.fr (O. Hili).

1631-073X/$ – see front matter © 2012 Académie des sciences. Publié par Elsevier Masson SAS. Tous droits réservés.
http://dx.doi.org/10.1016/j.crma.2012.07.005

Main results

Theorem 1. *Assume that* (A1)–(A6) *hold. Furthermore, let assume that* $b_n = n^\alpha L(n)$ *where* $L(n)$ *is a slowly varying function and* $\frac{-1}{3} < \alpha < \frac{1}{3}(d_0 - 1)$. *Then,* $\widehat{\lambda}_n$ *almost surely converges to* λ_0 *when* $n \longrightarrow \infty$.

Theorem 2. *Assume that* (A1)–(A6) *hold. Furthermore, let assume that* $b_n = n^\alpha L(n)$ *where* $L(.)$ *is a slowly varying function and* $\frac{-1}{3} < \alpha < \frac{1}{3}(d_0 - 1)$. *If*

(i) *the components of* $\dot{g}_{\lambda_0}$ *and* $\ddot{g}_{\lambda_0}$ *are in* L_2 *and the norms of these components are continuous functions of* λ_0,
(ii) $\int \ddot{g}_{\lambda_0}(x) g_{\lambda_0}(x)\,dx$ *is a non-singular* $((p+q+1)\times(p+q+1))$*-matrix,*

then, the limiting distribution of $\sqrt{n}(\widehat{\lambda}_n - \lambda_0)$ *is* $N(0, \Sigma^2)$ *where*

$$\Sigma^2 = \frac{1}{4}\left[\int \dot{g}_{\lambda_0}(x)\dot{g}^t_{\lambda_0}(x)\,dx\right]^{-1} \int_{\mathbb{R}} K^2(u)\,du,$$

t denotes the transpose.

1. Introduction

Le but de cette Note est d'estimer le paramètre λ du modèle ARFIMA défini en Éq. (1) par la méthode du minimum de distance de Hellinger. Les raisons du choix de cette technique d'estimation résident dans le fait que les estimateurs obtenus sont efficaces et surtout robustes (cf. Beran [1]).

Dans le paragraphe 2, nous définissons le modèle ARFIMA. Puis dans le paragraphe 3, nous établissons la convergence presque sûre et la normalité asymptotique de l'estimateur. Enfin, le paragraphe 4 constitue le schéma des preuves.

2. Le modèle

Dans cette Note, on considère le modèle ARFIMA (p, d, q) proposé par Odaki [4] pour définir une série temporelle présentant un caractère de courte ou de longue mémoire suivant d. Un tel processus de moyenne zéro est écrit comme suit :

$$\phi(B)(1-B)^d y_t = \theta(B)\varepsilon_t, \tag{1}$$

où B est l'opérateur retard défini par $By_t = y_{t-1}$.

$\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p$ et $\theta(z) = 1 - \theta_1 z - \cdots - \theta_q z^q$ sont des polynômes de degrés respectifs p et q ; (ε_t) est une suite de variables aléatoires mutuellement non correlées de moyenne zéro et de variance finie σ^2 ; d est réel et $(1-B)^d = \sum_{j=0}^{\infty} b_j(d) B^j$ avec $b_j(d) = \frac{\Gamma(j-d)}{\Gamma(j+1)\Gamma(-d)}$, pour $j \geqslant 0$.

Nous supposons dans la suite que $\{y_t\}$ est stationnaire au second-ordre. Odaki [4] a montré que $\{y_t\}$ est inversible et stationnaire tant que $d > -1$ et $d < 1/2$ respectivement. De plus, le processus est de courte mémoire si $-1/2 < d < 0$ et de longue mémoire si $0 < d < 1/2$.

Considérant que le processus est inversible, l'équation (1) peut être reécrite sous la forme d'un processus AutoRégressif infini $(AR(\infty))$ comme suit : $\varepsilon_t = \sum_{j=0}^{\infty} \alpha_j(\lambda) y_{t-j}$, où $\{\alpha_j(\lambda)\}_{j=0}^{\infty}$ sont les coefficients associés à $\phi(B)\theta(B)^{-1}(1-B)^d$ et $\lambda = (d, \phi, \theta) \in \Lambda \subset \mathbb{R}^{p+q+1}$.

Soient les observations $y_1, \ldots, y_n$. L'innovation ε_t n'étant pas observable, nous l'estimons par : $\widehat{\varepsilon}_t = \sum_{j=0}^{n} \alpha_j(\lambda) y_{t-j}$ pour $t = 1, \ldots, n$.

3. Estimation du paramètre

Ce paragraphe est consacré à la détermination de l'estimateur du minimum de distance de Hellinger (MDH) du paramètre λ et l'établissement des propriétés asymptotiques de cet estimateur. La vraie valeur du vecteur paramètre est $\lambda_0 = (d_0, \phi_0, \theta_0)$.

La méthode MDH consiste à minimiser la distance de Hellinger entre la fonction de densité du bruit blanc $f_{\lambda_0}(.)$ et la fonction aléatoire paramétrée $\widehat{f}_{\lambda,n}(x)$ suivante :

$$\widehat{f}_{\lambda,n}(x) = \frac{1}{nb_n}\sum_{t=1}^{n} K\left(\frac{x - \widehat{\varepsilon}_t}{b_n}\right), \tag{2}$$

dans laquelle (b_n) est une suite de fenêtres, K est un noyau et $\widehat{\varepsilon}_t$ est un estimateur de ε_t.

Soit $\widehat{\lambda}_n$ l'estimateur MDH de $\lambda_0 \in \Lambda$ défini comme suit :

$$\widehat{\lambda}_n = \arg\min_{\lambda\in\Lambda} H_2(\widehat{f}_{\lambda,n}, f_{\lambda_0}), \tag{3}$$

où

$$H_2(\widehat{f}_{\lambda,n}, f_{\lambda_0}) = \left\{ \int_{\mathbb{R}} \left| \widehat{f}^{\frac{1}{2}}_{\lambda,n}(x) - f^{\frac{1}{2}}_{\lambda_0}(x) \right|^2 \mathrm{d}x \right\}^{\frac{1}{2}}.$$

Afin d'établir les propriétés asymptotiques de $\widehat{\lambda}_n$, nous considérons les hypothèses suivantes :

(A1) ε_t admet une loi absolument continue par rapport à la mesure de Lebesgue sur $\mathbb{R}$ et sa densité $f_{\lambda_0}(.)$ est strictement positive presque partout.
(A2) $E(|\varepsilon_t|^s) < +\infty$ pour un nombre entier $s \geqslant 1$.
(A3) $\lambda_1 \neq \lambda_2$ implique que $f_{\lambda_1} \neq f_{\lambda_2}$ sur un ensemble de mesure de Lebesgue non nulle.
(A4) $f_{\lambda_0}(.)$ est uniformément continue et k-fois continûment différentiable pour un $k \geqslant 3$.
(A5) (i) $K \geqslant 0$ et $K(u) = K(-u)$ pour tout $u \in \mathbb{R}$,
(ii) il existe N_1 tel que $K \leqslant N_1 < +\infty$ et $\int_{\mathbb{R}} K(u)\,\mathrm{d}u = 1$,
(iii) il existe $C_1 > 0$ tel que $\sup_u |K(u+v) - K(u)| \leqslant C_1|v|$ pour tout $v \in \mathbb{R}$.
(A6) $\int_{\mathbb{R}} |uK(u)|\,\mathrm{d}u < \infty$.

Dans cette Note, nous considérons que la vraie valeur du paramètre de longue mémoire est d_0 et $0 < d_0 < 1/2$.

Théorème 1 *(Concergence presque sûre). Supposons que* (A1)–(A6) *sont satisfaites ; de plus supposons que $b_n = n^\alpha L(n)$ où $L(n)$ est une fonction à variation lente et $\frac{-1}{3} < \alpha < \frac{1}{3}(d_0 - 1)$. Alors, $\widehat{\lambda}_n$ converge presque sûrement vers λ_0 lorsque $n \longrightarrow \infty$.*

Soient $g_{\lambda_0} = f^{\frac{1}{2}}_{\lambda_0}$, $\dot{g}_{\lambda_0} = \frac{\partial g_{\lambda_0}}{\partial \lambda_0}$, $\ddot{g}_{\lambda_0} = \frac{\partial^2 g_{\lambda_0}}{\partial \lambda_0^2}$, $V_{\lambda_0} = [\int \dot{g}_{\lambda_0}(x)\dot{g}^t_{\lambda_0}(x)\,\mathrm{d}x]^{-1}\dot{g}_{\lambda_0}(x)$.

Théorème 2 *(Normalité asymptotique). Supposons que* (A1)–(A6) *sont satisfaites ; de plus supposons que $b_n = n^\alpha L(n)$ où $L(n)$ est une fonction à variation lente et $\frac{-1}{3} < \alpha < \frac{1}{3}(d_0 - 1)$. Si de plus*

(i) *les composantes de $\dot{g}_{\lambda_0}$ et $\ddot{g}_{\lambda_0}$ sont dans L_2 et les normes dans L_2 de ces composantes sont des fonctions continues en λ_0,*
(ii) *$\int \ddot{g}_{\lambda_0}(x) g_{\lambda_0}(x)\,\mathrm{d}x$ est une $((p+q+1)\times(p+q+1))$-matrice non singulière.*

Alors, la loi limite de $\sqrt{n}(\widehat{\lambda}_n - \lambda_0)$ est $N(0, \Sigma^2)$, où

$$\Sigma^2 = \frac{1}{4}\left[\int \dot{g}_{\lambda_0}(x)\dot{g}^t_{\lambda_0}(x)\,\mathrm{d}x\right]^{-1} \int_{\mathbb{R}} K^2(u)\,\mathrm{d}u.$$

4. Preuve des résultats

Ce paragraphe consiste à démontrer les théorèmes 1 et 2. Pour ce faire, nous utilisons les lemmes 1, 2, 3, 3.1 dans Hili [2] et les notations suivantes :

Soit $\mathfrak{F}$ l'ensemble des densités suivant la mesure de Lebesgue sur $\mathbb{R}$. On définit la fonctionnelle $T : \mathfrak{F} \longrightarrow \Lambda$ de la manière suivante : soit $g \in \mathfrak{F}$. On pose $A(g) = \{\lambda \in \Lambda:\ H_2(g, f_\lambda) = \min_{\gamma\in\Lambda} H_2(g, f_\gamma)\}$, où H_2 est la distance de Hellinger. Si $A(g)$ est un singleton, on note $T(g)$ cette valeur. Sinon, on choisit un élément arbitraire mais unique de ces minimums et on le note $T(g)$.

Lemme 1. *Supposons que* (A2) *et* (A4)–(A6) *sont satisfaites, de plus supposons que la fenêtre b_n est choisie comme dans les théorèmes* 1 *et* 2*. Alors, $\widehat{f}_{\lambda,n}(x)$ converge presque sûrement vers $f_{\lambda_0}(x)$, lorsque $n \longrightarrow \infty$, pour tout $x \in \mathbb{R}$.*

Schéma de la preuve du Lemme 1. Nous avons $\sup_x |\widehat{f}_{\lambda,n}(x) - f_{\lambda_0}(x)| \leqslant \sup_x |\widehat{f}_{\lambda,n}(x) - \widetilde{f}_n(x)| + \sup_x |\widetilde{f}_n(x) - f_{\lambda_0}(x)|$, où $\widetilde{f}_n(x) = \frac{1}{nb_n}\sum_{t=1}^n K(\frac{x-\varepsilon_t}{b_n})$. Selon Prakasa Rao [5] le deuxième terme de droite converge vers zéro puis on applique le théorème 1 dans Mayoral [3]. Voir la version longue pour plus de détails. □

Schéma de la preuve du Théorème 1. On applique les lemmes 1 et 3.1 dans Hili [2]. En effet, du lemme 1, $\sup_x |\widehat{f}_{\lambda,n}(x) - f_{\lambda_0}(x)| \longrightarrow 0$ presque sûrement quand $n \longrightarrow \infty$.

Alors

$$Prob\Big\{\lim_{n\longrightarrow\infty}\widehat{f}^{\frac{1}{2}}_{\lambda,n}(x)=f^{\frac{1}{2}}_{\lambda_0}(x)\text{ pour tout }x\Big\}=1$$

et puisque

$$\int_{\mathbb{R}}\widehat{f}_{\lambda,n}(x)\,\mathrm{d}x=\int_{\mathbb{R}}f_{\lambda_0}(x)\,\mathrm{d}x=1,$$

alors

$$H_2(\widehat{f}_{\lambda,n},f_{\lambda_0})=\bigg(\int_{\mathbb{R}}\big|\widehat{f}^{\frac{1}{2}}_{\lambda,n}(x)-f^{\frac{1}{2}}_{\lambda_0}(x)\big|^2\,\mathrm{d}x\bigg)^{\frac{1}{2}}\longrightarrow 0\quad\text{presque sûrement quand }n\longrightarrow\infty.$$

Du lemme 3.1 dans Hili [2], suivant la continuité de la foncionnelle T dans la topologie de Hellinger, $\widehat{\lambda}_n=T(\widehat{f}_{\lambda,n}(x))\longrightarrow T(f_{\lambda_0}(x))=\lambda_0$ presque sûrement quand $n\longrightarrow\infty$. D'où le résultat. □

Lemme 2. *On suppose que* (A3), (A4) *et les conditions* (i) *et* (ii) *du théorème* 2 *sont satisfaites et que* λ_0 *est un point interieur de* Λ. *Alors pour toute suite de densités* $\{\widehat{f}_{\lambda,n}\}$ *convergeant vers* f_{λ_0} *dans la métrique de Hellinger, on a* :

$$T(f_n)=\lambda_0+\int V_{\lambda_0}(x)\big[\widehat{f}^{\frac{1}{2}}_{\lambda,n}(x)-f^{\frac{1}{2}}_{\lambda_0}(x)\big]\,\mathrm{d}x+A_n\int\dot{g}_{\lambda_0}(x)\big[\widehat{f}^{\frac{1}{2}}_{\lambda,n}(x)-f^{\frac{1}{2}}_{\lambda_0}(x)\big]\,\mathrm{d}x,\tag{4}$$

où A_n *est une* $(p+q+1)\times(p+q+1)$*-matrice dont les composantes de* $\sqrt{n}A_n$ *tendent vers zéro quand* $n\longrightarrow\infty$.

Schéma de la preuve du Lemme 2. Voir Beran [1, théorème 2]. □

Lemme 3. *On suppose que* (A1), (A2), (A4) *et* (A6) *sont satisfaites, de plus on suppose que la fenêtre* b_n *est choisie comme dans les Théorèmes* 1 *et* 2. *Alors la loi limite de* $(nb_n)^{\frac{1}{2}}[\widehat{f}_{\lambda,n}(x)-f_{\lambda_0}(x)]$ *est* $N(0,f_{\lambda_0}(x)\int_{\mathbb{R}}K^2(u)\,\mathrm{d}u)$.

Schéma de la preuve du Lemme 3. Voir théorème 2 et corollaire 1 dans Wu et Mielniczuk [6]. □

Schéma de la preuve du Théorème 2. On applique les lemmes 1, 2 et 3. En effet, du lemme 2 on peut écrire

$$\sqrt{n}(\widehat{\lambda}_n-\lambda_0)=\sqrt{n}\int V_{\lambda_0}(x)\big[\widehat{f}^{\frac{1}{2}}_{\lambda,n}(x)-f^{\frac{1}{2}}_{\lambda_0}(x)\big]\,\mathrm{d}x+o_p(1)\tag{5}$$

Pour déterminer la loi limite de $\sqrt{n}(\widehat{\lambda}_n-\lambda_0)$, il suffit de déterminer la loi limite de $\sqrt{n}\int V_{\lambda_0}(x)[\widehat{f}^{\frac{1}{2}}_{\lambda,n}(x)-f^{\frac{1}{2}}_{\lambda}(x)]\,\mathrm{d}x$, avec $V_{\lambda_0}(x)\in L_2$, et $V_{\lambda_0}\perp f^{\frac{1}{2}}_{\lambda_0}$ où $\perp$ désigne l'orthogonalité en L_2.

De (A1), $f^{\frac{1}{2}}_{\lambda_0}(x)>0$ et de l'identité algebrique :

$$\widehat{f}^{\frac{1}{2}}_{\lambda,n}-f^{\frac{1}{2}}_{\lambda_0}=\frac{\widehat{f}_{\lambda,n}-f_{\lambda_0}}{2f^{\frac{1}{2}}_{\lambda_0}}-\frac{(\widehat{f}_{\lambda,n}-f_{\lambda_0})^2}{2f^{\frac{1}{2}}_{\lambda_0}(\widehat{f}^{\frac{1}{2}}_{\lambda,n}+f^{\frac{1}{2}}_{\lambda_0})^2},$$

nous avons,

$$\sqrt{n}\int V_{\lambda_0}(x)\big(\widehat{f}^{\frac{1}{2}}_{\lambda,n}-f^{\frac{1}{2}}_{\lambda_0}\big)\,\mathrm{d}x=\sqrt{n}\int\frac{V_{\lambda_0}(x)}{2f^{\frac{1}{2}}_{\lambda_0}}(\widehat{f}_{\lambda,n}-f_{\lambda_0})\,\mathrm{d}x+B_n,\tag{6}$$

où $B_n=-\sqrt{n}\int\frac{V_{\lambda_0}(x)}{2f^{\frac{1}{2}}_{\lambda_0}(\widehat{f}^{\frac{1}{2}}_{\lambda,n}+f^{\frac{1}{2}}_{\lambda_0})^2}(\widehat{f}_{\lambda,n}-f_{\lambda_0})^2\,\mathrm{d}x$.

Puisque $2f^{\frac{1}{2}}_{\lambda_0}(\widehat{f}^{\frac{1}{2}}_{\lambda,n}+f^{\frac{1}{2}}_{\lambda_0})^2>2f^{\frac{3}{2}}_{\lambda_0}$, alors

$$|B_n|\leqslant\int\frac{|V_{\lambda_0}(x)|\sqrt{n}(\widehat{f}_{\lambda,n}(x)-f_{\lambda_0}(x))^2}{2f^{3/2}_{\lambda_0}}\,\mathrm{d}x.$$

Du lemme 1, $\sqrt{n}(\widehat{f}_{\lambda,n}(x)-f_{\lambda_0}(x))^2\longrightarrow 0$ presque sûrement $n\longrightarrow\infty$. Les conditions (i) et (ii) du théorème 2 implique que V_{λ_0} est continu et borné (pour tout λ_0 fixé) de plus, on applique le théorème de Vitali sur la sequence $S_n(x)=$

$|V_{\lambda_0}(x)|\sqrt{n}(\widehat{f}_{\lambda,n}(x) - f_{\lambda_0}(x))^2$ assure que $|B_n| \longrightarrow 0$ en probabilité lorsque $n \longrightarrow \infty$. Donc, il revient à déterminer la loi limite du premier terme de (6) qui est

$$\sqrt{n}\int \frac{V_{\lambda_0}(x)}{2 f_{\lambda_0}^{\frac{1}{2}}}(\widehat{f}_{\lambda,n} - f_{\lambda_0})\,\mathrm{d}x. \tag{7}$$

Du lemme 3, la loi limite de (7) est $N(0, \Sigma^2)$ où

$$\Sigma^2 = \frac{1}{4}\int V_{\lambda_0}(x)V_{\lambda_0}^t(x)\,\mathrm{d}x \int_{\mathbb{R}} K^2(u)\,\mathrm{d}u.$$

D'où le résultat. □

Remerciements

Nous sommes très reconnaissants aux rapporteurs dont leurs pertinentes remarques ont permis d'améliorer la qualité de la Note.

Références

[1] R. Beran, Minimum Hellinger distance estimates for parametric models, The Annals of Statistics 5 (3) (1977) 445–463.
[2] O. Hili, On the estimation of nonlinear time series models, Stochastics and Stochastics Reports 52 (1995) 207–226.
[3] L. Mayoral, Minimum distance estimation of stationary and non-stationary ARFIMA processes, Econometrics Journal 10 (2007) 124–148.
[4] M. Odaki, On the invertibility of fractionally differenced ARIMA processes, Biometrika 80 (1993) 703–709.
[5] B.L.S. Prakasa Rao, Nonparametric Functional Estimation, Academic Press, New York, 1983.
[6] W.B. Wu, J. Mielniczuk, Kernel density estimation for linear process, The Annals of Statistics 30 (2002) 1441–1459.