C. R. Acad. Sci. Paris, t. 330, Série I, p. 889–892, 2000
Probabilités/*Probability Theory*

Modèles autorégressifs explosifs avec bruit longue mémoire

Mohamed BOUTAHAR

Département de mathématiques, case 901, 163, avenue de Luminy, 13288 Marseille cedex 9, France
GREQAM, Vieille Charité 13002 Marseille, France
Courriel : boutahar@lumimath.univ-mrs.fr

(Reçu le 25 janvier 2000, accepté après révision le 13 avril 2000)

Résumé. On précise la distribution limite de l'estimateur des moindres carrés dans les modèles autorégressifs explosifs gouvernés par un bruit longue mémoire, cette distribution limite n'est pas gaussienne, mais un mélange de variables gaussiennes. © 2000 Académie des sciences/Éditions scientifiques et médicales Elsevier SAS

Explosive autoregressive models with long-memory noise

Abstract. *Here we give the limiting distribution of the least squares estimates in the explosive autoregressive models driven by a long-memory process. We prove that with an appropriate normalization the estimation error converges, in distribution, to a mixture of normal distributions.* © 2000 Académie des sciences/Éditions scientifiques et médicales Elsevier SAS

1. Introduction

Dans cette Note, $(\Omega, \mathcal{A}, P)$ est un espace de probabilité sur lequel est défini le modèle autorégressifs d'ordre p suivant :

$$y_t = a_1 y_{t-1} + \cdots + a_p y_{t-p} + \varepsilon_t, \tag{1}$$

y_t et ε_t représentent respectivement la sortie et le bruit du modèle, $\boldsymbol{\beta} = {}^{\mathrm{t}}(a_1, \ldots, a_p)$ est le paramètre inconnu. Pour estimer $\boldsymbol{\beta}$, on considère l'estimateur des moindres carrés :

$$\boldsymbol{\beta}_n = \left(\sum_{k=1}^{n} \Phi_{k-1}^{\mathrm{t}} \Phi_{k-1} \right)^{-1} \sum_{k=1}^{n} \Phi_{k-1} y_k,$$

où $\Phi_{t-1} = {}^{\mathrm{t}}(y_{t-1}, \ldots, y_{t-p})$.

Les propriétés asymptotiques de l'estimateur des moindres carrés $\boldsymbol{\beta}_n$ ont été étudiées par Lai et Wei [6] en supposant que le bruit (ε_t) est une différence de martingales, ils ont montré qu'il est fortement consistant sans supposer de contraintes sur les racines du polynôme caractéristique $\varphi(z) = 1 - a_1 z - \cdots - a_p z^p$. La distribution limite $\boldsymbol{\beta}_n$ a été étudiée par Chan et Wei [2] dans le cas où $\varphi(z)$ est instable (i.e.,

Note présentée par Jean-Pierre KAHANE.

S0764-4442(00)00292-5/FLA

$\varphi(z)=0 \Longrightarrow |z| \geqslant 1$). Les modèles autorégressifs multivariés ont été étudiés par plusieurs auteurs, Duflo et al. [4] ont étendus les travaux de Lai et Wei [6] au modèle $\mathrm{AR}_d(1)$, Touati [8] précise la loi limite dans le cas d'un modèle mixte. Dans ce travail, on considère le modèle (1), en supposant que le bruit est un processus longue mémoire. On montre que la distribution limite de $\boldsymbol{\beta}_n$, dans le cas où $\varphi(z)$ est explosif, est non gaussienne mais un mélange de distributions gaussiennes. Dans le cas où $\varphi(z)$ est instable, la distribution limite de $\boldsymbol{\beta}_n$ est précisée dans Chan et Terrin [3].

2. Notations et hypothèses

On note $\xrightarrow{\mathcal{L}}$, $\xrightarrow{P}$ et $\xrightarrow{\mathrm{L}_p}$ respectivement la convergence en loi, en probabilité et dans $\mathrm{L}_p(\Omega)$; $X \rightsquigarrow N(m,\Gamma)$ signifie que X est un vecteur gaussien de moyenne m et de matrice de covariance Γ ; ${}^{\mathrm{t}}X$ désigne la transposée de la matrice X.

Si X est une matrice, on note $\mathrm{vec}(X)$ le vecteur obtenu en empilant les colonnes les unes au dessous des autres, c'est-à-dire que si $X=(X_1,X_2,\ldots,X_p)$, X_j est la $j^{\text{ème}}$ colonne de X, alors $\mathrm{vec}(X) = {}^{\mathrm{t}}({}^{\mathrm{t}}X_1, {}^{\mathrm{t}}X_2, \ldots, {}^{\mathrm{t}}X_p)$. $\lambda_{\min}(\mathbf{A})$ (resp. $\lambda_{\max}(\mathbf{A})$) est la plus petite (resp. grande) valeur propre de la matrice $\mathbf{A}$. $\mathbf{A}$ est la matrice compagne du polynôme $\varphi(z)$.

H.1 : le bruit (ε_t) est un processus stationnaire gaussien ayant une densité spectrale $f(\lambda)$ telle que :

$$f(\lambda) = |\lambda|^{1-2H} L\big(|\lambda|^{-1}\big), \quad 1/2 < H < 1,$$

L est une fonction telle que $L(na)/L(n) \to 1$ quand $n \to \infty$ pour tout $a>0$, bornée sur tout intervalle fini, et f est intégrable sur $[-\pi,\pi]$.

Remarque 1. – Le processus (ε_t) admet la représentation spectrale :

$$\varepsilon_n = \int_{-\pi}^{\pi} \exp(in\lambda) f^{1/2}(\lambda) W(\mathrm{d}\lambda),$$

où $W(\cdot)$ est une mesure aléatoire gaussienne.

H.2 : les variables aléatoires $y_{1-p},\ldots,y_0$ sont de carrés intégrables et indépendantes de (ε_t).

H.3 : le polynôme $\varphi(z)$ est explosif (i.e., $\varphi(z)=0 \Longrightarrow |z|<1$).

3. Résultats principaux

THÉORÈME 1. – *On suppose que les hypothèses* H.1–H.3 *sont satisfaites. Alors* :

(i) *on a*

$$\mathbf{A}^{-n}\Phi_n \xrightarrow{\mathrm{L}_2} L = \Phi_0 + \int_{-\pi}^{\pi} \big(\exp(-i\lambda)\mathbf{A} - \mathbf{I}_p\big)^{-1} f^{1/2}(\lambda) W(\mathrm{d}\lambda) e_1, \tag{2}$$

$e_1 = {}^{\mathrm{t}}(1,0,\ldots,0)$ *vecteur de taille* $(p,1)$;

(ii) *la loi de L ne charge pas les hyperplans de* $\mathbb{R}^p$;

(iii)

$$\mathbf{A}^{-n} \sum_{k=1}^{n} \Phi_{k-1} {}^{\mathrm{t}}\Phi_{k-1} {}^{\mathrm{t}}\big(\mathbf{A}^{-n}\big) \xrightarrow{\mathrm{L}_1} \boldsymbol{\Sigma}_2 = \sum_{k=1}^{\infty} \mathbf{A}^{-k} L\, {}^{\mathrm{t}}L\, {}^{\mathrm{t}}\big(\mathbf{A}^{-k}\big); \tag{3}$$

de plus,

$$P(\boldsymbol{\Sigma}_2 > 0) = 1. \tag{4}$$

Éléments de démonstration. – De (1) on obtient :

$$\mathbf{A}^{-n}\Phi_n = \Phi_0 + \sum_{k=1}^{n} \mathbf{A}^{-k} \varepsilon_k e_1 = \Phi_0 + \int_{-\pi}^{\pi} \sum_{k=1}^{n} \mathbf{A}^{-k} \exp(ik\lambda) f^{1/2}(\lambda) W(\mathrm{d}\lambda) e_1,$$

$$\sum_{k=1}^{n} \mathbf{A}^{-k} \exp(ik\lambda) \longrightarrow \big(\exp(-i\lambda)\mathbf{A} - \mathbf{I}_p\big)^{-1}, \quad \forall \lambda \in [-\pi, \pi],$$

et

$$\left\| \sum_{k=1}^{n} \mathbf{A}^{-k} \exp(ik\lambda) \right\| = \mathrm{O}(1),$$

donc

$$\left\| \sum_{k=1}^{n} \mathbf{A}^{-k} \exp(ik\lambda) - \big(\exp(-i\lambda)\mathbf{A} - \mathbf{I}_p\big)^{-1} \right\|_{\mathrm{L}^2([-\pi,\pi], f(\lambda)\mathrm{d}\lambda)} \longrightarrow 0;$$

par conséquent, la convergence (2) s'obtient par la continuité dans L_2 des intégrales stochastiques.

Remarque 2. – Les intégrales stochastiques considérées dans cette Note sont des intégrales par rapport à la mesure gaussienne $W(\cdot)$ et sont donc des intégrales de Wiener–Itô (*voir* Major [7]).

Pour tout $x \in \mathbb{R}^p - \{0\}$, ${}^{\mathrm{t}}xL = {}^{\mathrm{t}}x\Phi_0 + {}^{\mathrm{t}}x\eta$, où ${}^{\mathrm{t}}x\eta$ est une variable aléatoire gaussienne centrée et de variance non nulle $\int_{-\pi}^{\pi} \left| {}^{\mathrm{t}}x(\exp(-i\lambda)\mathbf{A} - \mathbf{I}_p)^{-1} e_1 \right|^2 f(\lambda)\mathrm{d}\lambda$, et donc la loi de η, et par conséquent celle de L, ne charge pas les hyperplans de $\mathbb{R}^p$.

La démonstration de (3) et (4), s'obtient en adaptant celle de (Lai et Wei [6], Theorem 2) et en utilisant le résultat :

$$\left\| \mathbf{A}^{-n} \right\| = \mathrm{O}\big(\rho^n n^{\nu-1}\big), \tag{5}$$

où

$$\rho = \max_{1 \leqslant j \leqslant p} |\lambda_j|, \quad \nu = \max\big\{m_j,\ \|\lambda_j\| = \rho\big\},$$

$(\lambda_j, 1 \leqslant j \leqslant p)$ sont les valeurs propres de la matrice $\mathbf{A}^{-1}$, et m_j est la multiplicité de λ_j. □

Le théorème suivant donne la distribution limite de $\boldsymbol{\beta}_n$ et établit sa consistance.

THÉORÈME 2. – *Sous les hypothèses* H.1–H.3, *on a* :

(i)

$${}^{\mathrm{t}}\big(\mathbf{A}^n\big)(\boldsymbol{\beta}_n - \boldsymbol{\beta}) \xrightarrow{\mathcal{L}} \boldsymbol{\Sigma}_2^{-1} N_2, \tag{6}$$

où

$$N_2 = \mathcal{A}_1 \left(-\Phi_0 + \int_{-\pi}^{\pi} \big(\mathbf{I}_p - \mathbf{A}\exp(-i\lambda)\big)^{-1} f^{1/2}(\lambda) W(\mathrm{d}\lambda) e_1 \right),$$

où $\mathcal{A}_1$ est une matrice gaussienne indépendante de Φ_0 et de la suite (ε_t) ; de plus, $\mathcal{A}_1$ est de moyenne nulle et le $(j,k)^{\text{ème}}$ bloc de sa matrice de covariance est donné par :

$$\mathrm{var}(\mathcal{A}_1)_{(j,k)} = \int_{-\pi}^{\pi} \big(\mathbf{I}_p - \mathbf{A}\exp(i\lambda)\big)^{-1} e_j^{\mathrm{t}} e_k \big(\mathbf{I}_p - {}^{\mathrm{t}}\mathbf{A}\exp(-i\lambda)\big)^{-1} f(\lambda)\mathrm{d}\lambda;$$

e_j est le $j^{\text{ème}}$ vecteur de la base canonique de $\mathbb{R}^p$ (vecteur dont les composantes sont nulles sauf la $j^{\text{ème}}$ qui est égale à 1).

(ii) *L'estimateur $\boldsymbol{\beta}_n$ est faiblement consistant avec la vitesse de convergence*

$$\|\boldsymbol{\beta}_n - \boldsymbol{\beta}\|^2 = \mathrm{o}_p\big(d(n)\rho^{2n} n^{2(\nu-1)}\big), \tag{7}$$

ρ et ν sont donnés par (5), *et pour toute suite $d(n) \uparrow \infty$.*

Éléments de démonstration. – On utilise le théorème de Riemann–Lebesgue (*voir* Zygmund [9], Theorem (4.4)), précisant que le coefficient de Fourier c_n de toute fonction intégrable est tel que : $c_n \to 0$ quand $|n| \to \infty$; et le lemme suivant dont la démonstration est facile :

LEMME 1. – *Posons*

$$\mathcal{A}_{n,1} = \int_{-\pi}^{\pi} \big(\mathbf{I}_p - \mathbf{A}\exp(i\lambda)\big)^{-1} \mathrm{e}^{i(n+1)\lambda} f^{1/2}(\lambda) W(\mathrm{d}\lambda),$$
$$\mathcal{A}_0 = \int_{-\pi}^{\pi} \big(\mathbf{I}_p - \mathbf{A}\exp(-i\lambda)\big)^{-1} f^{1/2}(\lambda) W(\mathrm{d}\lambda).$$

Alors $Z_n = {}^{\mathrm{t}}\big({}^{\mathrm{t}}\Phi_0, {}^{\mathrm{t}}\operatorname{vec}(\mathcal{A}_0), {}^{\mathrm{t}}\operatorname{vec}(\mathcal{A}_{n,1})\big) \xrightarrow{\mathcal{L}} Z = {}^{\mathrm{t}}\big({}^{\mathrm{t}}\Phi_0, {}^{\mathrm{t}}\operatorname{vec}(\mathcal{A}_0), {}^{\mathrm{t}}\operatorname{vec}(\mathcal{A}_1)\big)$.

Dans un premier temps, on montre que

$$\mathbf{A}^{-n} \sum_1^n \Phi_{k-1}\varepsilon_k \xrightarrow{\mathcal{L}} N_2. \tag{8}$$

Après un long calcul, on montre que

$$\mathbf{A}^{-n} \sum_1^n \Phi_{k-1}\varepsilon_k = \mathrm{o}_p(1) + V_n,$$

où

$$V_n = \mathcal{A}_{n,1}\left(-\Phi_0 + \int_{-\pi}^{\pi} \big(\mathbf{I}_p - \mathbf{A}\exp(-i\lambda)\big)^{-1} f^{1/2}(\lambda) W(\mathrm{d}\lambda) e_1\right).$$

Les composantes de ce dernier vecteur sont des fonctions continues de celles du vecteur Z_n défini dans le lemme précédent, donc la convergence jointe (8) découle de ce dernier lemme et du théorème des fonctions continues (Billingsley [1], Theorem 5.1).

Finalement, la convergence (6) s'obtient de (3)–(4) et (8).

Pour la démonstration de (7), on utilise (6) pour obtenir $[d(n)]^{-1}\,{}^{\mathrm{t}}(\boldsymbol{\beta}_n - \boldsymbol{\beta})(\mathbf{A}^n)^{\mathrm{t}}(\mathbf{A}^n)(\boldsymbol{\beta}_n - \boldsymbol{\beta}) \xrightarrow{P} 0$, qui implique que

$$\lambda_{\min}\big((\mathbf{A}^n)^{\mathrm{t}}(\mathbf{A}^n)\big)\|\boldsymbol{\beta}_n - \boldsymbol{\beta}\|^2 = \mathrm{o}_p\big(d(n)\big);$$

de plus,

$$1/\lambda_{\min}\big((\mathbf{A}^n)^{\mathrm{t}}(\mathbf{A}^n)\big) = \lambda_{\max}\big({}^{\mathrm{t}}(\mathbf{A}^{-n})(\mathbf{A}^{-n})\big) = \big\|\mathbf{A}^{-n}\big\|^2 = \mathrm{O}\big(\rho^{2n} n^{2(\nu-1)}\big). \qquad \square$$

Références bibliographiques

[1] Billingsley P., Convergence of Probability Measures, Wiley, New York, 1968.
[2] Chan N.H., Wei C.Z., Limiting distributions of least-squares estimates of unstable autoregressive processes, Ann. Statis. 16 (1988) 367–401.
[3] Chan N.H., Terrin N., Inference for unstable long-memory processes with applications to fractional Unit Root Autoregressions, Ann. Statis. 23 (1995) 1662–1683.
[4] Duflo M., Senoussi R., Touati A., Moindres carrés d'un modèle autorégressif, Ann. Inst. Henri-Poincaré 27 (1991) 1–25.
[5] Hall P., Heyde C.C., Martingales Limit Theory and its Application, Academic Press, 1980.
[6] Lai T.L., Wei C.Z., Asymptotic properties of general autoregressive models and strong consistency of least squares estimates of their parameters, J. Multiv. Anal. 13 (1983) 1–13.
[7] Major P., Multiple Wiener–Itô integrals, in: Lect. Notes in Math. 849, Springer-Verlag, New York, 1981.
[8] Touati A., Vitesse de convergence en loi de l'estimateur des moindres carrés d'un modèle autorégressif (cas mixte), Ann. Inst. Henri-Poincaré 32 (1996) 211–230.
[9] Zygmund A., Trigonometric Series, Cambridge University Press, 1959.