Statistical Inference for Stochastic Processes **5:** 321–333, 2002.

General Autoregressive Models with Long-memory Noise

MOHAMED BOUTAHAR
Département de mathématiques, case 901, 163 avenue de luminy, 13288 Marseille Cedex 9, and GREQAM, vieille charité, 13002 Marseille, France

Abstract. We give the limiting distribution of the least-squares estimator in the general autoregressive model driven by a long-memory process. We prove that with an appropriate normalization the estimation error converges, in distribution, to a random vector which contains: (1) a stochastic component, due to the presence of the unstable roots, which are multiple Wiener–Itô integrals and a non-linear functionals of stochastic integrals with respect to a Brownian motion; (2) a constant component due to the stable roots; (3) a stochastic component, due to the presence of the explosive roots, which is a mixture of normal distributions.

AMS 2000 Subject Classification: Primary 62M10, 62E20; secondary 60F17.

Key words: fractional Brownian motion, general autoregressive model, least-squares estimator, long-memory process, multiple Wiener–Itô integral, standard Brownian motion, stochastic integral.

1. Introduction

Consider the univariate autoregressive model

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

where y_t is the tth observation on the dependent variable, $y_t = 0$ if $t \leqslant 0$, and ε_t is a disturbance assumed to be a stationary Gaussian process with regularly varying spectral density $f(\lambda)$ of the form

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

where L is a slowly varying function (i.e. $L(na)/L(n) \to 1$ as $n \to \infty$ for any $a > 0$), bounded in all finite intervals and f is integrable on $[-\pi, \pi]$. It is well known that ε_t can be written as the Fourier transform of Gaussian random measure, that is,

$$\varepsilon_t = \int_{-\pi}^{\pi} e^{it\lambda} f^{1/2}(\lambda) W(\mathrm{d}\lambda), \tag{3}$$

where $W(.)$ is the Gaussian random measure corresponding to a white noise. $\beta = (a_1, \ldots, a_p)'$ is the unknown parameter which is estimated by the least-squares estimator (L.S.E.):

$$\beta_n = \left(\sum_{k=1}^{n} \Phi_{k-1}\Phi'_{k-1}\right)^{-1} \sum_{k=1}^{n} \Phi_{k-1} y_k, \tag{4}$$

where $\Phi_{t-1} = (y_{t-1}, \ldots, y_{t-p})'$. Recall that the least-squares estimation error satisfies

$$\beta_n - \beta = \left(\sum_{k=1}^{n} \Phi_{k-1}\Phi'_{k-1}\right)^{-1} \sum_{k=1}^{n} \Phi_{k-1} \varepsilon_k. \tag{5}$$

Models having a long-memory disturbance received considerable attention by researchers from various disciplines. The monograph by Beran [1] provides an updated survey of recent developments of long-memory processes in statistics (see also [15, 20]). Yajima [21, 22] considers the L.S.E. in the regression model with deterministic design, he proves the strong consistency of the L.S.E. and its limiting normal distribution. If model (1) is stationary, which is the case when the characteritic polynomial $\varphi(z) = 1 - a_1 z - \cdots - a_p z^p$ is stable (i.e. $\varphi(z) = 0 \Longrightarrow |z| > 1$), then β can also be estimated by other methods. Dahlhaus [6] considers the maximum likelihood estimator and proves its strong consistency and asymptotic normality by assuming that model (1) is Gaussian. Under the latter assumption, Giraitus and Taqqu [9] obtained the same results as Dahlhaus [6] for the Whittle estimator. If model (1) is non-stationary, which is the case if $\varphi(z)$ is unstable (i.e. $\varphi(z) = 0 \Longrightarrow |z| \geqslant 1$), then the above results no longer hold.

The almost sure properties of β_n were studied by Lai and Wei [11] when (ε_t) is a martingale difference sequence. They showed that β_n is strongly consistent without imposing any assumption on the roots of the polynomial $\varphi(z)$ (i.e. a general model). The limiting distribution of β_n for the unstable model (i.e. $\varphi(z) = 0 \Longrightarrow |z| \geqslant 1$) is given in [4]. Chan and Terrin [5] established the limiting distribution of β_n when the polynomial $\varphi(z)$ is unstable and the disturbance is a long-memory process. The fractional integrated autoregressive moving average (ARFIMA), popular in econometrics, indicates the usefulness of these theoretical developments (see [7, 16]).

We can write model (1) in a multivariate form

$$\Phi_n = \mathbf{A}\Phi_{n-1} + e_n, \tag{6}$$

where $\mathbf{A}$ is the companion matrix of the polynomial $\varphi(z)$, $e_n = (\varepsilon_n, 0, \ldots, 0)'$.

Model (6) was studied by many authors. Duflo et al. [8] considered model (6) with arbitrary matrix $\mathbf{A}$. They proved the consistency of the L.S.E. by assuming that the sequence (e_n) is a white noise. The limiting distribution of the L.S.E. was given by Touati [17] for the stable-explosive (i.e when some eigenvalues are within the unit circle and others are ouside) model, and by Touati [18] for the general model when (e_n) is an i.i.d. sequence.

Tsay and Tiao [19] considered a multivariate ARMA model with (e_n) a martingale difference sequence. They assumed that the characteristic polynomial of the model has roots on or outside the unit circle and proved the consistency of the L.S.E.

To study the cointegration of time series, Jeganathan [10] considered model (6) with a more general matrix $\mathbf{A}$ and (e_n) a fractionally integrated process. He gave a procedure to identify the approximate unit eigenvalues of matrix $\mathbf{A}$ based on a Wald-type approach. However the null and the contiguous alternatives that he considered are asymptotically equivalent. Moreover under the null, the eigenvalues of matrix $\mathbf{A}$ are bounded in absolute value by unity. When a contiguous alternative is accepted, then eigenvalues of $\mathbf{A}$ estimated close to unity in absolute values are adjusted to unity. Consequently, the final model that he retained is such that the eigenvalues of $\mathbf{A}$ are either on (and represent the non-stationary trend) or inside (and represent the cointegrating relationship) the unit circle. In model (6) above, we do not impose Jeganathan's condition on matrix $\mathbf{A}$ which can have eigenvalues greater than unity in absolute value. Moreover eigenvalues outside the unit circle are not necessary close to unity and will be called explosive eigenvalues.

The purpose of this paper is to extend the work of [4, 5] by letting the characteristic polynomial to be arbitrary. In other words, we do not make any assumption on the roots of $\varphi(z)$ (and consequently any assumption on the eigenvalues of its companion matrix $\mathbf{A}$).

The paper is organized as follows. Section 2 studies the explosive model (i.e. $\varphi(z) = 0 \Longrightarrow |z| < 1$), and gives the paper's main contribution namely: the consistency of β_n and its non-Gaussian limiting distribution. Section 3 considers the general model and gives the limiting distribution of β_n.

2. Explosive Model

In this section we assume that the polynomial $\varphi(z)$ is explosive (i.e. $\varphi(z) = 0 \Longrightarrow |z| < 1$). Denote by $\xrightarrow{L_p}$ and $\xrightarrow{\mathcal{L}}$ the convergence in $L_p(\Omega)$ and in law respectively, $||.||$ stands for the Euclidian norm, for a given matrix $\mathbf{A}$ we define $||\mathbf{A}|| = \sup_{||x||=1} ||\mathbf{A}x||$; and for a given random matrix or vector X, we denote its norm in $L_p(\Omega)$ by $||X||_p = (E(||X||^p))^{1/p}$.

$X \rightsquigarrow N_p(m, \Sigma)$ means that X is a p-dimensional Gaussian random vector with mean m and covariance Σ.

THEOREM 2.1.

$$\mathbf{A}^{-n}\Phi_n \xrightarrow{L_2} L = \int_{-\pi}^{\pi} (e^{-i\lambda}\mathbf{A} - \mathbf{I}_p)^{-1} f^{1/2}(\lambda) W(\mathrm{d}\lambda) e_1. \tag{7}$$

The random variable $x'L$ has a continuous distribution for all $x \in \mathbb{R}^p - \{0\}$.

$$\mathbf{A}^{-n} \sum_{k=1}^{n} \Phi_{k-1}\Phi_{k-1}' \mathbf{A}^{-n'} \xrightarrow{L_1} \mathbf{\Sigma}_2 = \sum_{k=1}^{\infty} \mathbf{A}^{-k} L L' \mathbf{A}^{-k'}. \tag{8}$$

Moreover

$$P(\mathbf{\Sigma}_2 > 0) = 1. \tag{9}$$

Proof. From (1) and (3),

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

Clearly

$$\sum_{k=1}^{n} \mathbf{A}^{-k} e^{ik\lambda} \to (e^{-i\lambda}\mathbf{A} - \mathbf{I}_p)^{-1}$$

pointwise, and

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

hence

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

consequently the convergence (7) follows by the L_2-continuity of the stochastic integrals.

For all $x \in \mathbb{R}^p - \{0\}$, $x'L$ is a Gaussian random variable with zero mean and variance $\int_{-\pi}^{\pi} \left| x'(e^{i\lambda}\mathbf{A} - \mathbf{I}_p)^{-1} e_1 \right|^2 f(\lambda)\,\mathrm{d}\lambda$, then it has a continuous distribution provided that it is non-degenerate, if we assume that $\mathrm{var}(x'L) = 0$ then

$$x'(\mathbf{I}_p e^{i\bullet} - \mathbf{A})^{-1} e_1 = 0, \quad \text{almost everywhere on } [0, \pi]. \tag{10}$$

Then there exist $(\lambda_j, 1 \leqslant j \leqslant p)$, a p distinct reals in $[0, \pi]$, such that

$$x'(\mathbf{I}_p e^{i\lambda_j} - \mathbf{A})^{-1} e_1 = 0, \qquad 1 \leqslant j \leqslant p; \tag{11}$$

now let

$$Z_j = [\varphi^*(e^{i\lambda_j})]^{-1}(e^{i(p-1)\lambda_j}, e^{i(p-2)\lambda_j}, \ldots, e^{i\lambda_j}, 1)',$$

where $\varphi^*(z) = z^p\varphi(z^{-1})$. Since the roots of $\varphi^*(z)$ are strictly outside the unit circle, it follows that $\varphi^*(e^{i\lambda_j}) \neq 0$ and hence the Z_j are well-defined; moreover $\det(Z_1, \ldots, Z_p) \neq 0$ since $(Z_1, \ldots, Z_p)$ is the Vandermonde matrix. It is easy to see that

$$(\mathbf{I}_p e^{i\lambda_k} - \mathbf{A}) Z_k = e_1 \quad \text{or} \quad Z_k = (\mathbf{I}_p e^{i\lambda_k} - \mathbf{A})^{-1} e_1,$$

hence (11) implies that

$$x'Z_k = 0, \quad \forall 1 \leqslant k \leqslant p,$$

and this implies that x is orthogonal to the whole space C^p and must be equal to zero.

To prove (8) we will adapt the proof of [11, Theorem 2] and use the following result:

$$\left|\left|\mathbf{A}^{-n}\right|\right| = \mathrm{O}(\rho^n n^{\nu-1}), \tag{12}$$

where

$$\rho = \max_{1 \leqslant j \leqslant p} \left|\lambda_j\right|, \qquad \nu = \max\{m_j, \left|\lambda_j\right| = \rho\}, \quad (\lambda_j, 1 \leqslant j \leqslant p)$$

are the eigenvalues of the matrix $\mathbf{A}^{-1}$, and m_j is the multiplicity of the eigenvalue λ_j.

Let $L_n = \mathbf{A}^{-n}\Phi_n$, $\mathbf{F}_n = \sum_{i=1}^{n} \mathbf{A}^{-i} L_n L_n' \mathbf{A}^{-i'}$, $\Delta_n = \mathbf{A}^{-n} \sum_{k=1}^{n} \Phi_{k-1}\Phi_{k-1}' \mathbf{A}^{-n'} - \mathbf{F}_n$. Then

$$\begin{aligned}
||\Delta_n||_1 &= \left|\left|\sum_{i=1}^{n} \mathbf{A}^{-i} L_{n-i} L_{n-i}' \mathbf{A}^{-i'} - \mathbf{F}_n\right|\right|_1 \\
&\leqslant \sum_{i=1}^{n} ||\mathbf{A}^{-i}|| \left|\left|\mathbf{A}^{-i'}\right|\right| \, ||L_{n-i}L_{n-i}' - L_n L_n'||_1 \\
&\leqslant \sum_{i=1}^{n} ||\mathbf{A}^{-i}|| \left|\left|\mathbf{A}^{-i'}\right|\right| (||L_{n-i}||_2 + ||L_n||_2) \, ||L_n - L_{n-i}||_2 \\
&\to 0,
\end{aligned}$$

since $||L_n||_2 < \infty$, $||L_n - L_{n-i}||_2 \to 0$ and by using (12).

$$\begin{aligned}
||\mathbf{F}_n - \mathbf{\Sigma}_2||_1 &= \left|\left|\sum_{i=1}^{n} \mathbf{A}^{-i}(L_n L_n' - LL')\mathbf{A}^{-i'} + \sum_{i=n+1}^{\infty} \mathbf{A}^{-i} LL' \mathbf{A}^{-i'}\right|\right|_1 \\
&\leqslant \left|\left|L_n L_n' - LL'\right|\right|_1 \sum_{i=1}^{n} ||\mathbf{A}^{-i}|| \left|\left|\mathbf{A}^{-i'}\right|\right| \\
&\quad + \left|\left|LL'\right|\right|_1 \sum_{i=n+1}^{\infty} ||\mathbf{A}^{-i}|| \left|\left|\mathbf{A}^{-i'}\right|\right| \\
&\to 0,
\end{aligned} \tag{13}$$

by using (12) and the convergence

$$\left|\left|L_n L_n' - LL'\right|\right|_1 \leqslant (||L_n||_2 + ||L||_2) \, ||L_n - L||_2 \to 0.$$

The proof of (9) is exactly the same as in [11, Theorem 2]. □

The following theorem gives the limiting distribution of the L.S.E. β_n and proves its consistency.

THEOREM 2.2.

$$(\mathbf{A}^n)'(\beta_n - \beta) \xrightarrow{\mathcal{L}} \mathbf{\Sigma}_2^{-1} N_2, \tag{14}$$

where $N_2 = \mathcal{A}_1 Z_2$, $\mathcal{A}_1 = \int_{-\pi}^{\pi} (\mathbf{I}_p - e^{-\mathrm{i}\lambda}\mathbf{A})^{-1} f(\lambda)^{1/2} W(\mathrm{d}\lambda)$, $Z_2 \rightsquigarrow N_p(0, \tilde{\Sigma}_2)$, $\tilde{\Sigma}_2 = \int_{-\pi}^{\pi} (\mathbf{I}_p e^{-i\lambda} - \mathbf{A})^{-1} e_1 e_1' (\mathbf{I}_p e^{i\lambda} - \mathbf{A}')^{-1} f(\lambda)\,\mathrm{d}\lambda$, $e_1 = (1, 0, \ldots, 0)'$, *and* Z_2 *and* $\mathcal{A}_1$ *are independent.*
The L.S.E. β_n *is consistent with the speed of convergence*

$$||\beta_n - \beta||^2 = \mathrm{o}_p(d(n)\rho^{2n} n^{2(\nu-1)}), \tag{15}$$

where ρ *and* ν *are given by (12), and for every sequence* $d(n) \uparrow \infty$.

Remark. If $p = 1$, then $N_2 = \sigma_1^2 X_1 X_2$, where $X_1 \rightsquigarrow N(0, 1)$, $X_2 \rightsquigarrow N(0, 1)$, X_1 and X_2 are independent, and $\sigma_1^2 = \int_{-\pi}^{\pi} \left|1 - ae^{i\lambda}\right|^{-2} f(\lambda)\,\mathrm{d}\lambda$, hence the characteristic function of N_2 is given by

$$\varphi N_2(t) = E\left(e^{-\sigma_1^4 X_2^2 t^2/2}\right),$$

the distribution of N_2 is called a mixture of normal distributions (see, [14]).
We will use the theorem of Riemann–Lebesgue (see [23, Theorem (4.4)]), which states that the Fourier coefficient c_n of every integrable function is such that $c_n \to 0$ as $|n| \to \infty$; and the following two lemmas. Hereafter, let $\mathbf{R}(z) = (z\mathbf{I}_p - \mathbf{A})^{-1}$.

LEMMA 2.3. $\forall g_1$ *and* $g_2 \in L^2([-\pi, \pi], \mathrm{d}\lambda)$; *such that* g_1, g_2 *are Hermitians, we have*

$$\begin{aligned} 2\int'' & g_1(\lambda_1) g_2(\lambda_2) W(\mathrm{d}\lambda_1) W(\mathrm{d}\lambda_2) \\ &= 2\alpha \int'' g_2(\lambda_1) g_2(\lambda_2) W(\mathrm{d}\lambda_1) W(\mathrm{d}\lambda_2) + \\ &\quad + \int_{-\pi}^{\pi} g_1(\lambda) W(\mathrm{d}\lambda) \int_{-\pi}^{\pi} g_2(\lambda) W(\mathrm{d}\lambda) - \alpha \left(\int_{-\pi}^{\pi} g_2(\lambda) W(\mathrm{d}\lambda) \right)^2, \end{aligned}$$

where $\alpha = \langle g_2, g_1 \rangle_{L^2([-\pi,\pi],\,\mathrm{d}\lambda)} / \, ||g_2||^2_{L^2([-\pi,\pi],\,\mathrm{d}\lambda)}$, *and* $\int''$ *is the multiple Wiener–Itô integral defined in* [13, § 4] *denoted by* $I_G(f)$.

The proof of this lemma follows immediately by application of the Itô formula (see [13, Theorem 4.2]) to the orthogonal system $(g_2, g_1 - \alpha g_2)$.

LEMMA 2.4. *Let* vec $(\mathcal{A}_1)$ *stands for the column vector obtained by stacking the columns of the matrix* $\mathcal{A}_1$,

$$Z_n = \left((\mathrm{vec}(\mathcal{A}_1))', \left(\int_{-\pi}^{\pi} \mathbf{R}(e^{-i\lambda}) e^{in\lambda} f(\lambda)^{1/2} W(\mathrm{d}\lambda) e_1 \right)' \right)'.$$

Then

$$Z_n \xrightarrow{\mathcal{L}} Z = \left((\mathrm{vec}(\mathcal{A}_1))', Z_2'\right)'.$$

Proof. It is sufficient to prove that for all $u \in \mathbb{R}^{p^2+p}$,

$$u'Z_n \xrightarrow{\mathcal{L}} u'Z.$$

Let $u = (u_1', u_2')'$, $u_1 = \mathrm{vec}(u_1(j,k), j = 1, \ldots, p, k = 1, \ldots, p)$, $u_2 = (u_2(j), j = 1, \ldots, p)'$; now observe that

$$u'Z_n = \int_{-\pi}^{\pi} H(\lambda) W(\mathrm{d}\lambda),$$

where

$$H(\lambda) = \left(\sum_{j,k} u_1(j,k) \mathbf{R}_{j,k}(e^{i\lambda}) e^{-i\lambda} + \sum_j u_2(j) \mathbf{R}_{j,1}(e^{-i\lambda}) e^{in\lambda} \right) f^{1/2}(\lambda),$$

and $\mathbf{R}_{j,k}(z)$ is the (j,k)th term of the matrix $\mathbf{R}(z)$, it follows that $u'Z_n$ is a Gaussian random variable with zero mean. We will prove that

$$\phi_{u'Z_n}(t) \longrightarrow \phi_{u'Z}(t),$$

where $\phi_{u'Z_n}(t)$ and $\phi_{u'Z}(t)$ are the characteristic functions of $u'Z_n$ and $u'Z$.

$$\phi_{u'Z_n}(t) = e^{-\mathrm{var}(u'Z_n)t^2/2},$$

where $\mathrm{var}(u'Z_n) = u' \,\mathrm{var}(Z_n) u = u' \mathbf{\Gamma}_n u$,

$$\mathbf{\Gamma}_n = \begin{pmatrix} \mathbf{\Gamma}_1 & \mathbf{\Gamma}_{1,2}(n) \\ \mathbf{\Gamma}_{1,2}'(n) & \tilde{\Sigma}_2 \end{pmatrix}.$$

$$\begin{aligned} \mathbf{\Gamma}_1 &= \mathrm{var}(\mathrm{vec}(\mathcal{A}_1)), \mathbf{\Gamma}_{1,2}(n) \\ &= \mathrm{cov}\left(\mathrm{vec}(\mathcal{A}_1), \int_{-\pi}^{\pi} \mathbf{R}(e^{-i\lambda}) e_1 e^{in\lambda} f(\lambda)^{1/2} W(\mathrm{d}\lambda)\right), \end{aligned}$$

hence

$$\phi_{u'Z_n}(t) = e^{-(u_1'\mathbf{\Gamma}_1 u_1 + u_2'\tilde{\Sigma}_2 u_2 + u_1'\mathbf{\Gamma}_{1,2}(n) u_2 + u_2'\mathbf{\Gamma}_{1,2}'(n) u_1)(t^2/2)}.$$

$\forall (j,k,l)$

$$\begin{aligned} &E\left(\left(\int_{-\pi}^{\pi} \mathbf{R}_{j,k}(e^{i\lambda}) e^{-i\lambda} f^{1/2}(\lambda) W(\mathrm{d}\lambda)\right) \times \right. \\ &\qquad \left. \times \left(\int_{-\pi}^{\pi} \mathbf{R}_{l,1}(e^{-i\lambda}) e^{in\lambda} f^{1/2}(\lambda) W(\mathrm{d}\lambda)\right)\right) \\ &\quad = \int_{-\pi}^{\pi} \mathbf{R}_{j,k}(e^{i\lambda}) e^{-i\lambda} \mathbf{R}_{l,1}(e^{i\lambda}) e^{-in\lambda} f(\lambda)\, \mathrm{d}\lambda \\ &\quad \longrightarrow 0, \end{aligned}$$

by Riemann–Lebesgue's theorem since $\mathbf{R}_{j,k}(e^{-i\lambda})e^{-i\lambda}\mathbf{R}_{l,1}(e^{-i\lambda})f(\lambda)$ is integrable on $[-\pi, \pi]$, therefore

$$\phi_{u'Z_n}(t) \longrightarrow e^{-\left(u_1'\boldsymbol{\Gamma}_1 u_1 + u_2'\tilde{\Sigma}_2 u_2\right)(t^2/2)} = \phi_{u'Z}(t).$$

□

Proof of Theorem 2.2. We shall prove that

$$\mathbf{A}^{-n}\sum_{1}^{n}\Phi_{k-1}\varepsilon_k \xrightarrow{\mathcal{L}} N_2. \tag{16}$$

$$\begin{aligned}
\mathbf{A}^{-n}\sum_{1}^{n}\Phi_{k-1}\varepsilon_k &= \mathbf{A}^{-n}\sum_{t=1}^{n}\sum_{k=1}^{t-1}\mathbf{A}^{t-k-1}\varepsilon_t\varepsilon_k e_1 \\
&= \mathbf{A}^{-n}\sum_{t=1}^{n}\sum_{k=1}^{t-1}\mathbf{A}^{t-k-1}\bigg[2\int''_{[-\pi,\pi]^2} e^{ik\lambda_1}e^{it\lambda_2}\times \\
&\quad \times f^{1/2}(\lambda_1)f^{1/2}(\lambda_2)W(\mathrm{d}\lambda_1)W(\mathrm{d}\lambda_2) + \\
&\quad + \int_{-\pi}^{\pi} e^{i(t-k)\lambda} f(\lambda)\,\mathrm{d}\lambda\bigg]e_1, \text{ by the Itô formula} \\
&= T_1 + T_2.
\end{aligned} \tag{17}$$

Consider first the second term in the right-hand side of (17); after some computations we obtain

$$\begin{aligned}
T_2 &= \int_{-\pi}^{\pi}\mathbf{A}^{-n}\sum_{t=1}^{n}\sum_{k=1}^{t-1}\mathbf{A}^{t-k-1}e^{i(t-k)\lambda}f(\lambda)\,\mathrm{d}\lambda e_1 \\
&= \int_{-\pi}^{\pi}(\mathbf{A} - e^{-i\lambda}\mathbf{I}_p)^{-1}(\mathbf{I}_p - e^{i\lambda}\mathbf{A})^{-1}\mathbf{A}^{-n}f(\lambda)\,\mathrm{d}\lambda e_1 + \\
&\quad + \int_{-\pi}^{\pi}(e^{-i\lambda}\mathbf{I}_p - \mathbf{A})^{-1}(\mathbf{I}_p - e^{i\lambda}\mathbf{A})^{-1}e^{in\lambda}f(\lambda)\,\mathrm{d}\lambda e_1 + \\
&\quad + \int_{-\pi}^{\pi} n(\mathbf{I}_p e^{-i\lambda} - \mathbf{A})^{-1}\mathbf{A}^{-n}f(\lambda)\,\mathrm{d}\lambda e_1 \\
&= I_1 + I_2 + I_3.
\end{aligned} \tag{18}$$

By using (12), it follows that

$$I_1 = \mathrm{o}(n^{-\delta}) \quad \text{and} \quad I_3 = \mathrm{o}(n^{-\delta}), \quad \forall\delta > 0. \tag{19}$$

Let $g(\lambda) = (\mathbf{I}_p e^{-i\lambda} - \mathbf{A})^{-1}(\mathbf{I}_p - e^{i\lambda}\mathbf{A})^{-1}e_1 f(\lambda)$, since $||g(\lambda)|| \leqslant Cf(\lambda)$, for some positive constant C and for all $\lambda \in [-\pi, \pi]$, it follows that all the components of $g(\lambda)$ are integrable on $[-\pi, \pi]$, consequently the Riemann–Lebesgue's theorem implies that

$$\begin{aligned} I_2 &= \int_{-\pi}^{\pi} (\mathbf{I}_p e^{-i\lambda} - \mathbf{A})^{-1} (\mathbf{I}_p - e^{i\lambda}\mathbf{A})^{-1} e^{in\lambda} f(\lambda)\, \mathrm{d}\lambda e_1 \\ &= \mathrm{o}(1). \end{aligned} \tag{20}$$

From (18) to (20) we deduce that

$$T_2 = \mathrm{o}(1). \tag{21}$$

The first term in (17) is equal to

$$\begin{aligned} T_1 &= 2 \int''_{[-\pi,\pi]^2} \mathbf{A}^{-n} \sum_{t=1}^{n} \mathbf{A}^{t-1} e^{it\lambda_2} \sum_{k=1}^{t-1} \mathbf{A}^{-k} e^{ik\lambda_1} \times \\ &\quad \times f^{1/2}(\lambda_1) f^{1/2}(\lambda_2) W(\mathrm{d}\lambda_1) W(\mathrm{d}\lambda_2) e_1, \end{aligned} \tag{22}$$

after some computations we obtain

$$\begin{aligned} &\mathbf{A}^{-n} \sum_{t=1}^{n} \mathbf{A}^{t-1} e^{it\lambda_2} \sum_{k=1}^{t-1} \mathbf{A}^{-k} e^{ik\lambda_1} \\ &\quad = (\mathbf{I}_p - e^{i\lambda_1}\mathbf{A}^{-1})^{-1} (\mathbf{I}_p - e^{i\lambda_2}\mathbf{A})^{-1} e^{i(\lambda_1+\lambda_2)} \mathbf{A}^{-(n+1)} + \\ &\quad\quad + (\mathbf{I}_p - e^{-i\lambda_1}\mathbf{A})^{-1} (\mathbf{I}_p - e^{i\lambda_2}\mathbf{A})^{-1} e^{i(n+1)\lambda_2} - \\ &\quad\quad - (\mathbf{I}_p - e^{i\lambda_1}\mathbf{A}^{-1})^{-1} \mathbf{A}^{-(n+1)} e^{i(\lambda_1+\lambda_2)} \frac{e^{in(\lambda_1+\lambda_2)} - 1}{e^{i(\lambda_1+\lambda_2)} - 1}. \end{aligned}$$

By using (12), it follows that

$$T_1 = \mathrm{o}_p(1) + N_n, \tag{23}$$

where

$$\begin{aligned} N_n &= 2 \int''_{[-\pi,\pi]^2} (\mathbf{I}_p - e^{-i\lambda_1}\mathbf{A})^{-1} (\mathbf{I}_p - e^{i\lambda_2}\mathbf{A})^{-1} e^{i(n+1)\lambda_2} \times \\ &\quad \times f^{1/2}(\lambda_1) f^{1/2}(\lambda_2) W(\mathrm{d}\lambda_1) W(\mathrm{d}\lambda_2) e_1. \end{aligned} \tag{24}$$

Now, we will prove that

$$\begin{aligned} N_n &= \mathrm{o}_p(1) + \int_{-\pi}^{\pi} (\mathbf{I}_p - e^{-i\lambda}\mathbf{A})^{-1} f^{1/2}(\lambda) W(\mathrm{d}\lambda) \times \\ &\quad \times \int_{-\pi}^{\pi} (\mathbf{I}_p - e^{i\lambda}\mathbf{A})^{-1} e^{i(n+1)\lambda} f^{1/2}(\lambda) W(\mathrm{d}\lambda) e_1. \end{aligned} \tag{25}$$

The (j,k)th term of the matrix in the right-hand side of (24) is equal to

$$\begin{aligned} I_{j,k} &= \sum_{l=1}^{p} \int''_{[-\pi,\pi]^2} \times \\ &\quad \times e^{\mathrm{i}\lambda_1} \mathbf{R}_{j,l}(e^{\mathrm{i}\lambda_1}) \mathbf{R}_{l,k}(e^{-\mathrm{i}\lambda_2}) e^{in\lambda_2} f^{1/2}(\lambda_1) f^{1/2}(\lambda_2) W(\mathrm{d}\lambda_1) W(\mathrm{d}\lambda_2), \end{aligned}$$

where $\mathbf{R}_{j,l}(z)$ is the (j,l)th term of the matrix $\mathbf{R}(z)$. Application of Lemma 2.3 with $g_1(\lambda) = e^{i\lambda}\mathbf{R}_{j,l}(e^{i\lambda}) f^{1/2}(\lambda)$, $g_2(\lambda) = \mathbf{R}_{l,k}(e^{-i\lambda}) f^{1/2}(\lambda) e^{in\lambda}$ implies that

$$I_{j,k} = \sum_{l=1}^{p} \alpha_n I_1(l) + I_2(l) - \alpha_n I_3(l),$$

where

$$\begin{aligned}\alpha_n &= \int_{-\pi}^{\pi} \mathbf{R}_{j,l}(e^{-i\lambda})\mathbf{R}_{l,k}(e^{-i\lambda}) e^{i(n-1)\lambda} f(\lambda)\,\mathrm{d}\lambda \times \\ &\quad \times \int_{-\pi}^{\pi} \left|\mathbf{R}_{l,k}(e^{-i\lambda})\right|^2 f(\lambda)\,\mathrm{d}\lambda \longrightarrow 0,\end{aligned}$$

$$\begin{aligned}I_1(l) &= 2\int'' \mathbf{R}_{l,k}(e^{-i\lambda_1})\mathbf{R}_{l,k}(e^{-i\lambda_2}) e^{in(\lambda_1+\lambda_2)} \times \\ &\quad \times f^{1/2}(\lambda_1) f^{1/2}(\lambda_2) W(\mathrm{d}\lambda_1) W(\mathrm{d}\lambda_2), \\ I_2(l) &= \left(\int e^{i\lambda}\mathbf{R}_{j,l}(e^{i\lambda}) f^{1/2}(\lambda) W(\mathrm{d}\lambda)\right)\left(\int \mathbf{R}_{l,k}(e^{-i\lambda}) f^{1/2}(\lambda) e^{in\lambda} W(\mathrm{d}\lambda)\right), \\ I_3(l) &= \left(\int \mathbf{R}_{l,k}(e^{-i\lambda}) f^{1/2}(\lambda) e^{in\lambda} W(\mathrm{d}\lambda)\right)^2;\end{aligned}$$

since $\mathbf{R}_{l,k}(e^{-i\lambda_1})\mathbf{R}_{l,k}(e^{-i\lambda_2}) e^{in(\lambda_1+\lambda_2)} f^{1/2}(\lambda_1) f^{1/2}(\lambda_2)$ is symmetric

$$\begin{aligned}||\alpha_n I_1(l)||_2^2 &= \alpha_n^2 \left(\int_{-\pi}^{\pi} \left|\mathbf{R}_{l,k}(e^{-i\lambda})\right|^2 f(\lambda)\,\mathrm{d}\lambda\right)^2 \longrightarrow 0, \\ ||\alpha_n I_3(l)||_1 &= |\alpha_n| \int_{-\pi}^{\pi} \left|\mathbf{R}_{l,k}(e^{-i\lambda})\right|^2 f(\lambda)\,\mathrm{d}\lambda \longrightarrow 0;\end{aligned}$$

therefore,

$$\begin{aligned}I(j,k) &= \mathrm{o}_p(1) + \sum_{l=1}^{p} \left(\int \mathbf{R}_{j,l}(e^{i\lambda}) e^{-i\lambda} f^{1/2}(\lambda) W(\mathrm{d}\lambda)\right) \times \\ &\quad \times \left(\int \mathbf{R}_{l,k}(e^{-i\lambda}) f^{1/2}(\lambda) e^{in\lambda} W(\mathrm{d}\lambda)\right),\end{aligned}$$

and the last term is the (j,k)th term of the matrix in the right-hand side of (25). Consequently,

$$N_n = \mathrm{o}_p(1) + \mathcal{A}_1 \int \mathbf{R}(e^{-i\lambda})\, e_1 f^{1/2}(\lambda) e^{in\lambda} W(\mathrm{d}\lambda).$$

The components of the last vector are continuous functionals of the components of the vector Z_n defined in Lemma 2.4, hence the joint convergence (16) follows from (17), (21), (23), the Lemma 2.4 and the continuous mapping theorem [2, Theorem 5.1].

Finally the convergence (14) is obtained from (8) and (16). To prove (15), let $\lambda_{\min}(\mathbf{A})$ [resp. $\lambda_{\max}(\mathbf{A})$] denotes the minimum [resp. maximum] eigenvalue of the matrix $\mathbf{A}$. We use (14) to obtain

$$[d(n)]^{-1}(\beta_n - \beta)'\mathbf{A}^n(\mathbf{A}^n)'(\beta_n - \beta) \xrightarrow{P} 0,$$

which implies that

$$\lambda_{\min}(\mathbf{A}^n(\mathbf{A}^n)')\,||\beta_n - \beta||^2 = \mathrm{o}_p(d(n)),$$

moreover, using (12),

$$\frac{1}{\lambda_{\min}(\mathbf{A}^n(\mathbf{A}^n)')} = \lambda_{\max}((\mathbf{A}^{-n})'\mathbf{A}^{-n}) = \left|\left|\mathbf{A}^{-n}\right|\right|^2 = \mathrm{O}(\rho^{2n}n^{2(\nu-1)}).$$

□

3. General Model

In this section we consider the general $AR(p)$ model. This means that we do not make any assumption on the roots of the characteristic polynomial $\varphi(z)$. To obtain the limiting distribution of β_n we use Lai and Wei's [11] classical technique: (i) To transform the original model into various components corresponding to the location of their roots relative to the unit circle, (ii) To analyze each component separately.

The polynomial $\varphi(z)$ can be written as

$$\varphi(z) = \varphi_u(z)\varphi_e(z), \qquad \deg(\varphi_u) = p_1, \qquad \deg(\varphi_e) = r, \qquad p = p_1 + r,$$

where $\varphi_e(z)$ is an explosive polynomial and $\varphi_u(z)$ is an unstable polynomial which can be written as

$$\varphi_u(z) = (1-z)^a(1+z)^b\prod_{m=1}^{l}(1 - 2z\cos\theta_m + z^2)^{d_m}\varphi_s(z),$$

$\varphi_s(z)$ is a stable polynomial, $\deg(\varphi_s) = q$, $p_1 = q + a + b + 2\sum_1^l d_m$. If we define

$$y_t^e = \varphi(z)\varphi_e^{-1}(z)y_t, \qquad y_t^u = \varphi(z)\varphi_u^{-1}(z)y_t,$$

then

$$\varphi_e(z)y_t^e = \varepsilon_t$$

(i.e. (y_t^e) is the explosive $AR(r)$ studied in Section 2), and

$$\varphi_u(z)y_t^u = \varepsilon_t$$

(i.e. (y_t^u) is the unstable $AR(p_1)$ considered in [4, 5]. There exists a non-singular matrix $\mathbf{M}$, (see [11]) such that

$$\mathbf{M}\Phi_{t-1} = (\Phi_{t-1}^{u'}, \Phi_{t-1}^{e'})',$$

where (Φ^u_{t-1}) and (Φ^e_{t-1}) are the regressor vectors corresponding to the unstable and the explosive model given by

$$\Phi^u_{t-1} = (y^u_{t-1}, \dots, y^u_{t-p_1})', \qquad \Phi^e_{t-1} = (y^e_{t-1}, \dots, y^e_{t-r})'.$$

Let

$$\mathcal{D}_n = \mathrm{diag}(A_n^{1/2}, \mathbf{I}_r), \qquad \mathcal{Q}_n = \mathrm{diag}(G_n Q, \mathbf{A}_e^{-n}),$$

$$A_n^{1/2} = \mathrm{diag}(n^{-H+1/2} L(n)^{-1/2}\mathbf{I}_a, L(n)^{-1/2}\mathbf{I}_{b+2\sum_{m=1}^{l} d_m}, n^{-1/2}\mathbf{I}_q),$$

where $\mathbf{A}_e$ is the companion matrix of the explosive polynomial $\varphi_e(z)$. The matrices Q and G_n are the same as in [4, 5]. Q is such that

$$Q\Phi^u_t = (\mathbf{u}'_t, \mathbf{v}'_t, \mathbf{x}'_t(1), \dots, \mathbf{x}'_t(l), \mathbf{z}'_t)', \qquad \mathbf{u}_t = (u_t, \dots, u_{t-a+1})',$$

$$\begin{aligned}
\mathbf{v}_t &= (v_t, \dots, v_{t-b+1})', \qquad \mathbf{x}_t(m) = (x_t(m), \dots, x_{t-2d_m+1}(m))',\\
\mathbf{z}_t &= (z_t, \dots, z_{t-q+1})',
\end{aligned}$$

where

$$\begin{aligned}
u_t &= \varphi_u(z)(1-z)^{-a} y^u_t, \qquad v_t = \varphi_u(z)(1+z)^{-b} y^u_t,\\
x_t(m) &= \varphi_u(z)(1 - 2z\cos\theta_m + z^2)^{-d_m} y^u_t, \quad 1 \leqslant m \leqslant l,\\
z_t &= \varphi_u(z)(\varphi_s(z))^{-1} y^u_t.
\end{aligned}$$

$G_n = \mathrm{diag}\,(\mathbf{J}_n, \mathbf{K}_n, \mathbf{L}_n(1), \dots, \mathbf{L}_n(l), \mathbf{M}_n)$, $\mathbf{J}_n = \mathrm{diag}(n^{-a+j-1}, 1 \leqslant j \leqslant a)M$, $\mathbf{K}_n = \mathrm{diag}(n^{-b+j-1}, 1 \leqslant j \leqslant b)\tilde{M}$, $\mathbf{L}_n(m) = \mathrm{diag}(n^{-j}\mathbf{I}_2, 1 \leqslant j \leqslant d_m)\mathbf{C}_m$, $\mathbf{M}_n = \mathbf{I}_q$, the matrices M, $\tilde{M}$ and $\mathbf{C}_m$, $1 \leqslant m \leqslant l$ are given in [4].

Combining the preceding results and Theorem 6.1 of [5] we obtain the following theorem.

THEOREM 3.1. $(\mathbf{M}'\mathcal{Q}'_n)^{-1}(\beta_n - \beta)$

$$\xrightarrow{\mathcal{L}} \left((\mathbf{F}^{-1}\xi)', (\tilde{\mathbf{F}}^{-1}\eta)', (\mathbf{D}_1^{-1}\zeta_1)', \dots, (\mathbf{D}_l^{-1}\zeta_l)', (\Sigma_1^{-1}N_1)', (\Sigma_2^{-1}N_2)' \right)',$$

the matrices $\mathbf{F}$ $\tilde{\mathbf{F}}$, $\mathbf{D}_m$, *and the vectors* ξ, η, ζ_m, $1 \leqslant m \leqslant l$ *are the same as in* [5], $N_1 = \int_{-\pi}^{\pi} (e^{i\lambda}\mathbf{I}_p - \mathbf{A})^{-1} f(\lambda)\,\mathrm{d}\lambda e_1$, $\Sigma_1 = \int_{-\pi}^{\pi} (e^{i\lambda}\mathbf{I}_p - \mathbf{A})^{-1} e_1 e'_1 (e^{-i\lambda}\mathbf{I}_p - \mathbf{A}')^{-1} f(\lambda)\mathrm{d}\lambda$, $e_1 = (1, 0, \dots, 0)'$, Σ_2 *and* N_2 *are given by Theorems* 2.1 and 2.2.

References

1. Beran, J.: *Statistics for Long-memory Processes*, Chapman & Hall, London, 1994.
2. Billingsley, P.:*Convergence of Probability Measures*, Wiley, New York, 1968.
3. Boutahar, M. and Deniau, C.: A proof of asymptotic normality for some VARX Model, *Metrika* **42** (1995), 331–339.

4. Chan, N. H. and Wei, C. Z.: Limiting distributions of least-squares estimates of unstable autoregressive processes, *Ann. Statist.* **16** (1988), 367–401.
5. Chan, N. H. and Terrin, N.: Inference for unstable long-memory processes with applications to fractional unit root autoregressions, *Ann. Statist.* **23** (1995), 1662–1683.
6. Dahlhaus, R.: Efficient parameter estimation for self-similar processes, *Ann. Statist.* **17** (1989), 1749–1766.
7. Diebold , F. X.: Random walk versus fractional integration: power comparisons of scalar and joint tests of the variance-time function, *Technical Report.* **41**, 1988, Federal Reserve Board, Washington, D.C.
8. Duflo, M., Senoussi, R. and Touati, A.: Moindres carrés d'un Modéle Autorégressif, *Annales de L' I.H.P.* **27** (1991), 1–25.
9. Giriatis L. and Taqqu M.: Whittle estimator for finite-variance non-gaussian time series with long memory, *Ann. Statist.* **27** (1999), 178–203.
10. Jeganathan, P.: On asymptotic inference in cointegrated time series with fractionally integrated errors, *Econometric Theory* **15** (1999), 583–621.
11. Lai, T. L. and Wei, C. Z.: Asymptotic properties of general autoregressive models and strong consistency of least squares estimates of their parameters, *J. Multivar. Anal.* **13** (1983), 1–13.
12. Lai, T. L. and Wei, C. Z: Asymptotic properties of weighted sums with applications to stochastic regression in linear dynamic systems, In: P. R. Krishnaiah (ed.), *Multivariate Analysis*, Vol. VI, North-Holland, Amsterdam, 1985, pp. 375–393.
13. Major, P.: *Multiple Wiener–Itô integrals*, *Lecture Notes in Math.* 849, Springer, New York, 1981.
14. Hall, P. and Heyde, C. C.: *Martingales Limit Theory and Its Application*, Academic Press, New York, 1980.
15. Hallin, M., Taniguchi, M., Serroukh, A. and Choy, K.: Local asymptotic normality for regression models with long memory disturbance, *Ann. Statist.* **6** (1999), 2054–2080.
16. Robinson, P. M.: Time series with strong dependance, In: C. A. Sims (ed.), *Advances in Econometrics, Sixth World Congress*, Vol. 1, Cambridge University Press, Cambridge, 1994, pp. 47–95.
17. Touati, A.: Moindres carrés d'un Modèle Autorégressif mixte, *Annales de L' I.H.P.* **32** (1996), 211–230.
18. Touati, A.: Moindres carrés d'un Modèle Autorégressif général, Thesis, Université Paris-Sud.
19. Tsay, R. S. and Tiao, G. C.: Asymptotic properties of multivariate non-stationary processes with applications to autoregressions, *Ann. Statist.* **18** (1990), 220–250.
20. Viano, M. C., Deniau, C. and Oppenheim, G.: Long range dependence and mixing for discrete time fractional processes, *J. Time Ser. Anal.* **16-3** (1995), 323–338.
21. Yajima, Y.: On estimation of regression model with long-memory stationary errors, *Ann. Statist.* **16** (1988), 791–807.
22. Yajima, Y.: Asymptotic properties of the LSE in a regression model with long-memory stationary errors, *Ann. Statist.* **19** (1991), 158–177.
23. Zygmund , A.: *Trigonometric Series*, Cambridge University Press, Cambridge, 1959.