Annales Univ. Sci. Budapest., Sect. Comp. 22 (2003) 207-223

NEARLY NONSTATIONARY AR PROCESSES WITH MIXING INNOVATON

J. Kormos (Debrecen, Hungary)

Dedicated to Professor Karl-Heinz Indlekofer
on the occasion of his 60th birthday

Abstract. In this paper we show an interesting application P.C.B Philips' [14] result for AR(p) process. In the first point we give a summary of results concerning the nearly unstable models. In the second point the notion of mixing properties are investigated and a suitable transformation is introduced which allows us to prove results for the nearly unstable AR(1) process with autoregressive innovation having strong mixing property. In the third point the theorem and the proof are given.

1. Introduction

1.1. Classical results

Consider the autoregressive AR(p) model

$$(1)\qquad \begin{aligned} &X_k = \beta_1 X_{k-1} + \ldots + \beta_p X_{k-p} + \epsilon_k, \quad k = 1, 2, \ldots, \\ &X_0 = X_{-1} = \ldots = X_{1-p} = 0, \end{aligned}$$

where ϵ_k is the (unobservable) random disturbance (noise) at time k and $\beta_1, \ldots, \beta_p$ are unknown parameters. The least-squares estimator (LSE) of the parameter $\mathcal{B} = (\beta_1, \ldots, \beta_p)'$ based on the observation $X_1, \ldots, X_n$ is given by

$$\hat{\mathcal{B}}_n = \left(\sum_{k=1}^{n} \tilde{X}_{k-1} \tilde{X}'_{k-1} \right)^{-1} \sum_{k=1}^{n} X_k \tilde{X}_{k-1},$$

where

$$\tilde{X}_k = (X_k, X_{k-1}, \ldots, X_{k-p+1})'.$$

The polynomial ϕ defined by

$$\phi(z) = 1 - \beta_1 z - \ldots - \beta_p z^p \tag{2}$$

is called the characteristic polynomial of the AR(p) model (1).

When all roots of ϕ are outside the unit circle, the model (1) is said to be asymptotically stationary. Under the assumption that the ϵ_k's are independent and identically distributed (i.i.d.) with $\mathbb{E}\epsilon_k^2 = \sigma^2$ the LSE of $\hat{\mathcal{B}}_n$ is asymptotically normal

$$\left(\sum_{k=1}^{n} \tilde{X}_{k-1}\tilde{X}'_{k-1}\right)^{-1/2} (\hat{\mathcal{B}}_n - \mathcal{B}) \xrightarrow{\mathcal{D}} \mathcal{N}(0, I), \quad \text{as} \quad n \to \infty,$$

where $\xrightarrow{\mathcal{D}}$ denotes convergence in distribution and I is the unit matrix (Mann, Wald [12], Anderson [2]). By another normalization

$$\sqrt{n}(\hat{\mathcal{B}}_n - \mathcal{B}) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \Sigma^{-1}), \quad n \to \infty,$$

where the matrix Σ can be expressed by the help of σ^2.

When ϕ has no roots inside the unit circle but has at least one root on the unit circle the model is said to be unstable. It was shown by White ([15],[16]) that in case of the unstable AR(1) model

$$X_k = \beta X_{k-1} + \epsilon_k, \quad k \geq 1$$

with $\beta = 1$

$$n(\hat{\beta}_n - \beta) \xrightarrow{\mathcal{D}} \frac{\int_0^1 W(t)dW(t)}{\int_0^1 W^2(t)dt},$$

where $W(t)$, $t \geq 0$ is a standard Wiener process.

In the "explosive" case, when $|\beta| > 1$ the serial $(\hat{\beta}_n)_{n\geq 1}$ limit distribution of $\hat{\beta}_n$ is not asymptotically normal. For example, if $\epsilon_1 \sim \mathcal{N}(0,1)$ then

$$n(\hat{\beta}_n - \beta) \xrightarrow{\mathcal{D}} \text{Cauchy}(0, \beta^2 - 1).$$

1.2. Nearly unstable model

These results led to the study of the following so-called nearly nonstationary (better to call it nearly unstable) AR(1) model

$$\begin{aligned} X_{n,k} &= \beta_n X_{n,k-1} + \epsilon_{n,k}, \quad k = 1, 2, \ldots, n, \\ X_{n,0} &= 0, \end{aligned}$$

where $\beta_n = 1 + \gamma/n$. It can be shown that

$$\left(\sum_{k=1}^{n} X_{n,k-1}^2 \right)^{1/2} (\hat{\beta}_n - \beta_n) \frac{\int_0^1 Y(t)dW(t)}{\left(\int_0^1 Y^2(t)dt \right)^{1/2}},$$

where $Y(t)$, $t \in [0, 1]$ is an Ornstein-Uhlenbeck process defined as the solution of the stochastic differential equation

$$dY(t) = \gamma Y(t)dt + dW(t), \quad Y(0) = 0.$$

By another normalization

$$n(\hat{\beta}_n - \beta_n) \xrightarrow{\mathcal{D}} \frac{\int_0^1 Y(t)dW(t)}{\int_0^1 Y^2(t)dt}.$$

Meer, Pap and Zuijlen [13] considered the following nearly unstable AR(p) model

$$\begin{aligned} X_{n,k} &= \beta_{1,n} X_{n,k-1} + \ldots + \beta_{p,n} X_{n,k-p} + \epsilon_{n,k}, \\ &\qquad k = 1, 2, \ldots, n, \\ X_{n,0} &= X_{n,-1} = \ldots = X_{n,-p} = 0, \end{aligned} \tag{3}$$

where the vector of parameters

$$\mathcal{B}_n = (\beta_{1,n}, \ldots, \beta_{p,n})'$$

is given by

$$\mathcal{B}_n = \mathcal{B} + \sigma_n \mathbf{h_n},$$

where

$$\mathcal{B} = (\beta_1, \ldots, \beta_p)'$$

is a vector such that the polynomial

$$\phi(z) = 1 - \beta_1 z - \ldots - \beta_p z^p$$

corresponds to an unstable AR(p) model, $\{\sigma_n\}$ are the normalizing matrices, and

$$\mathbf{h_n} = (h_{1,n}, \ldots, h_{n,n})'$$

is a sequence of vectors with $\mathbf{h_n} \to \mathbf{h}$. (Jeganathan [10] proved that the sequence $\sigma_n^{-1}(\hat{\mathcal{B}}_n - \mathcal{B}_n)$ converges in law and gave a complicated representation for the limiting distribution in terms of multiple stochastic integrals with respect to the Wiener processes.)

For the sake of simplicity they supposed that ϕ has all its roots on the unit circle. Then ϕ can be written as

$$\phi(z) = (1-z)^a (1+z)^b \prod_{j=1}^{l} ((1 - e^{i\alpha_j} z)(1 - e^{-i\alpha_j} z))^{m_j},$$

where $a, b, l, m_j, j = 1, \ldots, l$ are nonnegative integers, $\alpha_j \in (0, \pi)$, $j = 1 \ldots l$. They suggested writing ϕ in the form

$$\phi(z) = \prod_{j=1}^{q} (1 - a_j z)^{r_j},$$

where $q = 2 + 2l$, $a_j = e^{i\theta_j}$ and $\theta_1, \ldots, \theta_q \in (-\pi, \pi]$ are all different. They supposed that in the nearly unstable AR(p) model the characteristic polynomial ϕ_n can be written as

$$\phi_n(z) = \prod_{j=1}^{q} \prod_{k=1}^{r_j} (1 - a_{j,k,n} z),$$

where $a_{j,k,n} = e^{h_{j,k,n}/n + i\theta_j}$, $h_{j,k,n}$, $j = 1, \ldots, r_j$, $n \geq 1$ are complex numbers such that $h_{j,k,n} \to h_{j,k}$ as $n \to \infty$.

It is clear that

$$\phi_n(z) \to \phi(z) = \prod_{j=1}^{q} (1 - a_j z)^{r_j},$$

where $a_j = e^{i\theta_j}$, $j = 1, \ldots, q$. Obviously $p = \sum\limits_{j=1}^{q} r_j$.

They supposed that the system $\epsilon_{n,k}$ in (3), $k = 1, \ldots, n$, $n \geq 1$ is a triangular array of real squares integrable martingale differences with respect to the filtrations $(\mathcal{F}_{nk})_{k=0,1,\ldots,n;n\geq 1}$ such that for all $t \in [0,1]$

$$\frac{1}{nt}\sum_{k=1}^{[nt]} \mathbb{E}(\epsilon_{nk}^2 | \mathcal{F}_{n,k-1}) \xrightarrow{\mathbb{P}} 1,$$

(4) and

$$\forall \alpha > 0 \qquad \frac{1}{n}\sum_{k=1}^{[nt]} \mathbb{E}(\epsilon_{nk}^2 \chi_{\{|\epsilon_{nk}| > \alpha\sqrt{n}\}} | \mathcal{F}_{n,k-1}) \xrightarrow{\mathbb{P}} 0$$

as $n \to \infty$.

As the main result they described the asymptotic behaviour of the least-squares estimator of the coefficients. A convergence result was presented for the general complex-valued case. The limit distribution was given by the help of some continuous time AR processes.

They clarified the relationship between general complex-valued discrete and continuous time AR(p) models. As a consequence they were able to understand and to simplify the complicated expressions of Jeganathan [10] for the limit distribution of the LSE's in real-valued discrete settings. One of the advantages of their approach of studying complex-valued models is that they could avoid complicated formulas with sines and cosines. They showed how to use their results for real-valued AR(p) models.

In that case the limit distribution can be identified with the maximum likelihood estimator of the coefficients of the corresponding continuous time AR processes.

1.3. Examples

Now for illustration of the above mentioned results we give some examples. We shall study real-valued AR(2) models near to unstable model given by

$$\begin{aligned} X_{n,k} &= \beta_{1,n} X_{n,k-1} + \beta_{2,n} X_{n,k-2} + \epsilon_{n,k}, \quad k = 1, 2, \ldots, n, \\ X_{n,0} &= X_{n,-1} = 0, \end{aligned} \tag{5}$$

where $\{\epsilon_{n,k}\}$ is an array of real random variables satisfying the condition (4) and $\beta_{1,n}, \beta_{2,n}$ are real numbers.

Example 1. First consider the case when the limit unstable model has complex roots, i.e. its characteristic polynomial is

$$\phi(z) = (1 - e^{i\theta} z)(1 - e^{-i\theta} z) = 1 - 2z\cos\theta + z^2.$$

Then we have $\beta_1 = 2\cos\theta$ and $\beta_2 = -1$. The characteristic polynomial of (5)

$$\phi(z) = (1 - e^{h_n/n + i\theta} z)(1 - e^{\overline{h}_n/n - i\theta} z),$$

where $h_n \in \mathbb{C}$ such that $h_n \to h$, as $n \to \infty$ and $\theta \in (0, \pi)$.

As a consequence of the main theorems in [13] we conclude

$$n(\hat{\mathcal{B}}_n - \mathcal{B}) = \begin{pmatrix} n(\hat{\beta}_{1,n} - \beta_1) \\ n(\hat{\beta}_{2,n} - \beta_2) \end{pmatrix} \xrightarrow{\mathcal{D}} \begin{pmatrix} 2(\Re(\hat{c})\cos\theta - \Im(\hat{c})\sin\theta) \\ -2\Re(\hat{c}) \end{pmatrix},$$

where

$$\hat{c} = \frac{\int_0^1 \overline{Y(t)} dY(t)}{\int_0^1 |Y(t)|^2 dt},$$

and $Y(t)$, $t \in [0,1]$ is the continuous time complex-valued AR(1) process given by

$$dY(t) = hY(t)dt + dW(t), \quad Y(0) = 0,$$

where $W(t)$, $t \in [0,1]$ is a standard complex-values Wiener process.

The preceding convergence statement can be reformulated as

$$n(\hat{\mathcal{B}}_n - \mathcal{B}_n) = \begin{pmatrix} n(\hat{\beta}_{1,n} - \beta_{1,n}) \\ n(\hat{\beta}_{2,n} - \beta_{2,n}) \end{pmatrix} \xrightarrow{\mathcal{D}} \frac{2}{s_Y^2} \begin{pmatrix} r_{YW}^{+} \cos\theta - r_{YW}^{-} \sin\theta \\ -r_{YW}^{+} \end{pmatrix},$$

where

$$s_Y^2 = \int_0^1 (Y_1^2(t) + Y_2^2(t))dt,$$

$$r_{YW}^{+} = \int_0^1 (Y_1(t)dW_1(t) + Y_2(t)dW_2(t)),$$

$$r_{YW}^{-} = \int_0^1 (Y_1(t)dW_2(t) - Y_2(t)dW_1(t)),$$

$W_1(t)$ and $W_2(t)$, $t \in [0,1]$ are independent real-valued standard Wiener processes, and the process $(Y_1(t), Y_2(t))$, $t \in [0,1]$ is given by

$$\begin{pmatrix} dY_1(t) \\ dY_2(t) \end{pmatrix} = \begin{pmatrix} \lambda & -\omega \\ \omega & \lambda \end{pmatrix} \begin{pmatrix} Y_1(t)dt \\ Y_2(t)dt \end{pmatrix} + \begin{pmatrix} dW_1(t) \\ dW_2(t) \end{pmatrix}$$

with initial values $Y_1(0) = Y_2(0) = 0$ where $\lambda = \Re(h)$ and $\omega = \Im(h)$.

We remark that Chan and Wei [3] proved convergence of $n(\hat\beta_{2,n} + 1)$ in the stable case, i.e. when $h_n \equiv 0$.

Example 2. Now consider the case when the limit unstable model has double roots equal to 1, i.e. its characteristic polynomial is

$$\phi(z) = (1-z)^2 = 1 - 2z + z^2,$$

and we have $\beta_1 = 2$ and $\beta_2 = -1$. The characteristic polynomial of the model (5) has the form

$$\phi_n(z) = (1 - e^{h_{1,n}/n} z)(1 - e^{h_{2,n}/n} z),$$

where $h_{k,n} \in \mathbb{C}$ such that $h_{k,n} \to h_k$, as $n \to \infty$, for $k = 1, 2$, and the polynomial ϕ_n has real coefficients. This implies that $h_{1,n}$ and $h_{2,n}$ are real numbers or conjugated complex numbers. The same is valid for h_1 and h_2.

One can get

$$\begin{pmatrix} 0 & -n \\ n^2 & n^2 \end{pmatrix} \begin{pmatrix} \hat\beta_{1,n} - \beta_1 \\ \hat\beta_{2,n} - \beta_2 \end{pmatrix} = \begin{pmatrix} \hat c_{1,n} \\ \hat c_{2,n} \end{pmatrix} \xrightarrow{\mathcal{D}} \begin{pmatrix} \hat c_1 \\ \hat c_2 \end{pmatrix},$$

where

$$\begin{pmatrix} \hat c_1 \\ \hat c_2 \end{pmatrix} = S^{-1} \begin{pmatrix} \int_0^1 \dot Y(t) d\dot Y(t) \\ \int_0^1 Y(t) d\dot Y(t) \end{pmatrix},$$

$$S = \begin{pmatrix} \int_0^1 (\dot Y(t))^2 dt & \int_0^1 \dot Y(t) Y(t) dt \\ \int_0^1 Y(t) \dot Y(t) dt & \int_0^1 (Y(t))^2 dt \end{pmatrix},$$

and $Y(t)$, $t \in [0,1]$ is the continuous time real-valued AR(2) process

$$\begin{aligned} d\dot Y(t) &= ((h_1 + h_2)\dot Y(t) - h_1 h_2 Y(t))dt + dW(t), \\ dY(t) &= \dot Y(t) dt, \\ Y(0) &= \dot Y(0) = 0, \end{aligned}$$

where $W(t)$, $t \in [0,1]$ is a standard real-valued Wiener process. Moreover $\hat{c}_1, \hat{c}_2$ can be interpreted as the MLE of $c_1 = h_1 + h_2$ and $c_2 = -h_1 h_2$. By Ito's formula we can also derive

$$\begin{pmatrix} 0 & -n \\ n^2 & n^2 \end{pmatrix} \begin{pmatrix} \hat{\beta}_{1,n} - \beta_{1,n} \\ \hat{\beta}_{2,n} - \beta_{2,n} \end{pmatrix} \xrightarrow{\mathcal{D}} S^{-1} \begin{pmatrix} \int_0^1 \dot{Y}(t) dW(t) \\ \int_0^1 Y(t) dW(t) \end{pmatrix}.$$

Comparing the complex-valued AR(2) models with real-valued AR(2) models we observe that convergence of LSE's in the real-valued models can be derived from the complex-valued case by taking into account the extra requirement that the coefficients should be real numbers. However, the formulations in the context of complex-valued models are remarkably simpler.

As we have seen, a multiple root in the model implies a higher order autoregressive component in the corresponding continuous time model. Different but not conjugated roots imply components driven by independent Wiener processes in the continuous time model. In case the roots are conjugated pairs, then the components are driven by conjugated complex-valued Wiener processes. A real root is connected to a real-valued Wiener process, and a complex root is connected to a complex-valued Wiener process, even if the model has real coefficients!

We finally note, that convergence of LSE's in models with complex-valued disturbances $\{\epsilon_{n,k}\}$ can be handled similarly (see the AR(1) case in Kormos, van der Meer, Pap and van Zuijlen [11]).

2. The problem, preliminary

Let us consider the p-order autoregressive process

$$X_k + \beta_1 X_{k-1} + \ldots + \beta_p X_{k-p} = \epsilon_k \tag{1'}$$

and instead of (2) characteristic polynomial we deal with $\Phi(Z) = Z^p + \beta_1 Z^{p-1} + + \ldots + \beta_p$. Let us suppose that one of the zeros of the characteristic polynomial is on the unit circle, and the absolute value of the others is less than one, i.e.

$$1 \geq |Z_1| > |Z_2| \geq \ldots \geq |Z_p|. \tag{6}$$

The sharp inequality between $|Z_1|$ and $|Z_2|$ means that Z_1 is a real zero (the complex conjugate does not appear). Using the B backshift operator we can write (1') in the following form of

$$A_p(B)X_k = \epsilon_k,$$

where

$$A_p(B) = (1 - Z_1B)(1 - Z_2B)\dots(1 - Z_pB);$$

where p is the degree of the operator-polynomial.

From this it follows that

$$(1 - Z_1B)X_k = (1 - Z_pB)^{-1}\dots(1 - Z_2B)^{-1}\epsilon_k.$$

Denoting the expression on the right hand side of the equation by Y_k we can get

$$A_{p-1}(B)Y_k = \epsilon_k,$$

which is a stochastic difference equation defining a $(p-1)$-order autoregressive process, where (of course) $A_{p-1}(B) = (1 - Z_2B)\dots(1 - Z_pB)$. It results that the last two equations can be written with coefficients in the form of

$$X_k - \rho X_{k-1} = Y_k, \tag{7}$$

$$Y_k + c_1Y_{k-1} + \dots + c_{p-1}Y_{k-p+1} = \epsilon_k, \tag{8}$$

which is an equivalent form of the equation (1') taking into consideration that

$$\begin{aligned}\beta_1 &= -\rho + c_1,\\ \beta_i &= -c_{i-1}\rho + c_i, \quad i = 2,\dots,(p-1),\\ \beta_p &= -c_{p-1}\rho.\end{aligned} \tag{9}$$

We can interpret the above also in the following way: The examination of the stationarity properties of the (1') AR(p) process is traced back to the behaviour of such AR(1) process, where the innovation process Y_k is an AR(p--1) stationary process. (We have implicitly used the fact that the stationarity of an autoregressive process could be determined by the zeros of its characteristic polynomial with maximum absolute values.)

A stationary (in strong sense) process (X_t) satisfies the strong mixing condition, if

$$\sup_{A,B} |P(AB) - P(A)P(B)| = \alpha(k) \to 0,$$

when $k \to \infty$, where
$$A \in F^X_{(-\infty,0]}, \quad B \in F^X_{(k,\infty]}.$$

Regarding the equations (7)-(8), let us examine now, when is a stationary finite-order autoregressive model strong mixing. The complexitiy of the problem is underlined by the fact that C.S. Withers [17] created an example for stationary first-order processes, that are not strong mixing.

Let $Y = (Y_k)$ be q-order stationary autoregressive process, with zero mean, that is let the following equation be fulfilled

$$Y_k + c_1 Y_{k-1} + \ldots + c_q Y_{k-q} = \epsilon_k. \tag{10}$$

The stationarity condition results that for the zeros of the characteristic polynomial it is true that

$$|Z_i| < 1, \qquad i = 1, \ldots, q,$$

and the moving average representation of the process (Y_k)

$$Y_k = \sum_{i=0}^{\infty} g_i \epsilon_{k-i} \tag{11}$$

does exist.

Let

$$\underline{b}' = (b_l, \ldots, b_{l+m-1}); \qquad \underline{\gamma}' = (\gamma_l, \ldots, \gamma_{l+m-1});$$
$$\underline{\alpha}_n' = (\alpha_{nl}, \ldots, \alpha_{n(l+m-1)}); \qquad \underline{\beta}_n' = (\beta_{nl}, \ldots, \beta_{n(l+m-1)})$$

be m-dimensional vectors with scalar elements.

Let V_t from the (11) be defined by

$$V_t = \sum_{i=0}^{t-1} g_i \epsilon_{t-i}, \qquad t = l, \ldots, l+m-1,$$

and compose the vector $\underline{V}' = (V_l, \ldots, V_{l+m-1})$. Let us define the open interval D_n in the form of
$$D_n = \{\underline{b} \mid \alpha_{nt} < b_t < \beta_{nt}\}$$

and let
$$D = \bigcup_{n=1}^{s} D_n.$$

Let us consider the following conditions for (ϵ_l), $(l \in \mathbb{Z})$:
– the random variables ϵ_l are independent,
– for some $\delta > 0$

$$\max_l E|\epsilon_l|^\delta < \infty,$$

– if Φ_l denotes the characteristic function of ϵ_k then

(12)
$$\max_l \int_{-\infty}^{\infty} |\Phi_l(t)| dt < \infty,$$

$$\sup_{m,s,l} \sup_{\underline{\alpha},\underline{\beta},\underline{\gamma}} \max_t \left| \frac{\partial}{\partial \gamma_t} P(\underline{V} + \underline{\gamma} \in D) \right| < \infty,$$

where $m \in \mathbb{N}$, $s \in \mathbb{N}$, $l \in \mathbb{Z}$.

After all this the following lemma can be formulated.

Lemma. *Let the process $Y = (Y_k)$ be the (10) q-order stationary autoregressive process. Let the conditions (12) be satisfied for the innovation process ϵ_k. Then the process (Y_k) is strong mixing and for the coefficients $\alpha(l)$*

$$\alpha(l) = \mathcal{O}(Z_0^{\lambda l}),$$

where

$$\lambda = \frac{\delta}{1+\delta}$$

and

$$\max_j |Z_j| < Z_0 < 1.$$

The detailed proof of the lemma can be found in the papers of K.C. Chanda [4], V.V. Gorodetskij [6] and C.S. Withers [17]. These authors found the solution of the general but not trivial problem, when a linear process shows strong mixing characteristics. The main topic of the present paper and the limited space do not allow us to treat these questions in due depth.

Corollary. *It follows from the lemma that*

$$\sum_{l=1}^{\infty} \alpha(l)^{1-2/\beta} < \infty,$$

where $\beta > 2$.

Note. Concerning the behaviour of coefficients $\alpha(m)$ the following relations are known. Denote the maximal correlation between the past and the future of (Y_k) by $r^*(m)$, i.e. let

$$r^*(m) = \sup_{\xi,\eta} E\xi\eta,$$

where ξ is any $F^y_{(-\infty,0]}$ measureable random variable, and η is a $F^y_{[m,\infty)}$ measureable random variable. The condition $r^*(m) \to 0$ $(m \to \infty)$ is the condition of the asymptotic independence between the past and the future of the process.

The followings are true for the $r^*(m)$ sequence

$$r^*(m) \geq 4\alpha(m), \qquad m = 1, 2, \ldots$$

If (Y_k) is a Gaussian process, then $r^*(m)$ and $\alpha(m)$ are asymptotically equivalent, more precisely

$$4\alpha(m) \leq r^*(m) \leq 2\pi\alpha(m)$$

(see I.A. Ibragimov - Yu.V. Linnik [9] and P. Hall - C.C. Heyde [7]).

3. Results, proofs

Then we can form our statement concerning AR(p) processes from which a real zero could be separated.

Theorem. *Let $X = (X_k)$ be a p-order autoregressive process defined by the stochastic difference equation (1') and let $X_k = 0$, $k \leq 0$. The innovation process (ϵ_k) satisfies the conditions (12), and the zeros of the characteristic polynomial satisfy the condition (6). Let us perform the transformations (7)-(9) and denote by*

$$\hat{\rho} = \frac{\sum\limits_{k=1}^{n} X_k X_{k-1}}{\sum\limits_{k=1}^{n} X_{k-1}^2}$$

the least squares estimator of the unknown ρ parameter in the first-order autoregressive process according to (7)

$$X_k = \rho X_{k-1} + Y_k$$

and let

$$\sup_k E|Y_k|^\beta < \infty$$

be fulfilled for some $\beta > 2$.

Then, if $\rho = 1$,

$$n(\hat{\rho} - 1) \underset{n\to\infty}{\overset{\mathcal{D}}{\to}} \frac{1}{2}\left(W^2(1) - \frac{\sigma_y^2}{\sigma^2}\right) \Big/ \int_0^1 W^2(s)ds, \tag{13}$$

and if $\rho = \rho_n = e^{-\lambda/n}$, *then*

$$n(\hat{\rho} - \rho) \underset{n\to\infty}{\overset{\mathcal{D}}{\to}} \frac{\int_0^1 X(s)dW(s) + \frac{1}{2}\left(1 - \frac{\sigma_y^2}{\sigma^2}\right)}{\int_0^1 X^2(s)ds}, \tag{14}$$

where

$$\sigma_y^2 = \lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^n E(Y_k^2), \tag{15}$$

$$\sigma^2 = \lim_{n\to\infty} E\left(\frac{1}{2}\left(\sum_{k=1}^n Y_k\right)^2\right),$$

and the process $(X(s))$ *is defined by the following equation*

$$X(s) = W(s) - \lambda\int_0^s e^{-(s-r)\lambda}W(r)dr, \tag{16}$$

where $W(s)$ *is the standard Wiener process.*

Proof. The technique of the proof is standard, it is not different from the methods used in similar problems (Donsker-Prohorov invariance principle, application of the continuous mapping theorem), we will only show the most important steps.

Notice that

$$n(\hat{\rho} - \rho) = \frac{\frac{1}{n}\sum_{k=1}^n X_{k-1}Y_k}{\frac{1}{n^2}\sum_{k=1}^n X_{k-1}^2}, \tag{17}$$

and for $\rho = 1$

$$n(\hat{\rho} - 1) = \frac{\frac{1}{n} \sum\limits_{k=1}^{n} X_{k-1}(X_k - X_{k-1})}{\frac{1}{n^2} \sum\limits_{k=1}^{n} X_{k-1}^2}.$$

Denote by $X_k(t)$ the random element in $D[0,1]$

$$X_n(t) = \frac{1}{\sqrt{n}\sigma} \sum_{k=1}^{[nt]} Y_k = \frac{1}{\sqrt{n}\sigma} \sum_{k=1}^{k-1} Y_k,$$

$$\frac{k-1}{n} \leq t < \frac{k}{n} \qquad (k = 1, \ldots, n),$$

$$X_n(1) = \frac{1}{\sqrt{n}\sigma} \sum_{k=1}^{n} Y_k.$$

One can show that, if $n \to \infty$

$$X_n(t) \xrightarrow{\mathcal{D}} W(t), \tag{18}$$

where $W(t)$ is the standard Wiener process, see N. Herrndorf [8].

After simple transformations it can be seen that

$$\frac{1}{n^2} \sum_{k=1}^{n} X_{k-1}^2 = \sigma^2 \int_0^1 X_n^2(t) dt,$$

and

$$\frac{1}{n} \sum_{k=1}^{n} X_{k-1}(X_k - X_{k-1}) = \frac{\sigma^2}{2} X_n^2(1) - \frac{1}{2n} \sum_{k=1}^{n} Y_k^2,$$

thus applying the continuous mapping theorem for (18) and taking into consideration that

$$\frac{1}{n} \sum_{k=1}^{n} Y_k^2 \to \sigma_y^2, \quad \text{with probability 1}, \tag{19}$$

if $n \to \infty$ we get the statement (13).

In the case of $\rho = e^{-\frac{\lambda}{n}}$

$$X_k = \sum_{j=1}^{k} e^{-\frac{(k-j)\lambda}{n}} Y_j,$$

hence

$$\frac{1}{\sqrt{n}} X_{[nt]} = \sigma \sum_{j=1}^{[nt]} e^{-\frac{([nt]-j)\lambda}{n}} \int\limits_{(j-1)/n}^{j/n} dX_n(s) = \sigma \int\limits_0^t e^{-(t-s)\lambda} dX_n(s).$$

By partial integration we get the expression

$$\sigma \left(X_k(t) - \lambda \int\limits_0^t e^{-(t-s)\lambda} X_n(s) ds \right),$$

for which in case of $n \to \infty$ the following statement is true

$$\sigma \left(X_n(t) - \lambda \int\limits_0^t e^{-(t-s)\lambda} X_n(s) ds \right) \xrightarrow{\mathcal{D}} \sigma X(t),$$

where $X(t)$ is defined by (16).

From these it follows directly by the re-application of the continuous mapping theorem

$$\frac{1}{n^2} \sum_{k=1}^{n} X_{k-1}^2 \xrightarrow{\mathcal{D}} \sigma^2 \int\limits_0^1 X^2(s) ds, \tag{20}$$

when $n \to \infty$.

Now only the limit distribution specification of the numerator $n(\hat{\rho} - \rho)$ in the expression (17) is to be done. Let us consider the following identity

$$\frac{1}{n} X_n^2 = -\frac{2\lambda}{n^2} \sum_{k=1}^{n} X_{k-1}^2 + \frac{1}{n} \sum_{k=1}^{n} Y_k^2 + \frac{2}{n} \sum_{k=1}^{n} X_{k-1} Y_k.$$

From this applying (19), (20) and the continuous mapping theorem we get

$$\frac{2}{n} \sum_{k=1}^{n} X_{k-1} Y_k \xrightarrow{\mathcal{D}} \sigma^2 X^2(1) + 2\lambda\sigma^2 \int\limits_0^1 X^2(s) ds - \sigma_y^2, \tag{21}$$

when $n \to \infty$. (21) and (20) result the statement (14) of the theorem directly.

Corollary. (i) Because (Y_k) is a stationary sequence, the condition (15) needs the satisfaction of $EY_1^2 < \infty$, which comes from the conditions trivially. The limit distribution (13) also exists and is positive, and can be written in the form of

$$\sigma^2 = EY_1^2 + 2\sum_{k=2}^{\infty} EY_1 Y_k.$$

(ii) If the innovation process (Y_k) is a sequence of independent and identically distributed random variables the (13) gives back the classical result of J.S. White [15] and T.W. Anderson [1] exactly.

(iii) The process (16) satisfies the following first-order stochastic differential equation

$$dX(s) + \lambda X(s)ds = dW(s).$$

We supposed that we had some preliminary knowledge about the behaviour of the process (1'), i.e. about the zeros of the process and we supposed that they could be characterized by the condition (6). So whether the coefficients $\beta_2, \ldots, \beta_p (c_1, \ldots, c_{p-1})$ are known or unknown, the behaviour of the process is determined by the value of the parameter ρ, or more precisely by the distance of its absolute value from 1.

References

[1] **Anderson T.W.,** *The statistical analysis of time series,* Wiley, New York, 1971.

[2] **Anderson T.W.,** On asymptotic distributions of estimates of parameters of stochastic difference equations, *Ann. Math. Statist.,* **30** (1959), 676-687.

[3] **Chan N.H. and Wei C.Z.,** Limiting distributions of least squares estimates of unstable autoregressive processes, *Ann. Statist.,* **16** (1988), 367-401.

[4] **Chanda K.C.,** Strong mixing properties of linear stochastic processes, *J. Appl. Prob.,* **11** (1974), 401-408.

[5] **Dzhaparidze K., Kormos J., Meer T.v.d. and Zuijlen M.v.,** Parameter estimation for nearly nonstationary AR(1) processes, *Mathematical and Computer Modelling,* **19** (1992), 29-41.

[6] **Городецкий В.В.** О свойстве сильного перемещивания для линейно порожденных последовательностей, *Теория вер. и ее прим.,* **22** (1977), 411-413.

[7] **Hall P. and Heyde C.C.,** Martingale limit theory and its application, Academic Press, New York, 1980.

[8] **Herrndorf N.,** A functional central limit theorem for weakly dependent sequences of random variables, *Ann. Prob.,* **12** (1984), 141-153.

[9] **Ибрагимов И.А. и Линник Ю.В.,** *Независимые и стационарно связанные величины,* Наука, Москва, 1965.

[10] **Jeganathan P.,** On the asymptotic behaviour of least-squares estimators in AR time series with roots near the unit circle, *Econometric Theory,* **7** (1991), 269-306.

[11] **Kormos J., Meer T. v.d., Pap G. and Zuijlen M.v.,** Asymptotic inference of nearly nonstationary complex-valued AR(1) processes, *Exploring stochastic laws. Festschrift in honour of the 70th birthday of academician V.S. Korolyuk,* VSP, The Netherlands, 1995, 201-218.

[12] **Mann H.B. and Wald A.,** On the statistical treatment of linear stochastic difference equations, *Econometrica,* **11** (1943), 173-220.

[13] **Meer T.v.d., Pap G. and Zuijlen M.v.,** Asymptotic inference for nearly unstable AR(p) processes, *Econometric Theory,* **15** (1999), 184-217.

[14] **Philips P.C.B.,** Time series regression with unit roots, *Econometrica,* **55** (1987), 277-301.

[15] **White J.S.,** The limit distribution of the serial correlation coefficient in the explosive case I., *Ann. Math. Statist.,* **29** (1958), 1188-1197.

[16] **White J.S.,** The limiting distribution of the serial correlation coefficient in the explosive case II., *Ann. Math. Statist.,* **30** (1959), 831-834.

[17] **Withers C.S.,** Conditions for linear processes to be strong mixing, *Z. Wahr. verw. Geb.,* **57** (1981), 477-480.

J. Kormos
Institute of Informatics
University of Debrecen
H-4010 Debrecen, Pf. 12
Hungary