Теорія Ймовір. та Матем. Статист.
Вип. 72, 2005

Theor. Probability and Math. Statist.
No. 72, 2006, Pages 177–191
S 0094-9000(06)00675-2
Article electronically published on September 6, 2006

ESTIMATION OF A MATRIX-VALUED PARAMETER OF AN AUTOREGRESSIVE PROCESS WITH NONSTATIONARY NOISE

UDC 519.21

A. P. YURACHKIVSKIĬ AND D. O. IVANENKO

Abstract. Suppose that $\check{A}_n$ is the least squares estimator constructed from n observations of an unknown matrix A in an autoregressive process $\xi_k = A\xi_{k-1} + \varepsilon_k$. Under the assumption that the sequence (ε_k) is a martingale difference, not necessarily stationary and ergodic, we find the limit distribution as $n \to \infty$ of the statistic $\sqrt{n}(\check{A}_n - A)$ by using methods of stochastic analysis. This limit distribution may be different from the normal distribution.

1. Introduction

Consider the following autoregressive process:

$$\xi_k = A\xi_{k-1} + \varepsilon_k, \qquad k \in \mathbb{N}, \tag{1}$$

where A is a square matrix, and (ε_k) is a martingale difference with respect to a filtration of σ-algebras $(\mathcal{F}_k, k \in \mathbb{Z}_+)$ such that the random variable ξ_0 is $\mathcal{F}_0$-measurable. Unless otherwise stated, all vectors in the sequel are interpreted as column vectors. Then $a^\top b$ is the scalar product, while $ab^\top$ is the tensor product of vectors. The latter is also denoted by $a \otimes b$; in particular, $a^{\otimes 2} = aa^\top$. The vector norm, denoted in the same way as the module, is supposed to be Euclidean, while the matrix norm is assumed to be the operator norm. The pseudoinverse matrix of the matrix B is denoted by $B^\dagger$, and O stands for the zero matrix. For symmetric matrices, the inequality $S \le T$ means that $T - S$ is positive definite. The limit in probability is denoted by l. i. p.

We study the asymptotic behavior, as $n \to \infty$, of the normalized deviation

$$\sqrt{n}\left(\check{A}_n - A\right)$$

of the true value of the matrix-valued parameter A from its least squares estimator

$$\check{A}_n = \left(\sum_{i=1}^n \xi_i \xi_{i-1}^\top\right)\left(\sum_{i=1}^n \xi_{i-1}\xi_{i-1}^\top\right)^\dagger. \tag{2}$$

Our approach is based on an application of some methods of stochastic analysis and enables us to avoid the assumption of ergodicity and even of asymptotic normality of the sequence (ε_k). Therefore the limit distribution of the underlying statistic may be different from the normal distribution.

2000 *Mathematics Subject Classification.* Primary 62F12; Secondary 60F05.

2. Preliminary results

Lemma 1. *Let D and F be arbitrary square matrices of the same order. Then*

$$DFF^\top D^\top \le \|F\|^2 DD^\top.$$

Proof. For any row vector z, we have

$$zDFF^\top D^\top z^\top = |zDF|^2 \le |zD|^2\|F\|^2 = \|F\|^2 zDD^\top z^\top. \qquad \square$$

Let $\mathsf{E}^0 = \mathsf{E}[\dots|\,\mathcal{F}_0]$.

By iterating equality (1), we obtain

$$\xi_k = \sum_{i=0}^{k-1} A^i \varepsilon_{k-i} + A^k \xi_0, \tag{3}$$

giving immediately the following result.

Lemma 2. *Assume that for each positive integer i,*

$$\mathsf{E}^0\,|\varepsilon_i| < \infty \quad \text{a.s.}, \tag{4}$$

$$\mathsf{E}[\varepsilon_i|\,\mathcal{F}_{i-1}] = 0. \tag{5}$$

Then

$$\mathsf{E}^0\,\xi_k^{\otimes 2} = \sum_{i=0}^{k-1} A^i\,\mathsf{E}^0\,\varepsilon_{k-i}^{\otimes 2} A^{\top i} + (A^k\xi_0)^{\otimes 2}$$

for all positive integers k.

Put $\sigma_k^2 = \mathsf{E}[\varepsilon_k^{\otimes 2}|\,\mathcal{F}_{k-1}]$, $\chi_k^N = I\{|\xi_k| > N\}$, $B_k^N = \mathsf{E}^0\,\xi_k^{\otimes 2}\chi_k^N$, and

$$I_k^N = I\{|\varepsilon_k| > (1-\|A\|)N\}.$$

Lemma 3. *Assume that conditions* (4) *and* (5) *hold and suppose that $\|A\| < 1$. Then*

$$B_k^N \le AB_{k-1}^N A^\top + 2((1-\|A\|)N)^{-2} A\,\mathsf{E}^0\left(\xi_{k-1}^{\otimes 2}\operatorname{tr}\sigma_k^2\right)A^\top + \mathsf{E}^0\,\varepsilon_k^{\otimes 2}\chi_{k-1}^N + 2\,\mathsf{E}^0\,\varepsilon_k^{\otimes 2} I_k^N$$

for any positive integer k.

Proof. By (1), we can write

$$\xi_k^{\otimes 2} = A\xi_{k-1}^{\otimes 2}A^\top + A\xi_{k-1}\varepsilon_k^\top + \varepsilon_k(A\xi_{k-1})^\top + \varepsilon_k^{\otimes 2}, \tag{6}$$

$$\chi_k^N \le \chi_{k-1}^N + I_k^N.$$

Then we multiply both sides of the inequality by a positive definite matrix $\xi_k^{\otimes 2}$ and transform the right-hand side according to (6). Therefore, by averaging both sides with respect to $\mathcal{F}_0$ and by condition (5), we obtain

$$\begin{aligned} B_k^N \le{}& AB_{k-1}^N A^\top + A\,\mathsf{E}^0\,\xi_{k-1}^{\otimes 2} I_k^N A^\top + \mathsf{E}^0\left(A\xi_{k-1}\varepsilon_k^\top + \varepsilon_k(A\xi_{k-1})^\top\right) I_k^N \\ &+ \mathsf{E}^0\,\varepsilon_k^{\otimes 2}\chi_{k-1}^N + \mathsf{E}^0\,\varepsilon_k^{\otimes 2} I_k^N. \end{aligned}$$

It remains to apply the following immediate (in)equalities:

$$ab^\top + ba^\top \le a^{\otimes 2} + b^{\otimes 2}, \qquad \mathsf{E}^0\,\xi_{k-1}^{\otimes 2} I_k^N = \mathsf{E}^0\left(\xi_{k-1}^{\otimes 2}\mathsf{P}\{|\varepsilon_k| > (1-\|A\|)N|\,\mathcal{F}_{k-1}\}\right),$$

$$\mathsf{E}\left[|\varepsilon_k|^2|\,\mathcal{F}_{k-1}\right] = \operatorname{tr}\sigma_k^2. \qquad \square$$

Now we assume that A satisfies a weaker condition: For some positive integer m,

$$\|A^m\| < 1. \tag{7}$$

We also introduce the following notation: $\xi_k^{[j]} = \xi_{km+j}$, $\mathcal{F}_k^{[j]} = \mathcal{F}_{km+j}$,

$$\varepsilon_k^{[j]} = \sum_{i=0}^{m-1} A^i \varepsilon_{km+j-i},$$

$$\sigma_k^{[j]2} = \sum_{i=0}^{m-1} A^i \,\mathsf{E}\left[\varepsilon_{km+j-i}^{\otimes 2} \middle| \mathcal{F}_{k-1}^{[j]}\right] A^{\top i}, \qquad \chi_k^{jN} = I\left\{\left|\xi_k^{[j]}\right| > N\right\},$$

and $B_k^{jN} = \mathsf{E}^0\, \xi_k^{[j]\otimes 2} \chi_k^{jN}$, $I_k^{jN} = I\{\left|\varepsilon_k^{[j]}\right| > (1 - \|A^m\|)N\}$, $k \in \mathbb{Z}_+, j \in \{1, \dots, m\}$. It is clear that

$$\xi_k^{[j]} = A^m \xi_{k-1}^{[j]} + \varepsilon_k^{[j]}$$

for $k \geq 1$. Taking into account that $\mathsf{E}^0 \left|\varepsilon_k^{[j]}\right|^2 < \infty$ and $\mathsf{E}\left[\varepsilon_k^{[j]\otimes 2} \middle| \mathcal{F}_{k-1}^{[j]}\right] = \sigma_k^{[j]2}$ under conditions (4) and (5), one can repeat the proof of Lemma 3 with some evident modifications to obtain the following generalization of this result.

Lemma 4. *If conditions* (3), (4), *and* (7) *hold, then*

$$B_k^{jN} \leq A^m B_{k-1}^{jN} A^{\top m} + 2((1 - \|A^m\|)N)^{-2} A\, \mathsf{E}^0 \left(\xi_{k-1}^{[j]\otimes 2} \operatorname{tr} \sigma_k^{[j]2}\right) A^\top$$
$$+ \mathsf{E}^0\, \varepsilon_k^{[j]\otimes 2} \chi_{k-1}^{jN} + 2\, \mathsf{E}^0\, \varepsilon_k^{[j]\otimes 2} I_k^{jN}$$

for any $k \in \mathbb{N}$ *and* $j \in \{1, \dots, m\}$.

3. Reducing to a simpler problem

By (1), (4), and (5), we have $\mathsf{E}^0(\xi_n - A\xi_{n-1})\xi_{n-1}^\top = O$, implying that

$$A = \mathsf{E}^0\, \xi_n \xi_{n-1}^\top \left(\mathsf{E}^0\, \xi_{n-1} \xi_{n-1}^\top\right)^{-1},$$

since the matrix $\mathsf{E}^0\, \xi_{n-1}^{\otimes 2}$ is nondegenerate. This gives a hint to use estimate (2) even without the least squares method.

Let

$$K_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \varepsilon_i \xi_{i-1}^\top, \qquad Q_n = \frac{1}{n} \sum_{i=0}^{n-1} \xi_i^{\otimes 2}, \qquad L_n = \frac{1}{n} \sum_{i=1}^{n} \varepsilon_i^{\otimes 2},$$

$$R_n = \frac{\xi_{n-1}^{\otimes 2} - \xi_0^{\otimes 2}}{n} + \frac{AK_n + (AK_n)^\top}{\sqrt{n}}.$$

Then it follows from (2) that

$$\sqrt{n}(\check{A}_n - A) = K_n Q_n^\dagger. \tag{8}$$

Substituting the right-hand side of equality (1) instead of ξ_i in the definition of Q_n we obtain after some algebra that

$$Q_n = AQ_nA^\top + L_n + R_n. \tag{9}$$

In order to obtain Q_n from this equality, the following rough result is sufficient.

Lemma 5. *Assume that condition* (7) *holds for some positive integer* m. *Then for any* F, *the matrix equation*

$$X = AXA^\top + F \tag{10}$$

has a unique solution.

Proof. It is clear that condition (7) implies that the series $\sum_{n=0}^{\infty} A^n F A^{\top n}$ obtained by iterating equality (10) is convergent. □

Necessary and sufficient conditions for the solution of this equation to exist and be unique in the class of positive definite matrices are given in [1, p. 92].

Denote by $\mathfrak{A}F$ the solution of equation (10). Then, by Lemma 5, formulas (9) and (7) imply that

$$Q_n = \mathfrak{A}L_n + \mathfrak{A}R_n. \tag{11}$$

Assume that B is a symmetric matrix of rank r having nonzero eigenvalues $\lambda_1, \dots, \lambda_r$ and suppose that $a_1, \dots, a_r$ are the corresponding eigenvectors. Then it is clear that

$$B^\dagger = \sum_{i=1}^{r} \lambda_i^{-1} a_i^{\otimes 2}.$$

This representation implies the following result.

Lemma 6. *Assume that a sequence (B_n) of symmetric matrices converges to a nondegenerate matrix B. Then $B_n^\dagger \to B^{-1}$.*

Theorem 1. *Suppose that condition* (7) *holds for some positive integer m,*

$$|\xi_n|^2/n \xrightarrow{\mathsf{P}} 0, \tag{12}$$

$$(K_n, L_n) \xrightarrow{\mathrm{d}} (K, L), \tag{13}$$

and the matrix $\mathfrak{A}L$ is nondegenerate with probability 1*. Then*

$$\sqrt{n}(\check{A}_n - A) \xrightarrow{\mathrm{d}} K(\mathfrak{A}L)^{-1}.$$

Proof. By (11), (12), and (13) we obtain that

$$(K_n, Q_n) \xrightarrow{\mathrm{d}} (K, \mathfrak{A}L),$$

since the operator $\mathfrak{A}$ is continuous. Based on the latter relation and Lemma 6 we apply the method of a common probability space and obtain that

$$K_n Q_n^\dagger \xrightarrow{\mathrm{d}} K(\mathfrak{A}L)^{-1},$$

since $\mathfrak{A}L$ is nondegenerate. It remains to use equality (8). □

Convergence to zero of the normalized solution of the recurrent equation (1) is studied in [2]. In our case, condition (12) can readily be checked using Lemma 2.

Checking condition (13) is the focus of the next section.

4. Solving an easier problem using stochastic analysis tools

Put

$$Y_n(t) = \frac{1}{\sqrt{n}} \sum_{k=1}^{[nt]} \varepsilon_k \otimes \xi_{k-1}, \qquad Z_n(t) = \frac{1}{\sqrt{n}} \sum_{k=1}^{[nt]} \varepsilon_k \tag{14}$$

(here, square brackets stand for the integer part),

$$G_n = \frac{1}{n} \sum_{k=0}^{n-1} \sigma_{k+1}^2 \otimes \xi_k^{\otimes 2}, \qquad H_n = \frac{1}{n} \sum_{k=0}^{n-1} \sigma_{k+1}^2 \otimes \xi_k, \qquad J_n = \frac{1}{n} \sum_{k=1}^{n} \sigma_k^2,$$

and let * be a linear operation on the space of 4-rank tensors acting on tensor products of vectors according to the rule

$$(a \otimes b \otimes c \otimes d)^* = a \otimes c \otimes b \otimes d.$$

We understand the joint characteristic $\langle M, V \rangle$ of tensor-valued local square integrable martingales M and V as the compensator [3, 4] of the process $M \otimes V$ (the notation $\langle M, M \rangle$

is abbreviated to $\langle M\rangle$), the square variation of the stochastic process X is interpreted as a tensor-valued process $[X]$ whose components are joint square variations [3, 4] of the components of X. In this way, the interpretation of the square brackets depends on the context in which they are used.

It is clear that

$$K_n = Y_n(1), \qquad L_n = [Z_n](1). \tag{15}$$

By construction and by conditions (4) and (5), Y_n and Z_n are locally square integrable martingales and

$$\begin{gathered}\langle Y_n\rangle(t) = n^{-1}[nt]G^*_{[nt]}, \qquad \langle Z_n, Y_n\rangle(t) = n^{-1}[nt]H_{[nt]},\\ \langle Z_n\rangle(t) = n^{-1}[nt]J_{[nt]}.\end{gathered} \tag{16}$$

The joint characteristic $\langle Y_n, Z_n\rangle$ is obtained from $\langle Z_n, Y_n\rangle$ by changing the numeration of the components of the tensor (an operation similar to *). Therefore we can exclude it from our argument.

Introduce the following condition:

CP. The sequences (G_n), (H_n), and (J_n) converge in probability to some random tensors G, H, and J, respectively.

In what follows $C, C_1, \dots$ stand for nonrandom constants. Weak convergence in C (denoted by $\xrightarrow{\mathrm{C}}$) of a sequence of stochastic processes without discontinuities of the second kind is understood as weak convergence of the measures generated by these processes on the Borel σ-algebra of the space D to the measure generated by an underlying continuous process.

Theorem 2. *Assume that conditions* (5) *and* **CP** *hold. Let*

$$\|\sigma_k^2\| \le C, \tag{17}$$

$$\lim_{N\to\infty} \overline{\lim_{n\to\infty}} \frac{1}{n}\sum_{k=1}^{n} \mathsf{E}\,|\varepsilon_k|^2 I\{|\varepsilon_k| > N\} = 0 \tag{18}$$

and suppose that condition (7) *is satisfied for some positive integer m. Then the sequence of vector stochastic processes (Y_n, Z_n) converges weakly in* C *to a continuous martingale (Y, Z) starting from zero and such that $\langle Y\rangle(t) = G^*t$, $\langle Z, Y\rangle(t) = Ht$, and $\langle Z\rangle(t) = Jt$.*

Proof. By (16) and by condition **CP**, we have

$$\langle Y_n\rangle(t) \xrightarrow{\mathsf{P}} G^*t, \qquad \langle Z_n, Y_n\rangle(t) \xrightarrow{\mathsf{P}} Ht, \qquad \langle Z_n\rangle(t) \xrightarrow{\mathsf{P}} Jt$$

for any t. If the terms in (14) satisfy the Lindeberg condition with respect to $\mathcal{F}_0$, then Theorem 2 follows from Theorem 7.1.11 in [3]. The Lindeberg condition applied to the second sum means that

$$\frac{1}{n}\sum_{k=1}^{[nt]} \mathsf{E}^0\,|\varepsilon_k|^2 I\{|\varepsilon_k|^2 > \delta n\} \xrightarrow{\mathsf{P}} 0 \tag{19}$$

for any $t, \delta > 0$. It is clear that condition (19) follows from (18). Using (19) and (17), the Lindeberg condition for the first sum means that

$$\lim_{N\to\infty} \overline{\lim_{n\to\infty}} \frac{1}{n}\sum_{k=1}^{n} \operatorname{tr} B_k^N = 0 \tag{20}$$

with probability 1. In order to prove equality (20) note that $\|a\otimes b\|^2=|a|^2|b|^2$ and

$$\begin{aligned}&\mathsf{E}^0\,|\varepsilon_k|^2|\xi_{k-1}|^2 I\big\{|\varepsilon_k|^2|\xi_{k-1}|^2>\delta n\big\}\,(I\{|\xi_{k-1}|\le N\}+I\{|\xi_{k-1}|>N\})\\ &\qquad\le N^2\,\mathsf{E}^0\,|\varepsilon_k|^2 I\big\{|\varepsilon_k|^2>N^{-2}\delta n\big\}+\mathsf{E}^0\left(|\xi_{k-1}|^2 I\{|\xi_{k-1}|>N\}\,\mathsf{E}\left[|\varepsilon_k|^2|\,\mathcal{F}_{k-1}\right]\right),\\ &\qquad\qquad \mathsf{E}\left[|\varepsilon_k|^2|\,\mathcal{F}_{k-1}\right]=\operatorname{tr}\sigma_k^2\le d\|\sigma_k^2\|.\end{aligned}$$

Put

$$U^{jN}=\varlimsup_{n\to\infty}\frac{1}{n}\sum_{k=0}^{[n/m]}B_k^{jN}$$

and show that

$$\lim_{N\to\infty}U^{jN}=O\quad\text{a.s.}\tag{21}$$

for $j=1,\dots,m$. Since

$$\sum_{k=1}^{n}\operatorname{tr}B_k^N\le\sum_{j=1}^{m}\sum_{k=0}^{[n/m]}\operatorname{tr}B_k^{jN},$$

equality (21) implies (20).

Let $h=\|A^m\|$, $v_{np}(u)=\mathsf{E}^0\,\varepsilon_n^{\otimes2}I\{|\varepsilon_p|>u\}$, $q_n(u)=v_{nn}(u)$, and

$$N_l=(1-h)\big(m\|A\|^l\big)^{-1}N.$$

By Lemma 4, condition (17), and by the Chebyshev inequality, we have

$$B_k^{jN}\le h^2B_{k-1}^{jN}+C_1N^{-2}\,\mathsf{E}^0\,\xi_{k-1}^{[j]\otimes2}+2\,\mathsf{E}^0\,\varepsilon_k^{[j]\otimes2}I_k^{jN},\qquad k\ge1.\tag{22}$$

Then the immediate inequalities

$$\left(\sum_{i=1}^{m}a_i\right)^{\otimes2}\le m\sum_{i=1}^{m}a_i^{\otimes2},\qquad I_k^{jN}\le\sum_{l=0}^{m-1}I\{|\varepsilon_{km+j-l}|>N_l\}$$

imply that

$$\mathsf{E}^0\,\varepsilon_k^{[j]\otimes2}I_k^{jN}\le m\sum_{i,l=0}^{m-1}v_{km+j-i,km+j-l}(N_l).\tag{23}$$

Thus

$$v_{np}(u)=\mathsf{E}^0\left(I\{|\varepsilon_p|>u\}\,\mathsf{E}[\varepsilon_n^{\otimes2}|\,\mathcal{F}_{n-1}]\right)$$

for $n>p$. For $n<p$, we obtain

$$v_{np}(u)=\mathsf{E}^0\left(\varepsilon_n^{\otimes2}\mathsf{P}\{|\varepsilon_p|>u|\,\mathcal{F}_{p-1}\}\right).$$

By condition (17) and the Chebyshev inequality, this implies that

$$v_{np}(u)\le Cu^{-2}\,\mathsf{E}^0\,|\varepsilon_p|^2$$

in the first case (from now on, matrices proportional to the identity matrix are written as scalars) and

$$v_{np}(u)\le Cu^{-2}\,\mathsf{E}^0\,\varepsilon_n^{\otimes2}$$

in the second case. Using again inequality (17), we derive that

$$v_{np}(u)\le dC^2u^{-2},\qquad n\ne p,$$

whence

$$\mathsf{E}^0\,\varepsilon_k^{[j]\otimes2}I_k^{jN}\le m\sum_{i=0}^{m-1}q_{km+j-i}(N_i)+mdC^2\sum_{l=0}^{m-1}N_l^{-2}\tag{24}$$

by (23).

According to Lemma 1, we get

$$\sum_{i=0}^{\infty} A^i A^{\top i} \equiv \sum_{n=0}^{\infty} \sum_{j=0}^{m-1} A^j A^{mn} A^{\top mn} A^{\top j} \le \sum_{n=0}^{\infty} h^{2n} \sum_{j=0}^{m-1} A^j A^{\top j}$$

and

$$\mathsf{E}^0\, \xi_k^{\otimes 2} \le C_2(\xi_0^{\otimes 2} + 1) \tag{25}$$

in view of Lemma 2 and condition (7). Combining (22)–(24) we obtain

$$U^{jN} \le h^2 U^{jN} + C_3 \left(\sum_{i=0}^{m-1} \varlimsup_{n\to\infty} \frac{1}{n} \sum_{k=1}^{[n/m]} q_{km+j-i}(N_i) + N^{-2}\big(\xi_0^{\otimes 2} + 1\big) \right).$$

Now (21) follows from (18) and (7). □

Corollary 1. *Under the assumptions of Theorem 2,*

$$(Y_n, [Z_n]) \xrightarrow{\mathrm{C}} (Y, [Z]).$$

Proof. If $\mathsf{E}\,|\xi_0|^2 < \infty$, then

$$\sup_n \frac{1}{n} \sum_{k=1}^{[nt]} \mathsf{E}\,|\varepsilon_k|^2 |\xi_{k-1}|^2 < \infty$$

by (17) and (25). Moreover,

$$\sup_n \frac{1}{n}\, \mathsf{E} \max_{k\le nt} \|\varepsilon_k \otimes \xi_{k-1}\| < \infty,$$

and hence Corollary V1.6.7 in [4] gives the required result.

The general case is reduced to the above case by substituting $\xi_0 I\{|\xi_0| \le N\}$ instead of ξ_0 and letting N tend to infinity. □

By the well-known property of a continuous local martingale, $[Z] = \langle Z \rangle$. Therefore Corollary 1, in particular, states that

$$(Y_n(1), [Z_n](1)) \xrightarrow{\mathrm{d}} (Y(1), J).$$

This together with Theorem 1, equality (15), formula (25), and the notation

$$V = (\mathfrak{A} J)^{-1}$$

proves the following result.

Corollary 2. *Suppose that the assumptions of Theorem 2 hold and the matrix $\mathfrak{A}J$ is nondegenerate with probability* 1. *Then*

$$\sqrt{n}(\check{A}_n - A) \xrightarrow{\mathrm{d}} Y(1)V. \tag{26}$$

Below is an important particular case of Corollary 2.

Corollary 3. *Suppose that the assumptions of Corollary 2 hold and the tensors G and J are nonrandom. Then*

$$\sqrt{n}(\check{A}_n - A) \xrightarrow{\mathrm{d}} \Psi,$$

where Ψ is a normally distributed random matrix with zero mean and covariance tensor

$$\mathsf{E}\,\Psi^{\otimes 2} = V^{\otimes 2} \cdot G^*, \tag{27}$$

and where $\cdot$ is the convolution of $(2,2)$-tensors with respect to upper indices of the first operand and lower indices of the second operand.

Proof. Let $S = Y(1)$. By Theorem 2, Y is a continuous local martingale with square characteristic $\langle Y\rangle(t) = G^* t$. Since this function is nonrandom by assumption, the Lévy theorem implies that Y is a homogeneous Wiener process. Therefore S has a normal distribution. In the standard tensor notation, $\Psi \equiv SV = (s^i_k v^k_j)$ and

$$\Psi^{\otimes 2} = \left(s^{i_1}_{k_1} s^{i_2}_{k_2} v^{k_1}_{j_1} v^{k_2}_{j_2}\right),$$

implying $\mathsf{E}\,\Psi^{\otimes 2} = V^{\otimes 2} \cdot \mathsf{E}\,S^{\otimes 2}$. It remains to observe that the homogeneous Wiener process Y is such that

$$\langle Y\rangle(t) = t\,\mathsf{E}\,Y(1)^{\otimes 2}$$

and therefore $\mathsf{E}\,S^{\otimes 2} = G^*$. □

The above results cannot be considered to have a final form, since there is no indication of how Condition **CP** can be checked. Let us reduce this condition to the following simpler ones:

CP1. The sequence (J_n) converges in probability to a random matrix J.

CP2. For any nonnegative integers i, j and an integer $l > i + j$ the sequences

$$\left(\frac{1}{n}\sum_{k=l}^{n-1} \sigma^2_{k+1} \otimes \varepsilon_{k-i} \otimes \varepsilon_{k-j}\right) \quad \text{and} \quad \left(\frac{1}{n}\sum_{k=l}^{n-1} \sigma^2_{k+1} \otimes \varepsilon_{k-i}\right)$$

converge in probability.

Lemma 7. *Conditions* (5), (17), **CP1**, **CP2**, *and* (7) *for some positive integer* m *imply condition* **CP**.

Proof. Let

$$\xi^{(r)}_k = \sum_{i=0}^{(mr)\wedge k-1} A^i \varepsilon_{k-i}, \qquad G^{(r)}_n = \frac{1}{n}\sum_{k=0}^{n-1} \sigma^2_{k+1} \otimes \xi^{(r)\otimes 2}_k. \tag{28}$$

Taking into account (3) we have

$$\xi_k - \xi^{(r)}_k = A^k \xi_0 + \sum_{i=(mr)\wedge k}^{k-1} A^i \varepsilon_{k-i}.$$

Using (17), (5), and (7), we obtain

$$\mathsf{E}^0\left|\xi_k - \xi^{(r)}_k\right|^2 \le C_4\left(h^{2k/m}|\xi_0|^2 + h^{2r}\right). \tag{29}$$

In a similar manner, we can check that

$$\mathsf{E}^0\left(|\xi_k|^2 + \left|\xi^{(r)}_k\right|^2\right) \le C_5\left(|\xi_0|^2 + 1\right). \tag{30}$$

Introduce a norm in the space of 4-rank tensors in a way that $\|A_1 \otimes A_2\| = \|A_1\| \cdot \|A_2\|$ for arbitrary matrices A_1 and A_2. Then, by (17) and (28),

$$\mathsf{E}^0\left\|G_n - G^{(r)}_n\right\| \le \frac{C}{n}\sum_{k=0}^{n-1} \mathsf{E}^0\left\|\xi^{\otimes 2}_k - \xi^{(r)\otimes 2}_k\right\|.$$

This, together with

$$\left\|aa^\top - bb^\top\right\| \le (|a| + |b|)|a - b|,$$

the Cauchy–Schwarz inequality, (29), (30), and (7), shows that

$$\lim_{r\to\infty} \overline{\lim_{n\to\infty}}\, \mathsf{E}^0\left\|G_n - G^{(r)}_n\right\| = 0 \quad \text{a.s.} \tag{31}$$

According to (28),

$$G_n^{(r)} = \frac{1}{n}\sum_{k=0}^{mr-1} \sigma_{k+1}^2 \otimes \xi_k^{(r)\otimes 2} + \sum_{i,j=0}^{mr-1} \frac{1}{n}\sum_{k=mr}^{n-1} \sigma_{k+1}^2 \otimes \big(A^i(\varepsilon_{k-i}\otimes\varepsilon_{k-j})A^{\top j}\big) \tag{32}$$

for $n \geq mr$. Now we deduce from Condition **CP2** that the sequence $\big(G_n^{(r)}, n \in \mathbb{N}\big)$ converges for any r. Then, by (31), the sequence (G_n) is also convergent.

A similar argument applies to the sequence (H_n). □

In order to use Corollary 2, one should calculate the joint distribution of Y and V or, which is equivalent, the joint distribution of G and J. Section 6 shows the way to do it.

5. Estimating a matrix as a vector

A square $d \times d$ matrix A can be considered as a vector θ formed by the elements of A. There are lots of references on estimation of a vector parameter of an autoregressive process; see, for example, [5]–[7]. Let us compare our results in Section 4 to those in book [7] which are, at our knowledge, the most general. In this book, as in other sources, restrictions on the noise are imposed and this implies the asymptotic normality of the statistic $\sqrt{n}(\check{\theta}_n - \theta)$. Therefore we suppose that the assumptions of Corollary 3 hold and use the notation introduced therein.

By writing the rows of the matrix $\Psi = SV$ one after another, we obtain the following d^2-dimensional row vector:

$$\psi = \big(s^1 V \dots s^d V\big),$$

where s^i stands for the i-th row of the matrix S. Since the matrices J_n are symmetric by construction, the matrices J and V have the same property. Therefore,

$$\psi^\top \psi = \begin{pmatrix} V s^{1\top} s^1 V & \dots & V s^{1\top} s^d V \\ \dots & \dots & \dots \\ V s^{d\top} s^1 V & \dots & V s^{d\top} s^d V \end{pmatrix}.$$

Now, introduce the following $d^2 \times d^2$ matrices:

$$\mathbf{G} = \begin{pmatrix} \mathsf{E}\, s^{1\top} s^1 & \dots & \mathsf{E}\, s^{1\top} s^d \\ \dots & \dots & \dots \\ \mathsf{E}\, s^{d\top} s^1 & \dots & \mathsf{E}\, s^{d\top} s^d \end{pmatrix}, \qquad \mathbf{V} = \begin{pmatrix} V & O & \cdots & O \\ O & V & \dots & O \\ \dots & \dots & \dots & \dots \\ O & \dots & O & V \end{pmatrix}$$

giving a "plain" version

$$\mathsf{E}\, \psi^\top \psi = \mathbf{VGV}$$

of equality (27) (recall that $G^* = \mathsf{E}\, S^{\otimes 2}$ in this equality). This form of the covariance matrix of the limit normal distribution for $\sqrt{n}(\check{\theta}_n - \theta)$ is obtained in [7]. If the vector θ is not interpreted as the linearly extended matrix, then the matrix $\mathbf{V}$ will not be block-diagonal. In our case, this structure of the matrix is explained by the fact that the first component of the vector ξ_n depends on ξ_{n-1}, since the dependence is determined by the first d components of the vector θ; the dependence of the second component of ξ_n on ξ_{n-1} is due to the next d components of θ and so on.

6. An illustration: A scalar-valued second-order autoregression

Let a sequence (ζ_n) be defined by the following difference equation:

$$\zeta_n = \theta_1 \zeta_{n-1} + \theta_2 \zeta_{n-2} + \eta_n, \qquad n \in \mathbb{N}, \tag{33}$$

where (η_n) is a sequence of random variables, ζ_{-1} and ζ_0 are some random variables, θ_1 and θ_2 are two scalar parameters, $\theta_2 \neq 0$. The recurrence equation (33) is equivalent to (1), where

$$\xi_n = \begin{pmatrix} \zeta_n \\ \zeta_{n-1} \end{pmatrix}, \qquad \varepsilon_n = \begin{pmatrix} \eta_n \\ 0 \end{pmatrix}, \qquad A = \begin{pmatrix} \theta_1 & \theta_2 \\ 1 & 0 \end{pmatrix}. \tag{34}$$

Throughout below the symbol A denotes the latter matrix. Estimate (2) becomes of the following form in this case:

$$\check{A}_n = \begin{pmatrix} \sum \zeta_k\zeta_{k-1} & \sum \zeta_k\zeta_{k-2} \\ \sum \zeta_{k-1}^2 & \sum \zeta_{k-1}\zeta_{k-2} \end{pmatrix} \begin{pmatrix} \sum \zeta_k^2 & \sum \zeta_{k-1}\zeta_{k-2} \\ \sum \zeta_{k-1}\zeta_{k-2} & \sum \zeta_{k-2}^2 \end{pmatrix}^{\dagger} \tag{35}$$

(the index k in all sums runs from 1 to n). If the pseudoinverse matrix is an ordinary inverse matrix, then the second row of the matrix $\check{A}_n$ is equal to $(1 \quad 0)$.

It is known for square matrices D that a number $m \in \mathbb{N}$ such that $\|D^m\| < 1$ exists if and only if the spectrum of D belongs to the open unit circle. Using the characteristic equation

$$\lambda^2 - \theta_1\lambda - \theta_2 = 0 \tag{36}$$

we obtain the following result.

Lemma 8. *A natural number m such that $\|A^m\| < 1$ exists if and only if*

$$|\theta_1| < 1 - \theta_2 < 2. \tag{37}$$

Put $\varkappa_k^2 = \mathsf{E}[\eta_k^2 | \mathcal{F}_{k-1}]$, $\eta_n = 0$ for $n \leq 0$,

$$M = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad \alpha_n^{(rij)} = \frac{1}{n} \sum_{k=mr}^{n-1} \varkappa_{k+1}^2\, \eta_{k-i}\eta_{k-j}, \qquad \beta_n^{(ri)} = \frac{1}{n} \sum_{k=mr}^{n-1} \varkappa_{k+1}^2\, \eta_{k-i}.$$

The second of the equalities (34) implies that

$$\sigma_k^2 = \varkappa_k^2\, M; \tag{38}$$

thus the equality in (32) becomes of the form

$$G_n^{(r)} = M \otimes \left(\frac{1}{n} \sum_{k=0}^{mr-1} \varkappa_{k+1}^2\, \xi_k^{(r)\otimes 2} + \sum_{i,j=0}^{mr-1} A^i M A^{\top j} \alpha_n^{(rij)} \right). \tag{39}$$

Similarly we obtain for $H_n^{(r)} = n^{-1} \sum_{k=0}^{n-1} \sigma_{k+1}^2 \otimes \xi_k^{(r)}$ that

$$H_n^{(r)} = M \otimes \left(\frac{1}{n} \sum_{k=0}^{mr-1} \varkappa_{k+1}^2\, \xi_k^{(r)} + \sum_{i=0}^{mr-1} A^i \begin{pmatrix} 1 \\ 0 \end{pmatrix} \beta_n^{(ri)} \right). \tag{40}$$

In view of the definition of the matrix A we have

$$A^2 = \theta_1 A + \theta_2,$$

and this simplifies the evaluation of the limits. By induction,

$$A^n = c_{n1} A + c_{n2}, \qquad n \in \mathbb{Z}_+, \tag{41}$$

where the row $c_n = (c_{n1} \quad c_{n2})$ is obtained from the recurrence relation

$$c_n = c_{n-1} A, \qquad n \in \mathbb{N}, \tag{42}$$

starting with $c_0 = (0 \quad 1)$. Using (42) we derive the following recurrence relation:

$$c_{n1} - \theta_1 c_{n-1,1} - \theta_2 c_{n-2,1} = 0, \qquad n \geq 2,$$

for the first component. Denoting by

$$\lambda_1 = \frac{\theta_1 + \sqrt{\theta_1^2 + 4\theta_2}}{2}, \qquad \lambda_2 = \frac{\theta_1 - \sqrt{\theta_1^2 + 4\theta_2}}{2} \tag{43}$$

the roots of equation (36) and taking into account that $c_{10} = 0$, $c_{11} = 1$, we get

$$c_{n1} = \begin{cases} (\theta_1^2 + 4\theta_2)^{-1/2}(\lambda_1^n - \lambda_2^n) & \text{if } \theta_1^2 \neq -4\theta_2, \\ n(\theta_1/2)^{n-1} & \text{if } \theta_1^2 = -4\theta_2. \end{cases} \tag{44}$$

According to (42) we have

$$c_{n2} = \theta_2 c_{n-1,1}, \qquad n \in \mathbb{N}. \tag{45}$$

Assume that for all $i, j \in \mathbb{Z}_+$ there exist random variables $\alpha^{(ij)}$ and $\beta^{(i)}$ such that

$$\frac{1}{n}\sum_{k=0}^{n-1} \varkappa_{k+1}^2\, \eta_{k-i}\eta_{k-j} \xrightarrow{\mathsf{P}} \alpha^{(ij)}, \tag{46}$$

$$\frac{1}{n}\sum_{k=0}^{n-1} \varkappa_{k+1}^2\, \eta_{k-i} \xrightarrow{\mathsf{P}} \beta^{(i)} \tag{47}$$

as $n \to \infty$. Put

$$\phi_1 = \sum_{i,j=0}^{\infty} c_{i1}c_{j1}\alpha^{(ij)}, \qquad \phi_2 = \sum_{i,j=0}^{\infty} c_{i1}c_{j2}\alpha^{(ij)}, \qquad \phi_3 = \sum_{i,j=0}^{\infty} c_{i2}c_{j2}\alpha^{(ij)}, \tag{48}$$

$$\mu_l = \sum_{i=0}^{\infty} c_{il}\beta^{(i)}, \qquad l = 1, 2, \tag{49}$$

$$\Phi = \phi_1 \begin{pmatrix} \theta_1^2 & \theta_1 \\ \theta_1 & 1 \end{pmatrix} + \phi_2 \begin{pmatrix} 2\theta_1 & 1 \\ 1 & 0 \end{pmatrix} + \phi_3 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}. \tag{50}$$

Lemma 9. *Let conditions* (37), (46), *and* (47) *hold. Assume that*

$$\varkappa_k^2 \leq C \tag{51}$$

for all k. Then

1) *the series on the right-hand sides of* (48) *and* (49) *absolutely converge in probability;*
2) $\phi_1 \geq 0$ *and* $\phi_2^2 \leq \phi_1\phi_3$.

Proof. 1) It is obvious that

$$\left(\frac{1}{n}\sum_{k=0}^{n-1} \varkappa_{k+1}^2\, \eta_{k-i}\eta_{k-j}\right)^2 \leq \frac{1}{n}\sum_{k=0}^{n-1} \varkappa_{k+1}^4\, \eta_{k-i}^2\eta_{k-j}^2,$$

whence

$$\mathsf{E}\big(\alpha^{(ij)}\big)^2 \leq C^4 \quad \text{for } i \neq j$$

by (51) and (46). Similarly relations (51) and (46) imply that $\mathsf{E}\,\alpha^{(ii)} \leq C^2$. Thus

$$\mathsf{E}\sum_{i,j=0}^{\infty} |c_{i1}| \cdot |c_{j1}| \cdot \big|\alpha^{(ij)}\big| \leq C^2 \left(\sum_{i=0}^{\infty} |c_{i1}|\right)^2. \tag{52}$$

According to Lemma 8, the absolute values of the roots of equation (36) do not exceed 1. Moreover, if the roots are equal (this is the case if $\theta_1^2 = -4\theta_2$; see (43)), then $|\theta_1| < 2$. Thus the convergence of the series on the right-hand side of (52) follows from (44), whence we conclude that the series in the definition of ϕ_1 absolutely converges in probability and

in L_1. The reasoning for the rest of the series in (48) and (49) is the same (we apply (45); in the case of μ_l we use (47) instead of (46)).

2) Put $\nu_{lk}^{(N)} = \sum_{i=0}^{N} c_{il}\eta_{k-i}$. It follows from (46) and the first assertion of the lemma that

$$\begin{aligned}\phi_1 &= \underset{N\to\infty}{\mathrm{l.\,i.\,p.}}\,\underset{n\to\infty}{\mathrm{l.\,i.\,p.}}\,\frac{1}{n}\sum_{k=0}^{n-1}\kappa_{k+1}^2\sum_{i,j=0}^{N}c_{i1}c_{j1}\eta_{k-i}\eta_{k-j}\\ &\equiv \underset{N\to\infty}{\mathrm{l.\,i.\,p.}}\,\underset{n\to\infty}{\mathrm{l.\,i.\,p.}}\,\frac{1}{n}\sum_{k=0}^{n-1}\varkappa_{k+1}^2\,\nu_{1k}^{(N)2},\end{aligned}\tag{53}$$

whence $\phi_1 \geq 0$. Similar equalities for ϕ_2 and ϕ_3 yield

$$\phi_1\phi_3 - \phi_2^2 = \underset{N\to\infty}{\mathrm{l.\,i.\,p.}}\,\underset{n\to\infty}{\mathrm{l.\,i.\,p.}}\,n^{-2}T_n$$

(we omit the superscript N), where

$$\begin{aligned}T_n &= \sum_{k,p=0}^{n-1}\varkappa_{k+1}^2\,\varkappa_{p+1}^2(\nu_{1k}^2\nu_{2p}^2 - \nu_{1k}\nu_{2k}\nu_{1p}\nu_{2p})\\ &\equiv \sum\varkappa_{k+1}^2\,\nu_{1k}^2\sum\varkappa_{k+1}^2\,\nu_{2k}^2 - \Big(\sum\varkappa_{k+1}^2\,\nu_{1k}\nu_{2k}\Big)^2 \geq 0. \qquad \square\end{aligned}$$

Lemma 10. *Let conditions* (37), (46), (47), *and* (51) *hold. Assume that*

$$\mathsf{E}[\eta_k|\,\mathcal{F}_{k-1}] = 0\tag{54}$$

for all k. Then

$$G_n \xrightarrow{\mathsf{P}} G \equiv M\otimes\Phi,\tag{55}$$

$$H_n \xrightarrow{\mathsf{P}} H \equiv M\otimes\left(\mu_1\begin{pmatrix}\theta_1\\1\end{pmatrix} + \mu_2\begin{pmatrix}1\\0\end{pmatrix}\right)\tag{56}$$

as $n\to\infty$.

Recall that μ_1, μ_2, and Φ are defined by equalities (49) and (50), respectively.

Proof. Put $\Phi^{(r)} = \sum_{i,j=0}^{mr-1}\alpha^{(ij)}A^iMA^{\top j}$. Using (39), (40), (46), and (47) we get

$$\underset{n\to\infty}{\mathrm{l.\,i.\,p.}}\,G_n^{(r)} = G^{(r)} \equiv M\otimes\Phi^{(r)},\tag{57}$$

$$\underset{n\to\infty}{\mathrm{l.\,i.\,p.}}\,H_n^{(r)} = H^{(r)} \equiv M\otimes\sum_{i=0}^{mr-1}\beta^{(i)}A^i\begin{pmatrix}1\\0\end{pmatrix}.\tag{58}$$

In view of equalities (34) and (38), relations (53) and (51) for η_k become of the form (5) and (17) for ε_k. Note that we used conditions (5), (17), and (7) to prove (30). Moreover, (7) is equivalent to (37) in the case under consideration. To derive (55) from (30) and (57), it is sufficient to show that

$$\underset{r\to\infty}{\mathrm{l.\,i.\,p.}}\,\Phi^{(r)} = \Phi.\tag{59}$$

Since the product F_1MF_2 is equal to the product of the first column of F_1 to the first column of F_2 for arbitrary 2×2 matrices F_1 and F_2, we deduce from (41) that

$$A^iMA^{\top j} = c_{i1}c_{j1}\begin{pmatrix}\theta_1^2 & \theta_1\\ \theta_1 & 1\end{pmatrix} + c_{i1}c_{j2}\begin{pmatrix}\theta_1 & 1\\ 1 & 0\end{pmatrix} + c_{i2}c_{j1}\begin{pmatrix}\theta_1 & 0\\ 1 & 0\end{pmatrix} + c_{i2}c_{j2}\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}.$$

Multiplying both sides of the equality by $\alpha^{(ij)}$, then summing with respect to i and j from 0 to $mr-1$ (since $\alpha^{(ij)} = \alpha^{(ji)}$ in view of (46), we get $\sum c_{i1}c_{j2}\alpha^{(ij)} = \sum c_{i2}c_{j1}\alpha^{(ij)}$), and applying Lemma 9 we obtain (59).

Relation (56) is obtained from (58) in a similar manner. □

Theorem 3. *Let conditions* (37), (46), (47), (51), *and* (54) *hold for an autoregressive process defined by* (33). *Assume that*

$$\frac{1}{n}\sum_{k=1}^{n} \varkappa_k^2 \xrightarrow{\mathsf{P}} \gamma^2, \tag{60}$$

$$\mathsf{P}\{\gamma^2 > 0\} = 1, \tag{61}$$

$$\lim_{N\to\infty} \overline{\lim_{n\to\infty}} \frac{1}{n}\sum_{k=1}^{n} \mathsf{E}\,\eta_k^2 I\{|\eta_k| > N\} = 0. \tag{62}$$

Then the estimator (35) *of the matrix A defined by* (34) *is such that*

$$\sqrt{n}(\check{A}_n - A) \xrightarrow{\mathrm{d}} \frac{1+\theta_2}{\gamma^2}\begin{pmatrix} (1-\theta_2)\tau_1 - \theta_1\tau_2 & (1-\theta_2)\tau_2 - \theta_1\tau_1 \\ 0 & 0 \end{pmatrix},$$

where $\tau_1 = \sum_{l=1}^{3} \rho_l \mathfrak{n}_l$, $\tau_2 = \sqrt{\phi_1}\mathfrak{n}_1$,

$$\rho_1 = \big(\phi_1^{1/2}\theta_1 + \phi_1^{-1/2}\phi_2\big) I\{\phi_1 > 0\}, \qquad \rho_2 = \big(\phi_3 - \phi_2^2/\phi_1\big)^{1/2} I\{\phi_1 > 0\}, \qquad \rho_3 = \sqrt{\phi_3} I\{\phi_1 = 0\}, \tag{63}$$

$\mathfrak{n}_1$, $\mathfrak{n}_2$, *and* $\mathfrak{n}_3$ *are independent standard normal random variables that are independent of* γ *and all* $\alpha^{(ij)}$ *and* $\beta^{(i)}$.

Proof. It follows from Lemma 10 that condition **CP** holds for (G_n) and (H_n). For (J_n), condition **CP** follows from (60) and (38), namely $J_n \xrightarrow{\mathsf{P}} J \equiv \gamma^2 M$, whence

$$\mathfrak{A}J = \gamma^2 \mathfrak{A}M. \tag{64}$$

Put

$$\Delta = (1-\theta_2)(1-\theta_1^2-\theta_2^2) - 2\theta_1^2\theta_2 \equiv (1+\theta_2)\big((1-\theta_2)^2 - \theta_1^2\big). \tag{65}$$

By (37) we have $\Delta \neq 0$. Then

$$\mathfrak{A}M = \frac{1}{\Delta}\begin{pmatrix} 1-\theta_2 & \theta_1 \\ \theta_1 & 1-\theta_2 \end{pmatrix}. \tag{66}$$

We obtain the latter equality by substituting the matrix $\mathfrak{A}M$ instead of X in equation (10) with $F = M$. Then (64) together with (61), (66), and (37) implies that $\mathfrak{A}J$ is nondegenerate. Taking into account (65) we evaluate its inverse matrix

$$V = \frac{1}{\gamma^2(1+\theta_2)}\begin{pmatrix} 1-\theta_2 & -\theta_1 \\ -\theta_1 & 1-\theta_2 \end{pmatrix}. \tag{67}$$

Therefore all the assumptions of Corollary 2 hold (relation (18) coincides with (62) in this case). We show that (26) is equivalent to the conclusion of the theorem.

Put

$$\bar{\rho}_1 = \begin{pmatrix} \rho_1 \\ \sqrt{\phi_1} \end{pmatrix}, \qquad \bar{\rho}_2 = \begin{pmatrix} \rho_2 \\ 0 \end{pmatrix}, \qquad \bar{\rho}_3 = \begin{pmatrix} \rho_3 \\ 0 \end{pmatrix}. \tag{68}$$

Then the second assertion of Lemma 9 together with (63) allows one to rewrite equality (50) as follows:

$$\Phi = \sum_{l=1}^{3} \bar{\rho}_l^{\otimes 2}. \tag{69}$$

Let the sequence (Y_n) converge to a continuous local martingale Y (the convergence is due to Theorem 2). Then Y can be represented in the following form:

$$Y = \sum_{l=1}^{3} w_l \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \bar{\rho}_l, \tag{70}$$

where w_1, w_2, and w_3 are standard Wiener processes that are independent of γ and all $\alpha^{(ij)}$ and $\beta^{(i)}$. Indeed, the right-hand side of equality (70) (for a moment, we denote it by Y') is a continuous local martingale such that

$$(\langle Y' \rangle(t))^* = t \sum_{l=1}^{3} \begin{pmatrix} 1 \\ 0 \end{pmatrix}^{\otimes 2} \otimes \bar{\rho}_l^{\otimes 2} = tM \otimes \sum_{l=1}^{3} \bar{\rho}_l^{\otimes 2}.$$

This, together with (69) and (55), shows that $\langle Y' \rangle(t) = G^* t$, whence $Y' \overset{\mathrm{d}}{=} Y$. Since Theorem 2 deals with the convergence of distributions, one can identify Y with Y'.

Using (68) and (70) we get the following representation for the first row of the matrix Y (the second row is zero):

$$(1 \quad 0) Y = \sum_{l=1}^{3} w_l \bar{\rho}_l^{\top}.$$

Putting $\mathfrak{n}_l = w_l(1)$ and considering (67), we derive the theorem from Corollary 2. □

Example. Let $\eta_k = \gamma_{k-1}\chi_k$, where (γ_k) and (χ_k) are independent sequences of random variables, χ_k are jointly independent random variables, $|\gamma_n| \le C$, $\mathsf{E}\,\chi_k = 0$, $\mathsf{E}\,\chi_k^2 = 1$, and

$$\underset{n\to\infty}{\mathrm{l.\,i.\,p.}}\, \gamma_n^2 = \gamma^2, \tag{71}$$

$$\lim_{N\to\infty} \overline{\lim_{n\to\infty}} \frac{1}{n} \sum_{k=1}^{n} \mathsf{E}\,\chi_k^2 I\{|\chi_k| > N\} = 0. \tag{72}$$

Assume that conditions (37) and (61) hold. Let $\mathcal{F}_k$ be the σ-algebra generated by the random variables $\zeta_{-1}, \zeta_0, \gamma_0, \dots, \gamma_{k-1}, \chi_1, \dots, \chi_k$. Then

$$\varkappa_{k+1}^2 = \gamma_k^2. \tag{73}$$

Conditions (54), (51), and (60) are satisfied in this case. Relation (62) follows from (72) and the uniform boundedness of γ_n^2.

Relations (71) and (73) imply that

$$\frac{1}{n} \sum_{k=0}^{n-1} \varkappa_{k+1}^2\, \eta_{k-i}^2 - \frac{\gamma^4}{n} \sum_{k=1}^{n} \chi_k^2 \xrightarrow{\mathsf{P}} 0$$

for all i. This together with (72) shows that condition (46) holds for $j = i$ and $\alpha^{(ii)} = \gamma^4$. Since the terms in the sum $\sum_{k=1}^{n} \gamma_k^2 \gamma_{k-i-1}\gamma_{k-j-1}\chi_{k-i}\chi_{k-j}$ are uncorrelated for $j \ne i$, we have $\alpha^{(ij)} = 0$. The same reasoning proves that $\beta^{(i)} = 0$. As a result, equalities (48) become of the form

$$\phi_1 = \gamma^4 \sum_{i=0}^{\infty} c_{i1}^2, \qquad \phi_2 = \gamma^4 \sum_{i=0}^{\infty} c_{i1}c_{i2}, \qquad \phi_3 = \gamma^4 \sum_{i=0}^{\infty} c_{i2}^2.$$

Equalities (44) and (45) help to evaluate the sums of the corresponding series. This gives the expressions for τ_1 and τ_2 in Theorem 3.

Bibliography

1. M. Arató, *Linear stochastic systems with constant coefficients: A statistical approach*, Springer-Verlag, Berlin, 1982. MR1011456 (90i:62108)
2. V. O. Koval, *Limit theorems for operator-normalized random vectors. II*, Teor. Ĭmovīr. Mat. Stat. **62** (2000), 37–47; English transl. in Theory Probab. Math. Statist. **62** (2001), 39–49. MR1871507 (2002j:60051
3. R. Sh. Liptser and A. N. Shiryaev, *Theory of Martingales*, "Nauka", Moscow, 1986; English transl., Kluwer, Dordrecht, 1989. MR1022664 (90j:60046)
4. J. Jacod and A. N. Shiryaev, *Limit Theorems for Stochastic Processes*, Springer-Verlag, Berlin, 1987. MR0959133 (89k:60044)
5. T. W. Anderson, *The Statistical Analysis of Time Series*, John Wiley & Sons, Inc., New York, 1971. MR0283939 (44:1169)
6. I. Vuchkov, L. Boyadzhieva, and E. Solakov, *Applied Linear Regression Analysis*, "Finansy i Statistika", Moscow, 1987. (Russian) MR0935130 (89f:62058)
7. A. Ya. Dorogovtsev, *The Theory of Estimates of the Parameters of Random Processes*, "Vishcha Shkola", Kiev, 1982. (Russian) MR0668517 (84h:62122)

Department of Mathematics and Theoretical Radiophysics, Faculty of Radiophysics, Taras Shevchenko National University, Glushkov Ave. 2, Building 5, 03127 Kyïv, Ukraine

E-mail address: yap@univ.kiev.ua

Department of Mathematics and Theoretical Radiophysics, Faculty of Radiophysics, Taras Shevchenko National University, Glushkov Ave. 2, Building 5, 03127 Kyïv, Ukraine

E-mail address: ida@univ.kiev.ua

Received 24/MAY/2004

Translated by V. ZAYATS