Theory of Stochastic Processes
Vol. 12 (28), no. 1–2, 2006, pp. 154–161

UDC 519.21

A. P. YURACHKIVSKY AND D. O. IVANENKO

MATRIX PARAMETER ESTIMATION IN AN AUTOREGRESSION MODEL

The vector difference equation $\xi_k = Af(\xi_{k-1}) + \varepsilon_k$, where (ε_k) is a square integrable difference martingale, is considered. A family of estimators $\check{A}_n$ depending, besides the sample size n, on a bounded Lipschitz function is constructed. Convergence in distribution of $\sqrt{n}\left(\check{A}_n - A\right)$ as $n \to \infty$ is proved with the use of stochastic calculus. Ergodicity and even stationarity of (ε_k) is not assumed, so the limiting distribution may be, as the example shows, other than normal.

Introducton

We consider the vector autoregression process

$$\xi_k = Af(\xi_{k-1}) + \varepsilon_k, \quad k \in \mathbb{N}. \tag{1}$$

Here, A is an unknown square matrix, f is a prescribed function, and (ε_k) is a square integrable difference martingale with respect to some flow $(\mathcal{F}_k, k \in \mathbb{Z}_+)$ of σ-algebras such that the random variable ξ_0 is $\mathcal{F}_0$-measurable. In the detailed form, the assumption about (ε_k) means that for any k ε_k is $\mathcal{F}_k$-measurable,

$$\mathrm{E}|\varepsilon_k|^2 < \infty \tag{2}$$

and

$$\mathrm{E}(\varepsilon_k | \mathcal{F}_{k-1}) = 0. \tag{3}$$

All vectors are regarded, unless otherwise stated, as columns. Then $a^\top b$ and $ab^\top$ signify scalar and tensor product respectively. The latter is otherwise denoted $a \otimes b$ (this is a $(0,2)$-tensor), in particular $a^{\otimes 2} = aa^\top$. We use the Euclidean norm of vectors, denoting it $|\cdot|$, and the operator norm of matrices. Other notation: $B^\dagger$ – the pseudoinverse to B; O – the null matrix; l.i.p. – limit in probability; $\xrightarrow{\mathrm{d}}$ – the weak convergence of the finite-dimensional distributions of random functions, in particular the convergence in distribution of random vectors.

Let h be a vector function such that for some n

$$\mathrm{E}\left(|\xi_n| + |Af(\xi_n)|\right)|h(\xi_{n-1})| + \mathrm{E}|h(\xi_{n-1})| < \infty.$$

Then from (1) – (3) we have $\mathrm{E}\left(\xi_n - Af(\xi_{n-1})\right) \otimes h(\xi_{n-1}) = \mathrm{O}$, whence

$$A = \left(\mathrm{E}\xi_n \otimes h(\xi_{n-1})\right)\left(\mathrm{E}f(\xi_{n-1}) \otimes h(\xi_{n-1})\right)^{-1}$$

2000 *AMS Mathematics Subject Classification.* Primary 62F12; Secondary 60F05.
Key words and phrases. Autoregression, martingale, estimator, tensor, convergence.

provided the inverse exists. This prompts the estimator

$$\check{A}_n = \left(\sum_{k=1}^{n} \xi_k \otimes h(\xi_{k-1})\right) \left(\sum_{k=1}^{n} f(\xi_{k-1}) \otimes h(\xi_{k-1})\right)^{\dagger} \tag{4}$$

coinciding in the case $f(x) = x$ with the LSE.

The goal of the article is to study the asymptotic behaviour of the normalized deviation $\sqrt{n}\,\big(\check{A}_n - A\big)$ as $n \to \infty$. The use of stochastic calculus underlying our approach allows us to dispense with the assumptions of ergodicity and even asymptotic stationarity of the sequence (ε_k), thereat the limiting distribution of the studied statistic may be other than normal. This is the main distinction of our results from A.Ya. Dorogovtsev's ones [1] essentially based on the ergodicity assumption.

Preliminaries

Let E^0 denote $\mathrm{E}(\cdots|\mathcal{F}_0)$.

Lemma 1. *Let conditions (2) and (3) be fulfilled and there exist a number q such that for all x*

$$|Af(x)| \le q|x|. \tag{5}$$

Then, for any k,

$$\mathrm{E}^0|\xi_k|^2 \le q^{2k}|\xi_0|^2 + \sum_{i=0}^{k-1} q^i \mathrm{E}^0|\varepsilon_{k-i}|^2.$$

Proof. Writing, on the basis of (1),

$$|\xi_k|^2 = |Af(\xi_{k-1})|^2 + 2Af(\xi_{k-1})^\top \varepsilon_k + |\varepsilon_k|^2 \tag{6}$$

we deduce our assertion from (2), (3) and (5) by induction.

Denote further $\sigma_k^2 = \mathrm{E}\left(\varepsilon_k^{\otimes 2}|\mathcal{F}_{k-1}\right)$, $\chi_k^N = I\{|\xi_k| > N\}$, $I_k^N = I\{|\varepsilon_k| > (1-q)N\}$, $b_k^N = \mathrm{E}^0|\xi_k|^2\chi_k^N$. Obviously,

$$\mathrm{E}\left(|\varepsilon_k|^2|\mathcal{F}_{k-1}\right) = \mathrm{tr}\sigma_k^2. \tag{7}$$

Lemma 2. *Let conditions (2), (3) and (5) be fulfilled and*

$$q < 1. \tag{8}$$

Then for any k

$$b_k^N \le q^2 b_{k-1}^N + \mathrm{E}^0|\varepsilon_k|^2\chi_{k-1}^N + 2(q/(1-q))^2 N^{-2}\mathrm{E}^0|\xi_{k-1}|^2\mathrm{tr}\sigma_k^2 + 2\mathrm{E}^0|\varepsilon_k|^2 I_k^N.$$

Proof. Due to (1) and (5),

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

which together with (6), (5) and the obvious inequality $|a^\top b| \le |a|^2 + |b|^2$ yields

$$|\xi_k|^2\chi_k^N \le q^2|\xi_{k-1}|^2\chi_{k-1}^N + 2Af(\xi_{k-1})^\top\varepsilon_k\chi_{k-1}^N + |\varepsilon_k|^2\chi_{k-1}^N + 2(q^2|\xi_{k-1}|^2 + |\varepsilon_k|^2)I_k^N.$$

By Lemma 1 and condition (5), $\mathrm{E}^0|Af(\xi_{k-1})|^2 < \infty$. Hence, because of (2) and (3), $\mathrm{E}^0\left(Af(\xi_{k-1})^\top\varepsilon_k|\mathcal{F}_{k-1}\right) = 0$. The equality

$$\mathrm{E}^0|\xi_{k-1}|^2 I_k^N = \mathrm{E}^0\left(|\xi_{k-1}|^2\mathrm{P}\{|\varepsilon_k| > (1-q)N|\mathcal{F}_{k-1}\}\right),$$

together with condition (8), Chebyshev's inequality, and equality (7) completes the proof.

In what follows, C is a generic constant.

Obviously, $\mathrm{E}^0\chi_i^N \le N^{-2}b_i^N$. Hence and from the previous lemmas we deduce (the details can be found in the proof of Theorem 2 [2])

Corollary 1. *Let conditions (2), (3), (5), and (8) be fulfilled,*

$$\lim_{N\to\infty}\overline{\lim_{n\to\infty}}\frac{1}{n}\sum_{k=1}^{n}\mathrm{E}|\varepsilon_k|^2 I\{|\varepsilon_k|>N\}=0, \tag{9}$$

and let there exist an $\mathcal{F}_0$-measurable random variable υ such that for all k

$$\mathrm{E}\left(|\varepsilon_k|^2|\mathcal{F}_{k-1}\right)\le\upsilon. \tag{10}$$

Then with probability 1

$$\lim_{N\to\infty}\overline{\lim_{n\to\infty}}\frac{1}{n}\sum_{k=1}^{n}b_k^N=0.$$

The main results

Let h be a Borel function such that

$$|h(x)|\le C|x|. \tag{11}$$

Denote $\eta_k=h(\xi_k)$, $K_n=\frac{1}{\sqrt{n}}\sum_{k=1}^{n}\varepsilon_k\otimes\eta_{k-1}$, $Q_n=\frac{1}{n}\sum_{k=0}^{n-1}f(\xi_k)\otimes\eta_k$, $T_n=\sqrt{n}(AQ_nQ_n^{\dagger}-A)$, $G_n=\frac{1}{n}\sum_{k=1}^{n}\sigma_k^2\otimes\eta_{k-1}^{\otimes 2}$,

$$Y_n(t)=\frac{1}{\sqrt{n}}\sum_{k=1}^{[nt]}\varepsilon_k\otimes\eta_{k-1}. \tag{12}$$

Then, because of (4),

$$\sqrt{n}(\check{A}_n-A)=K_nQ_n^{\dagger}+T_n. \tag{13}$$

By construction and conditions (2), (3) and (5), Y_n is a locally square integrable martingale with quadratic characteristic

$$\langle Y_n\rangle(t)=n^{-1}[nt]G^{*}_{[nt]}, \tag{14}$$

where $*$ is a linear operation in the space of 4-valent tensors such that

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

Theorem 1. *Let conditions (2), (3), (5) and (8) – (11) be fulfilled, and let there exist a random $(0,4)$-tensor G such that*

$$G_n\xrightarrow{\mathrm{d}}G. \tag{15}$$

Then $Y_n\xrightarrow{\mathrm{d}}Y$, where Y is a continuous local martingale with quadratic characteristic $\langle Y\rangle(t)=G^{*}t$.

Proof. According to Corollary in [3] and in view of (14) and (15), it suffices to show that for any t

$$\mathrm{E}\sup_{s\le t}\|Y_n(s)-Y_n(s-)\|^2\to 0. \tag{16}$$

The argument in [3] doess not change if the expectation is taken conditioned on $\mathcal{F}_0$, so in (16) E and $\to$ may be substituted by E^0 and $\xrightarrow{\mathrm{P}}$, respectively. This weakened version of (16) is equivalent, because of (12), to the following relation:

$$n^{-1}\mathrm{E}^0\max_{k\le nt}\rho_k\xrightarrow{\mathrm{P}}0,$$

where $\rho_k = |\varepsilon_k|^2|\eta_{k-1}|^2$. Since for any $\delta > 0$

$$\max_k \rho_k \le \delta n + \sum_k \rho_k I\{\rho_k > \delta n\},$$

it remains to prove that the random variables $\sqrt{\rho_k/n}$ satisfy the Lindeberg condition: for any $\delta > 0$

$$\frac{1}{n}\sum_{k\le nt} \mathrm{E}^0 \rho_k I\{\rho_k > \delta n\} \xrightarrow{\mathrm{P}} 0. \tag{17}$$

Writing on the basis of (11)

$$\rho_k I\{\rho_k > \delta n\}\left(I\{|\xi_{k-1}| \le N\} + I\{|\xi_{k-1}| > N\}\right)$$
$$\le C^2\left(N^2|\varepsilon_k|^2 I\{|\varepsilon_k|^2 > (CN)^{-2}\delta n\} + |\varepsilon_k|^2|\xi_{k-1}|^2\chi_{k-1}^N\right),$$

we deduce (17) from both the conditions and the conclusion of Corollary 1.

Applying Theorem 1 to the compound processes (Y_n, Q_n) where the second component does not depend on t, we obtain

Corollary 2. *Let conditions (2), (3), (5), and (8) – (11) be fulfilled, and let there exist given on a common probability space random $(0,4)$-tensor G and $(0,2)$-tensor Q such that*

$$(G_n, Q_n) \xrightarrow{\mathrm{d}} (G, Q). \tag{18}$$

Then $(Y_n, Q_n) \xrightarrow{\mathrm{d}} (Y, Q)$, where Y is a continuous local martingale w. r. t. some flow $(\mathcal{F}(t), t \in \mathbb{R}_+)$ such that $\langle Y\rangle(t) = G^* t$ and the tensor-valued r. v. Q is $\mathcal{F}(0)$-measurable.

Theorem 2. *Let the conditions of Corollary 2 be fulfilled and* $\det Q \ne 0$ *a. s. Then*

$$\sqrt{n}(\check{A}_n - A) \xrightarrow{\mathrm{d}} Y(1)Q^{-1}. \tag{19}$$

Proof. By Corollary 2,

$$(Y_n(1), Q_n) \xrightarrow{\mathrm{d}} (Y(1), Q).$$

But $Y_n(1) = K_n$, which together with the nondegeneracy of Q implies that $K_n Q_n^\dagger \xrightarrow{\mathrm{d}} K Q^{-1}$. Now, to obtain the assertion of the theorem from (13), it remains to note that

$$\mathrm{P}\{T_n \ne \mathrm{O}\} \le \mathrm{P}\{\det Q_n \ne \mathrm{O}\} \to 0.$$

Simpler versions of condition (18)

Denote $f_0(x) = x$ and, for $r \ge 1$,

$$f_r(x_0, \dots, x_r) = Af(f_{r-1}(x_0, \dots, x_{r-1})) + x_r. \tag{20}$$

Then

$$\xi_k = f_r(\xi_{k-r}, \varepsilon_{k-r+1}, \dots, \varepsilon_k), \quad r < k, \tag{21}$$

and

$$|f_r(x_0, \dots, x_r)| \le \sum_{i=0}^r q^i |x_{r-i}|. \tag{22}$$

Below X_r stands for $(x_1, \dots, x_r)$, and d is the dimensionality of each x_j.

Lemma 3. *Let for all x, y*

$$|Af(x) - Af(y)| \leq q|x-y|. \tag{23}$$

Then for all x, y, r, X_r

$$|f_r(x, X_r) - f_r(y, X_r)| \leq q^r|x-y|.$$

Proof. Due to (20) and (23),

$$|f_r(x, X_r) - f_r(y, X_r)| \leq q|f_{r-1}(x, X_{r-1}) - f_{r-1}(y, X_{r-1})|,$$

so it remains to apply the induction.

Corollary 3. *Under the conditions of Lemma 3, for any N*

$$\lim_{r\to\infty} \sup_{|x|\leq N, X_r\in\mathbb{R}^{rd}} |f_r(x, X_r) - f_r(0, X_r)| = 0.$$

Corollary 4. *Let conditions (5), (8), (11) and (23) be fulfilled and for any x, y*

$$|h(x) - h(y)| \leq C|x-y|. \tag{24}$$

Then for any $N > 0$

$$\lim_{r\to\infty} \sup_{|x|\leq N, X_r\in\mathbb{R}^{rd}} \|f(f_r(x, X_r)) \otimes h(f_r(x, X_r)) - f(f_r(0, X_r)) \otimes h(f_r(0, X_r))\| = 0, \tag{25}$$

$$\lim_{r\to\infty} \sup_{|x|\leq N, X_r\in\mathbb{R}^{rd}} \|h(f_r(x, X_r))^{\otimes 2} - h(f_r(0, X_r))^{\otimes 2}\| = 0. \tag{26}$$

Denote further $\xi_k^r = f_r(0, \varepsilon_{k-r+1}, \ldots, \varepsilon_k)$, $\eta_k^r = h(\xi_k^r)$, $Q_n^r = \frac{1}{n}\sum_{k=r}^{n-1} f(\xi_k^r) \otimes \eta_k^r$, $G_n^r = \frac{1}{n}\sum_{k=r}^{n} \sigma_k^2 \otimes (\eta_{k-1}^r)^{\otimes 2}$. We endow the space of $(0,4)$-tensors with such a norm that for any $(0,2)$-tensors A_1 and A_2, $\|A_1 \otimes A_2\| = \|A_1\|\|A_2\|$.

Lemma 4. *Let conditions (2), (3), (5), (8) – (11), (23) and (24) be fulfilled and*

$$|f(x)| \leq C|x|. \tag{27}$$

Then almost surely

$$\lim_{r\to\infty} \overline{\lim_{n\to\infty}} \mathrm{E}^0 \|Q_n - Q_n^r\| = 0, \tag{28}$$

$$\lim_{r\to\infty} \overline{\lim_{n\to\infty}} \mathrm{E}^0 \|G_n - G_n^r\| = 0. \tag{29}$$

Proof. By Corollary 4 for any $N > 0$,

$$\lim_{r\to\infty} \overline{\lim_{n\to\infty}} \frac{1}{n} \sum_{k=r}^{n-1} \mathrm{E}\|f(\xi_k) \otimes \eta_k - f(\xi_k^r) \otimes \eta_k^r\| I\{|\xi_k| \leq N\} = 0. \tag{30}$$

Due to (11) and (27),

$$\mathrm{E}^0 \|f(\xi_k) \otimes \eta_k\| \chi_k^N \leq C^2 b_k^N,$$

so, by Corollary 1,

$$\lim_{N\to\infty} \overline{\lim_{n\to\infty}} \frac{1}{n} \sum_{k=0}^{n-1} \mathrm{E}^0 \|f(\xi_k) \otimes \eta_k\| \chi_k^N = 0. \tag{31}$$

Further, for $k \geq r$

$$\mathrm{E}^0 \|f(\xi_k^r) \otimes \eta_k^r\| = \mathrm{E}^0 |f(f_r(0, \varepsilon_{k-r+1}, \ldots, \varepsilon_k))||h(f_r(0, \varepsilon_{k-r+1}, \ldots, \varepsilon_k))|,$$

whence, in view of (22), (27), and (11),

$$\mathrm{E}\|f(\xi_k^r)\otimes\eta_k^r\|\chi_k^N \le C^2\mathrm{E}\left(\sum_{i=0}^{r-1} q^i|\varepsilon_{k-i}|\right)^2\chi_k^N. \tag{32}$$

Writing the Cauchy–Buniakowsky inequality

$$\left(\sum_{i=0}^{r-1} q^i|\varepsilon_{k-i}|\right)^2 \le \sum_{j=0}^{r-1} q^j \sum_{i=0}^{r-1} q^i|\varepsilon_{k-i}|^2,$$

we get for an arbitrary $L > 0$

$$\begin{aligned}&\mathrm{E}\left(\sum_{i=0}^{r-1} q^i|\varepsilon_{k-i}|\right)^2\chi_k^N \\ &\quad\le (1-q)^{-1}\left(\mathrm{E}\sum_{i=0}^{r-1} q^i|\varepsilon_{k-i}|^2 I\{|\varepsilon_{k-i}|>L\} + L^2\mathrm{P}\{|\xi_k|>N\}\sum_{i=0}^{r-1} q^i\right).\end{aligned} \tag{33}$$

Lemma 1 together with (8) and (10) implies that

$$\lim_{N\to\infty}\overline{\lim_{n\to\infty}}\frac{1}{n}\sum_{k=0}^{n}\mathrm{P}\{|\xi_k|>N\} = 0. \tag{34}$$

Obviously, for arbitrary nonnegative numbers $u_0,\dots,u_{r-1},v_1,\dots,v_{n-1}$,

$$\sum_{k=r}^{n-1}\sum_{i=0}^{r-1} u_i v_{k-i} \le \sum_{i=0}^{r-1} u_i \sum_{j=1}^{n-1} v_j,$$

so conditions (8) and (9) imply that

$$\lim_{L\to\infty}\sup_r\overline{\lim_{n\to\infty}}\frac{1}{n}\sum_{k=r}^{n-1}\mathrm{E}\sum_{i=0}^{r-1} q^i|\varepsilon_{k-i}|^2 I\{|\varepsilon_{k-i}|>L\} = 0,$$

whence, in view of (32) – (34),

$$\lim_{N\to\infty}\sup_r\overline{\lim_{n\to\infty}}\frac{1}{n}\sum_{k=r}^{n-1}\mathrm{E}\|f(\xi_k^r)\otimes\eta_k^r\|\chi_k^N = 0.$$

Combining this with (30) and (31), we arrive at (28).

The proof of (29) is similar.

Corollary 5. *Let the conditions of Lemma 4 be fulfilled and for any $r\in\mathbb{N}$ there exist a pair (Q^r, G^r) of tensors such that*

$$(Q_n^r, G_n^r)\xrightarrow{\mathrm{d}}(Q^r, G^r) \quad \text{as} \quad n\to\infty.$$

Then the sequence $((Q^r, G^r), r\in\mathbb{N})$ converges in distribution to some limit (Q, G) and relation (18) holds.

Lemma 5. *Let the sequence* (ε_k) *satisfy conditions (9) and (10) and for any uniformly bounded sequence* (α_k) *of Borel functions on* $\mathbb{R}^{rd}$

$$\frac{1}{n}\sum_{k=r}^{n-1}\big(\alpha_k(\varepsilon_{k-r+1},\dots,\varepsilon_k)-\mathrm{E}^0\alpha_k(\varepsilon_{k-r+1},\dots,\varepsilon_k)\big)\xrightarrow{\mathrm{P}}\mathrm{O}. \tag{36}$$

Then this relation holds for any sequence (α_k) *of Borel functions such that*

$$|\alpha_k(x_1,\dots,x_r)|\le C\left(\sum_{i=1}^r|x_i|^2+1\right). \tag{37}$$

Proof. Denote $\zeta_k=\alpha_k(\varepsilon_{k-r+1},\dots,\varepsilon_k)$. Then for any $N>0$

$$\frac{1}{n}\sum_{k=r}^{n-1}\big(\zeta_k I\{|\varepsilon_{k-r+1}|\le N,\dots,|\varepsilon_k|\le N\}-\mathrm{E}^0\zeta_k I\{|\varepsilon_{k-r+1}|\le N,\dots,|\varepsilon_k|\le N\}\big)\xrightarrow{\mathrm{P}}0,$$

so it suffices to prove that, for $j=0,\dots,r-1$,

$$\lim_{N\to\infty}\overline{\lim_{n\to\infty}}\frac{1}{n}\sum_{k=1}^n\mathrm{E}^0|\zeta_k|I\{|\varepsilon_{k-j}|>N\}=0\quad\text{a.s.} \tag{38}$$

By assumption,

$$\mathrm{E}^0|\zeta_k|I\{|\varepsilon_{k-j}|>N\}\le C\left(\mathrm{P}\{|\varepsilon_{k-j}|>N|\mathcal{F}_0\}+\sum_{i=0}^{k-1}\mathrm{E}^0|\varepsilon_{k-i}|^2I\{|\varepsilon_{k-j}|>N\}\right). \tag{39}$$

Due to (10), $\mathrm{P}\{|\varepsilon_i|>N|\mathcal{F}_0\}\le\upsilon^2N^{-2}$ and $\mathrm{E}^0|\varepsilon_{k-i}|^2I\{|\varepsilon_{k-j}|>N\}\le\upsilon^4N^{-2}$ as $i\ne j$, which together with (39) and (9) implies (38).

Remark. Obviously, if relation (36) holds for any sequence of $\mathbb{R}$-valued functions (uniformly bounded or satisfying (37)), then for any $m\in\mathbb{N}$ it is valid for any sequence of $\mathbb{R}^m$-valued functions with the same property.

The proof of the following statement is similar.

Lemma 6. *Let the sequence* (ε_k) *satisfy conditions (9) and (10) and for any uniformly bounded sequence* (α_k) *of* $\mathbb{R}$*-valued Borel functions on* $\mathbb{R}^{rd}$

$$\frac{1}{n}\sum_{k=r}^{n-1}\big(\sigma_k^2\otimes\alpha_k(\varepsilon_{k-r+1},\dots,\varepsilon_k)-\mathrm{E}^0(\sigma_k^2\otimes\alpha_k(\varepsilon_{k-r+1},\dots,\varepsilon_k))\big)\xrightarrow{\mathrm{P}}\mathrm{O} \tag{40}$$

(here $\otimes$ *signifies the multiplication of a tensor by a real number). Then this relation holds for any sequence* (α_k) *of tensor-valued functions satisfying (37) (with* $\|\cdot\|$ *instead of* $|\cdot|$ *on the left-hand side).*

Corollary 6. *Let the conditions of Lemmas 4 and 5 be fulfilled and for any uniformly bounded sequence* (α_k) *of* $\mathbb{R}$*-valued Borel functions on* $\mathbb{R}^{rd}$ *the sequence*

$$\left(\frac{1}{n}\sum_{k=r}^{n-1}\mathrm{E}^0\alpha_k(\varepsilon_{k-r+1},\dots,\varepsilon_k),\quad n=r,r+1,\dots\right)$$

converge in probability. Then the sequence $(Q_n^r,\quad n=r,r+1\dots)$ *converges in probability.*

Corollary 7. *Let the conditions of Lemmas 4 and 6 be fulfilled and for any uniformly bounded sequence of* $\mathbb{R}$*-valued functions the sequence*

$$\left(\frac{1}{n}\sum_{k=r}^{n-1} \mathrm{E}^0\sigma_k^2\alpha_k(\varepsilon_{k-r+1},\dots,\varepsilon_k), \quad n=r,r+1,\dots\right)$$

converge in probability. Then the sequence $(G_n^r, \quad n=r,r+1\dots)$ *converges in probability.*

An example

Suppose that conditions (5), (8), (11), (23) and (24) are fulfilled. Let also $\varepsilon_k=\gamma_k\chi_k$, where (γ_k) and (χ_k) are independent sequences of random variables and i.i.d. random vectors, respectively, $|\gamma_k|\le C$, and let for any $r\in\mathbb{N}$ and bounded Borel function g the sequence

$$\left(\frac{1}{n}\sum_{k=r}^{n-1} g(\gamma_{k-r+1},\dots,\gamma_k), \quad n=r,r+1,\dots\right)$$

converge in probability; $\mathrm{E}\chi_1=0$, $\mathrm{E}\chi_1^{\otimes 2}=\mathrm{I}$.

For $\mathcal{F}_k$, we take the σ-algebra generated by $\xi_0;\chi_1,\dots,\chi_k;\gamma_1,\gamma_2,\dots$ (so that the whole sequence (γ_k) is $\mathcal{F}_0$-measurable). Then $\sigma_k^2=\gamma_k^2\mathrm{I}$,

$$G_n^r=\mathrm{I}\otimes\frac{1}{n}\sum_{k=r}^{n}(\gamma_k\eta_k^r)^{\otimes 2} \tag{41}$$

and conditions (2), (3), and (10) are fulfilled. So is (9), because the γ_k's are uniformly bounded and χ_k's are identically distributed.

To deduce (19) from Theorem 2 and Corollary 5 it suffices to verify the conditions of Corollaries 6 and 7. In view of (41) and the expressions for Q_n^r and η_k^r, we may confine ourselves with the case $\alpha_k=\alpha$.

By the Stone – Weierstrass theorem, α can be approximated uniformly on compacta with finite linear combinations of functions of the kind $h(y)h_1(x_1)\dots h_r(x_r)$ $(y\in\mathbb{R}^r, x_j\in\mathbb{R}^d)$. By the choice of $\mathcal{F}_k$ and the assumptions on (γ_k) and (χ_k),

$$\mathrm{E}^0 h(\overline{\gamma}_k)h_1(\chi_{k-r+1})\dots h_r(\chi_k)=h(\overline{\gamma}_k)\prod_{i=1}^{r}\mathrm{E}h_i(\chi_1),$$

where $\overline{\gamma}_k=(\gamma_{k-r+1},\dots,\gamma_k)$.

Hence and from the Chebyshev's inequality, (36) emerges. The last condition of Corollary 6 follows from (41) and the above assumption on (γ_k).

If $\det Q\ne 0$, then Theorem 2 asserts (19). If herein $\mathrm{l.i.p.}_{n\to\infty}\frac{1}{n}\sum_{k=r}^{n} g(\overline{\gamma}_k)$ is random, then the limiting distribution will not be Gaussian.

Bibliography

1. A.Ya. Dorogovtsev, *Estimation theory for parameters of random processes*, Kyiv University Press, Kyiv, 1982. (Russian)
2. A.P. Yurachkivsky, D.O. Ivanenko, *Matrix parameter estimation in an autoregression model with non-stationary noise*, Th. Prob. Math. Stat. **72** (2005), 158–172. (Ukrainian)
3. A.P. Yurachkivsky, *Conditions for convergence of a sequence of martingales in terms of their quadratic characteristics*, Reports of the Nat. Ac. Sc. Ukr. **1** (2003), 33–36.
4. R.Sh. Liptser, A.N. Shiryaev, *Theory of Martingales*, Nauka, Moscow, 1986. (Russian)

Kyiv University, 64 Volodymyrska Str.
E-mail: yap@univ.kiev.ua, ida@univ.kiev.ua