Theory of Stochastic Processes
Vol.13 (29), no.1-2, 2007, pp.77-85

DMYTRO IVANENKO

ASYMPTOTICALLY OPTIMAL ESTIMATOR OF THE PARAMETER OF SEMI-LINEAR AUTOREGRESSION

The difference equations $\xi_k = af(\xi_{k-1}) + \varepsilon_k$, where (ε_k) is a square integrable difference martingale, and the differential equation $\mathrm{d}\xi = -af(\xi)\mathrm{d}t + \mathrm{d}\eta$, where η is a square integrable martingale, are considered. A family of estimators depending, besides the sample size n (or the observation period, if time is continuous) on some random Lipschitz functions is constructed. Asymptotic optimality of this estimators is investigated.

1. Introduction

Discrete time

We consider the difference equation

$$\xi_k = af(\xi_{k-1}) + \epsilon_k, \quad k \in \mathrm{N}, \tag{1}$$

where ξ_0 is a prescribed random variable, f is a prescribed nonrandom function, a is an unknown scalar parameter and (ϵ_k) is a square integrable difference martingale with respect to some flow $(\mathrm{F}_k, k \in \mathrm{Z}_+)$ of σ-algebras such that the random variable ξ_0 is F_0-measurable. In the detailed form, the assumption about (ϵ_k) means that for any k ϵ_k is F_k-measurable,

$$\mathrm{E}\epsilon_k^2 < \infty \tag{2}$$

and

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

The word "semi-linear" in the title means that the right-hand side of (1) depends linearly on a but not on ξ_{k-1}.

2000 *Mathematics Subject Classifications.* Primary 62F12. Secondary 60F05.
Key words and phrases. Martingale, estimator, optimization, convergence.

We use the notation: l.i.p. – limit in probability; $\xrightarrow{\mathrm{d}}$ – the weak convergence of finite-dimensional distributions of random functions, in particular convergence in distribution of random variables.

Let for each $k \in \mathrm{Z}_+$ $h_k = h_k(\omega, x)$ be an $\mathrm{F}_{k-1} \otimes \mathrm{B}$-measurable function (that is the sequence (h_k) be predictable) such that

$$\mathrm{E}\left[\left(|\xi_{k+1}| + |af(\xi_{k+1})|\right)|h_k(\xi_k)|\right] + \mathrm{E}|h_k(\xi_k)| < \infty.$$

Then from (1) – (3) we have $\mathrm{E}\left(\xi_{k+1} - af(\xi_k)\right)h_k(\xi_k) = 0$, whence

$$a = \left(\mathrm{E}\xi_{k+1}h_k(\xi_k)\right)\left(\mathrm{E}f(\xi_k)h_k(\xi_k)\right)^{-1}$$

provided $(\mathrm{E}f(\xi_k)h_k(\xi_k)) \neq 0$. This prompts the estimator

$$\breve{a}_n = \left(\sum_{k=0}^{n-1} \xi_{k+1}h_k(\xi_k)\right)\left(\sum_{k=0}^{n-1} f(\xi_k)h_k(\xi_k)\right)^{-1}, \tag{4}$$

coinciding with the LSE if $h_k(x) = f(x)$ for all k.

Continuous time

We consider the differential equation

$$\mathrm{d}\xi(t) = -af(\xi(t))\mathrm{d}t + \mathrm{d}\eta(t), \quad t \in \mathrm{R}, \tag{5}$$

where $\eta(t)$ is a local square integrable martingale w.r.t. a flow $(\mathrm{F}(t))$ such that the random variable $\xi(0)$ is $\mathrm{F}(0)$-measurable.

Let $h(t, x)$ be a predictable random function such that for all $t \in \mathrm{R}_+$

$$\mathrm{E}\left[\left(|\xi(t)| + |af(\xi(t))|\right)|h(t, \xi(t))|\right] + \mathrm{E}|h(t, \xi(t))| < \infty.$$

Let us multiply (5) on $h(t, \xi(t))$ and integrate from 0 to T. The same rationale as in the discrete case yields the estimator

$$\breve{a}_T = -\left(\int_0^T h(t, \xi(t))\mathrm{d}\xi\right)\left(\int_0^T f(\xi(t))h(t, \xi(t))\mathrm{d}t\right)^{-1}, \tag{6}$$

coinciding with the LSE if $h(t, x) = f(x)$.

Asymptotic normality of $\sqrt{n}\left(\breve{A}_n - A\right)$, where $\breve{A}_n$ is the LSE of a matrix parameter A, was proved in [1] under the assumptions of ergodicity and stationarity of (ξ_n). Convergence in distribution of this normalized deviation was proved in [2] with the use of stochastic calculus. Ergodicity and even stationarity of (ϵ_k) was not assumed in [2], so the limiting distribution could be other than normal.

The goal of the article is to match a sequence (h_k) (if time is discrete) or a function $h(t, \cdot)$ (if time is continuous) so that to minimize the value of some random functional V_n which, as we shall see in Section 3, is asymptotical close in distribution to some numeral characteristic of the estimator (in case the latter is asymptotically normal this characteristic coincides with the variance).

2. The main results

Discrete time

Denote $\sigma_k^2 = \mathrm{E}[\epsilon_{\mathrm{k}}^2|\mathrm{F}_{\mathrm{k}-1}]$, $\mu_k = h_k(\xi_k)$. Let $\mathrm{Lip}(C)$ denote the class of functions satisfying the Lipschitz condition with some constant C and equal to zero at the origin, $\mathrm{Lip} = \bigcup_{C>0}\mathrm{Lip}(C)$, and let $\mathbf{H}(C)$ denote the class of all predictable random functions on $\mathrm{Z}_+ \times \mathrm{R}$ (discrete time) or $\mathrm{R}_+ \times \mathrm{R}$ (continuous time) whose realizations $h_k(\cdot)$ (respectively $h(t,\cdot)$) belong, as functions of x, to $\mathrm{Lip}(C)$, $\mathbf{H} = \bigcup_{C>0}\mathbf{H}(C)$. Predictability means $\mathrm{P} \otimes \mathrm{B}$-measurability in (ω, t, x) (the σ-algebra P is defined in [4, p. 28], [6, p. 13]).

We are seeking for $(\widetilde{h}_k) \in \mathbf{H}$ minimizing the functional

$$V_n(h_0, \ldots, h_{n-1}) = \frac{\frac{1}{n}\sum_{k=0}^{n-1} \sigma_{k+1}^2 \mu_k^2}{\left(\frac{1}{n}\sum_{k=0}^{n-1} f(\xi_k)\mu_k\right)^2}. \tag{7}$$

Theorem 1. *Let*

$$V_n(\widetilde{h}_0, \ldots, \widetilde{h}_{n-1}) = \min_{h_0,\ldots,h_{n-1}\in\mathbf{H}} V_n(h_0, \ldots, h_{n-1}). \tag{8}$$

Then

$$\sigma_{k+1}^2\widetilde{\mu}_k \sum_{i=0}^{n-1} f(\xi_i)\widetilde{\mu}_i = f(\xi_k)\sum_{i=0}^{n-1}\sigma_{i+1}^2\widetilde{\mu}_i^2, \qquad k = \overline{0, n-1}. \tag{9}$$

Proof. To obtain the necessary conditions for extremum of the functional V_n (9) we will vary [3] just one of functions h_k, $k = \overline{0, n-1}$, leaving the other functions without changes. Thus regarding $V_n(h_0, \ldots, h_{n-1})$ as a functional depending on only one function $V_n(h_0, \ldots, h_{n-1}) = \widetilde{V}_n(h_k)$.

Let's choose some scalar function $g \in \mathbf{H}$ and denote $g_\lambda(x) = \widetilde{h}_k(x) + \lambda(g(x) - \widetilde{h}_k(x))$, $v(\lambda) = \widetilde{V}_n(g_\lambda)$.

Obviously, $g_\lambda \in \mathbf{H}$ so the minimum of $v(\lambda)$ is attained at zero and therefore

$$v'(0) = 0. \tag{10}$$

The expression for the left-hand side is

$$v'(0) = \frac{2n(g(\xi_k) - \mu_k)\left(\sigma_{k+1}^2\widetilde{\mu}_k(\sum_{i=0}^{n-1} f(\xi_i)\widetilde{\mu}_i - f(\xi_k)\sum_{i=0}^{n-1}\sigma_{i+1}^2\widetilde{\mu}_i^2\right)}{\left(\sum_{i=0}^{n-1} f(\xi_i)\widetilde{\mu}_i\right)^3}.$$

Hence in view of (10) we obtain the i th equation of system (9).

It remains to apply this argument to each function h_k, $k = \overline{0, n-1}$.

Remark. The Lipschitz condition was not used in the proof. It will be required in Section 3.

Corollary 1. Let $f \in \mathrm{Lip}(C)$ and there exist a constant $q > 0$ such that $\sigma_k^2 \geq q$ for all k. Then $h_i(x) = f(x)/\sigma_{i+1}^2$, $i = \overline{0, n-1}$, is a solution to the problem (8).

Continuous time

Let m denote the quadratic characteristic of η.

We shall match $\widetilde{h} = \widetilde{h}(\omega, t, x)$ from $\mathbf{H}(C)$ (C is independent of t) so that to minimize the value of the functional

$$V_T(h) = \frac{\frac{1}{T}\int_0^T h(t,\xi(t))^2 \mathrm{d}m(t)}{\left(\frac{1}{T}\int_0^T f(\xi(t))h(t,\xi(t))\mathrm{d}t\right)^2}. \tag{11}$$

Theorem 2. *Let*

$$V_T(\widetilde{h}) = \min_{h\in\mathbf{H}} V_T(h). \tag{12}$$

Then for all $g \in \mathbf{H}$

$$\begin{aligned} &\textstyle\int_0^T \widetilde{h}(t,\xi(t))g(t,\xi(t))\mathrm{d}m(t) \int_0^T f(\xi(t))\widetilde{h}(t,\xi(t))\mathrm{d}t = \\ &\qquad\qquad \textstyle\int_0^T f(\xi(t))g(t,\xi(t))\mathrm{d}t \int_0^T \widetilde{h}(t,\xi(t))^2\mathrm{d}m(t). \end{aligned} \tag{13}$$

Proof. Let's choose some scalar function $g \in \mathbf{H}$ and denote $g_\lambda(t,x) = \widetilde{h}(t,x) + \lambda g(t,x)$, $v(\lambda) = V_T(g_\lambda)$.

Obviously $g_\lambda(t,\cdot) \in \mathbf{H}$ so the minimum of $v(\lambda)$ is attained in zero and therefore

$$v'(0) = 0. \tag{14}$$

The expression for the left-hand side is

$$v'(0) = 2T\left(\textstyle\int_0^T f(\xi(t))\widetilde{h}(t,\xi(t))\mathrm{d}t\right)^{-3} \times$$

$$\left(\int_0^T f(\xi(t))\widetilde{h}(t,\xi(t))\mathrm{d}t \int_0^T \widetilde{h}(t,\xi(t))g(t,\xi(t))\mathrm{d}m(t) - \right.$$

$$\left. \int_0^T f(\xi(t))g(t,\xi(t))\mathrm{d}t \int_0^T \widetilde{h}(t,\xi(t))^2\mathrm{d}m(t)\right).$$

Hence in view of (14) we come to (13).

Corollary 2. Let $f \in \mathrm{Lip}(C)$, m be absolutely continuous w.r.t. the Lebesgue measure and there exist a constant $q > 0$ such that for all t $\dot{m} \geq q$. Then $h(t,x) = f(x)/\dot{m}$ is a solution to the problem (12).

3. An illustration

Denote $\mathrm{E}^0 = \mathrm{E}(\cdots|\mathrm{F}_0)$, $Q_n = \frac{1}{n}\sum_{k=0}^{n-1} f(\xi_k)\mu_k$, $G_n = \frac{1}{n}\sum_{k=1}^{n}\sigma_k^2\mu_{k-1}^2$.

We denote $\mathrm{E}^0 = \mathrm{E}(\cdots|\mathrm{F}_0)$ and introduce the conditions

CP1. For any $r \in \mathrm{N}$ and any uniformly bounded sequence (α_k) of R-valued Borel functions on R^r

$$\frac{1}{n}\sum_{k=r}^{n-1}\left(\alpha_k(\epsilon_{k-r+1},\ldots,\epsilon_k) - \mathrm{E}^0\alpha_k(\epsilon_{k-r+1},\ldots,\epsilon_k)\right) \xrightarrow{\mathrm{P}} 0,$$

$$\frac{1}{n}\sum_{k=r}^{n-1}\left(\sigma_k^2\alpha_k(\epsilon_{k-r+1},\dots,\epsilon_k)-\mathrm{E}^0\sigma_k^2\alpha_k(\epsilon_{k-r+1},\dots,\epsilon_k)\right)\xrightarrow{\mathrm{P}}0.$$

CP2. For such r and (α_k) the sequences

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

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

converge in probability.

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

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

Then

$$\xi_k=f_r(\xi_{k-r},\epsilon_{k-r+1},\dots,\epsilon_k),\quad r<k.$$

Lemma 1. *Let conditions (2), (3),* **CP1** *and* **CP2** *be fulfilled. Suppose also that*

$$\lim_{N\to\infty}\overline{\lim_{n\to\infty}}\frac{1}{n}\sum_{k=1}^{n}\mathrm{E}\epsilon_k^2 I\{|\epsilon_k|>N\}=0 \tag{15}$$

and there exist an F_0*-measurable random variable* υ *such that for all* k

$$\sigma_k^2\le\upsilon \tag{16}$$

and positive numbers $C,\ C_1$ *such that*

$$|a|C<1, \tag{17}$$

$f\in\mathrm{Lip}(C)$, $(h_k)\in\mathbf{H}(C_1)$. *Then*

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

Proof. Denote $\xi_k^r=f_r(0,\epsilon_{k-r+1},\dots,\epsilon_k)$, $\mu_k^r=h_k(\xi_k^r)$, $Q_n^r=\frac{1}{n}\sum_{k=r}^{n-1}f(\xi_k^r)\mu_k^r$, $G_n^r=\frac{1}{n}\sum_{k=r}^{n}\sigma_k^2(\mu_{k-1}^r)^2$. We claim that conditions (2), (3), (15), (16), (17) and the relation

$$(Q_n^r,G_n^r)\xrightarrow{\mathrm{d}}(Q^r,G^r)\quad\text{as}\quad n\to\infty \tag{19}$$

imply (18).

Let X_r denote $(x_1,\dots,x_r)\in\mathrm{R}^r$. Then under the assumptions on f and h_k for any $N>0$

$$\lim_{r\to\infty}\sup_{|x|\le N,X_r\in\mathrm{R}^r}|f_r(x,X_r)-f_r(0,X_r)|=0,$$

whence with probability 1 for any k

$$\lim_{r\to\infty}\sup_{|x|\le N, X_r\in \mathrm{R}^r}|f(f_r(x,X_r))h_k(f_r(x,X_r))-f(f_r(0,X_r))h_k(f_r(0,X_r))|=0,$$
$$\lim_{r\to\infty}\sup_{|x|\le N, X_r\in \mathrm{R}^r}|h_k(f_r(x,X_r))^2-h_k(f_r(0,X_r))^2|=0. \tag{20}$$

These relations were proved in [5].

Let us prove that conditions (2), (3), (15), (16) and (17) imply that almost surely

$$\lim_{r\to\infty}\overline{\lim_{n\to\infty}}\mathrm{E}^0|Q_n-Q_n^r|=0, \qquad \lim_{r\to\infty}\overline{\lim_{n\to\infty}}\mathrm{E}^0|G_n-G_n^r|=0. \tag{21}$$

By (20) for any $N>0$

$$\lim_{r\to\infty}\overline{\lim_{n\to\infty}}\frac{1}{n}\sum_{k=r}^{n-1}\mathrm{E}|f(\xi_k)\otimes\mu_k-f(\xi_k^r)\mu_k^r|I\{|\xi_k|\le N\}=0. \tag{22}$$

Denote $\chi_k^N=I\{|\xi_k|>N\}$, $I_k^N=I\{|\epsilon_k|>(1-C)N\}$, $b_k^N=\mathrm{E}^0|\xi_k|^2\chi_k^N$. Due to (17) and because of $(h_k)\in\mathbf{H}(C_1)$

$$\mathrm{E}^0|f(\xi_k)\mu_k|\chi_k^N\le CC_1b_k^N.$$

Hence and from (2), (3), (15)–(17) we get by Corollary 1 [5]

$$\lim_{N\to\infty}\overline{\lim_{n\to\infty}}\frac{1}{n}\sum_{k=0}^{n-1}\mathrm{E}^0|f(\xi_k)\mu_k|\chi_k^N=0. \tag{23}$$

Further, for $k\ge r$,

$$\mathrm{E}^0|f(\xi_k^r)\mu_k^r|=\mathrm{E}^0|f(f_r(0,\epsilon_{k-r+1},\dots,\epsilon_k))||h_k(f_r(0,\epsilon_{k-r+1},\dots,\epsilon_k))|,$$

whence

$$\mathrm{E}|f(\xi_k^r)\mu_k^r|\chi_k^N\le CC_1\mathrm{E}\left(\sum_{i=0}^{r-1}C^i|\epsilon_{k-i}|\right)^2\chi_k^N. \tag{24}$$

Writing the Cauchy – Bunyakovsky inequality

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

we get for an arbitrary $L>0$
$\mathrm{E}\left(\sum_{i=0}^{r-1}C^i|\epsilon_{k-i}|\right)^2\chi_k^N\le$

$$(1-C)^{-1}\left(\mathrm{E}\sum_{i=0}^{r-1}C^i\epsilon_{k-i}^2I\{|\epsilon_{k-i}|>L\}+L^2\mathrm{P}\{|\xi_k|>N\}\sum_{i=0}^{r-1}C^i\right). \tag{25}$$

In view of (2) and (3) Lemma 1 [5] together with (17) and (15) implies that

$$\lim_{N\to\infty}\overline{\lim_{n\to\infty}}\frac{1}{n}\sum_{k=0}^{n}\mathrm{P}\{|\xi_k|>N\}=0. \tag{26}$$

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

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

so conditions (17) and (15) imply that

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

whence in view of (24) – (26)

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

Combining this with (22) and (23), we arrive at the first relation of (21).

The proof of the second relation of (21) is similar.

Condition **CP1** implies that

$$\lim_{r\to\infty}\overline{\lim_{n\to\infty}}\mathrm{E}^0|Q_n^r-\mathrm{E}^0Q_n^r|=0,\qquad\lim_{r\to\infty}\overline{\lim_{n\to\infty}}\mathrm{E}^0|G_n^r-\mathrm{E}^0G_n^r|=0.$$

Under condition **CP2** the sequences $(\mathrm{E}^0G_n^r,\ n\in\mathrm{N})$ and $(\mathrm{E}^0Q_n^r,\ n\in\mathrm{N})$ converge in probability for any $r\in\mathrm{N}$. Thus relation (19) holds.

From (19) and (21) we obtain that the sequence $((Q^r,G^r),r\in\mathrm{N})$ converges in distribution to some limit (Q,G) and relation (18) holds.

By construction $V_n(h_0,\dots,h_{n-1})=G_nQ_n^{-2}$. The value $Q_n=0$ is excluded by the choice of the tuple $(h_0,\dots,h_{n-1})$ minimizing V_n.

Corollary 3. Let the conditions of Lemma 1 be fulfilled and $Q\neq0$ a.s. Then $V_n\xrightarrow{\mathrm{d}}V$, where $V=GQ^{-2}$.

Having in mind the use of stochastic analysis, we introduce the processes $\check{a}_n(t)=\check{a}_{[nt]}$ and the flows $\mathrm{F}_n(t)=\mathrm{F}_{[nt]}$ with continuous time.

Theorem 3. *Let conditions of Lemma 1 be fulfilled. Then* $\sqrt{n}\,(\check{a}_n(\cdot)-a)\xrightarrow{\mathrm{d}}\beta(\cdot)$*, where* β *is a continuous local martingale with quadratic characteristic*

$$\langle\beta\rangle(t)=tV, \tag{27}$$

and initial value 0.

Proof. Denote $Y_n(t)=\frac{1}{\sqrt{n}}\sum_{k=1}^{[nt]}\epsilon_k\mu_{k-1}$. Then because of (4)

$$\sqrt{n}(\check{a}_n(t)-a)=Y_n(t)Q_n^{-1}. \tag{28}$$

By construction and conditions (2), (3), (17) Y_n is a locally square integrable martingale with quadratic characteristic $\langle Y_n\rangle(t) = n^{-1}[nt]G_{[nt]}$.

It was proved in [5] that under conditions (2), (3), (15), (16), (17) and (18) $\sqrt{n}(\breve{a}_n(\cdot) - a) \xrightarrow{\mathrm{d}} Y(\cdot)Q^{-1}$, where Y is a continuous local martingale w.r.t. some flow $(\mathrm{F}(t), t \in \mathrm{R}_+)$ such that $\langle Y\rangle(t) = tG$ and the random variable Q is $\mathrm{F}(0)$-measurable (and so does G, which can be seen from the expression for $\langle Y\rangle$). In view of Lemma 1 it remains to note that $V_n = \langle Y_n\rangle(1)Q_n^{-2}$ and $V = \langle Y\rangle(1)Q^{-2}$.

Remark. This theorem explains the form of functional (7). In the most general case (without conditions **CP1** and **CP2**) the denominator (28) in limit is an $\mathrm{F}(0)$-measurable random variable, and the numerator tends to quadratic characteristic at the point $t = 1$ of the continuous local martingale Y. Thus, the numerator (7) is the quadratic characteristic at $t = 1$ of the pre-limit martingale Y_n, and the denominator satisfies the law of large numbers. Minimizing the pre-limit variance in $(h_k) \in \mathbf{H}(C_1)$, we lessen the value of the limit variance of the normalized deviation of estimator (4).

Let further $h_k(x) = f(x)/\sigma_{k+1}^2$. Recall that $(h_k,\ k = \overline{0, n-1})$ is a solution to the problem (8). For such h_k we have

Corollary 4. Let the conditions of Corollary 1 and Theorem 3 be fulfilled. Then

$$V = \left(\lim_{r\to\infty} \operatorname*{l.i.p.}_{n\to\infty} \frac{1}{n}\sum_{k=r}^{n-1} \mathrm{E}^0 \frac{f(\xi_k^r)^2}{\sigma_{k+1}^2}\right)^{-1}.$$

Proof. Obviously $V_n = Q_n^{-1}$. By Lemma 1 $Q_n \xrightarrow{\mathrm{d}} Q$, where $Q = \lim_{r\to\infty} \operatorname*{l.i.p.}_{n\to\infty} \mathrm{E}^0 Q_n^r$. To complete the proof it remains to note that $Q_n^r = \frac{1}{n}\sum_{k=r}^{n-1} \mathrm{E}^0 \frac{f(\xi_k^r)^2}{\sigma_{k+1}^2}$.

4. An example

Suppose that $f \in \mathrm{Lip}(C)$, $h_k \in \mathbf{H}(C_1)$ condition (17) be fulfilled. Let also $\epsilon_n = \gamma_n b_n(\xi_{n-1})$, where (γ_n) be a sequence of independent random variables with zero mean and variances ς_n^2, $|\gamma_k| \le C_2$, $b_n \in \mathbf{H}(C_3)$ and $C + C_2C_3 < 1$. Let also $\mathrm{E}\xi_0^2 < \infty$. For F_k we take the σ-algebra generated by $\xi_0; \gamma_1, \dots, \gamma_k$. Then $\sigma_k^2 = \varsigma_k^2 b_k(\xi_{k-1})^2$ and (ϵ_n) satisfies (2), (3).

Denote further

$$\widehat{f}_r(x_0, \dots, x_r) = af(\widehat{f}_{r-1}(x_0, \dots, x_{r-1})) + x_r b_r(\widehat{f}_{r-1}(x_0, \dots, x_{r-1})),$$

$$\widehat{\xi}_k^r = \widehat{f}_r(0, \gamma_{k-r+1}, \dots, \gamma_k), \quad \widehat{\mu}_k^r = h_k(\widehat{\xi}_k^r), \quad \widehat{Q}_n^r = \frac{1}{n}\sum_{k=r}^{n-1} f(\widehat{\xi}_k^r)\widehat{\mu}_k^r,$$

$\widehat{G}_n^r = \frac{1}{n}\sum_{k=r}^{n-1} \varsigma_{k+1}^2 b_{k+1}(\widehat{\xi}_k^r)^2(\widehat{\mu}_k^r)^2$. Similarly to the proof of Lemma 1 we obtain

$$\lim_{r\to\infty}\varlimsup_{n\to\infty}\mathrm{E}^0|G_n-\widehat{G}_n^r|=0, \qquad \lim_{r\to\infty}\varlimsup_{n\to\infty}\mathrm{E}^0|Q_n-\widehat{Q}_n^r|=0.$$

Summands in $\widehat{G}_n^r$ and $\widehat{Q}_n^r$ are nonrandom functions of $\gamma_{k-r+1},\dots,\gamma_k$, so they satisfy the law of large numbers in Bernstein's form.

If besides ϵ_n satisfies **CP2** and $Q\neq 0$ a.s. then Theorem 3 asserts (27). If herein $\frac{f(x)}{\varsigma_k^2 b_k(x)^2}\in \mathrm{Lip}$ then $\widetilde{h}_k(x)=\frac{f(x)}{\varsigma_k^2 b_k(x)^2}$ is a solution to the problem (8) and

$$V=\left(\lim_{r\to\infty}\operatorname*{l.i.p.}_{n\to\infty}\frac{1}{n}\sum_{k=r}^{n-1}\mathrm{E}^0\frac{f(\widehat{\xi}_k^r)^2}{\varsigma_{k+1}^2 b_{k+1}(\widehat{\xi}_k^r)^2}\right)^{-1}.$$

Example. Let $b_n=b$, $h_n=h$ and γ_n be i.i.d. random variables. In view of expressions for $\widehat{Q}_n^r$ and $\widehat{G}_n^r$ we may confine ourselves with the case $\alpha_k=\alpha$.

By the Stone – Weierstrass theorem for σ-compact spaces [7, p. 317] α can be uniformly on compacta approximated with finite linear combinations of functions of the kind $g_1(x_1)\dots g_r(x_r)$. By the choice of F_k and the assumptions on (γ_n) $\mathrm{E}^0 g_1(\gamma_{k-r+1})\dots g_r(\gamma_k)=\prod_{i=1}^r \mathrm{E}g_i(\gamma_1)$. Whence condition **CP2** emerges.

Acknowledgement. The author is grateful to A. Yurachkivsky for helpful advices.

Bibliography

1. Dorogovtsev A. Ya., *Estimation theory for parameters of random processes (Russian).* Kyiv University Press. Kyiv (1982).
2. Yurachkivsky A. P., Ivanenko D. O., *Matrix parameter estimation in an autoregression model with non-stationary noise (Ukranian),* Th. Prob. Math. Stat. 72 (2005), 158–172.
3. Elsholz, L. E., *Differential equations and calculus of variations (Russian),* Nauka, Moscow (1969).
4. Chung K. L., Williams R. J., *Introduction to stochastic integration (Russian),* Mir, Moscow (1987).
5. Yurachkivsky A. P., Ivanenko D. O., *Matrix parameter estimation in an autoregression model,* Theory of Stochastic Processes **12(28)** No 1-2 (2006), 154-161.
6. Liptser R. Sh., Shiryaev A. N., *Theory of martingales (Russian),* Nauka, Moscow (1986).
7. Kelley, J., *General topology (Russian).* Nauka, Moscow (1981).

Department of Mathematics and Theoretical Radiophysic, Kyiv National Taras Shevchenko University, Kyiv, Ukraine
E-mail address: ida@univ.kiev.ua