

THE ASYMPTOTIC PROPERTIES OF ESTIMATES OF THE PARAMETERS OF NONLINEAR TIME SERIES

V. V. Anisimov^a and Kh. S. Keibakh^b

UDC 519.21

Asymptotic properties of nonlinear time series parameter estimators constructed on trajectories of stochastic systems under stationary and transient conditions are studied with the use of the least-squares method. The investigation method is based on the study of asymptotic properties of extremal sets of random functions.

Keywords: parameter estimation, asymptotic properties, nonlinear time series.

INTRODUCTION

In real models, data, as a rule, arise in observations in trajectories of stochastic systems and are basically interdependent and nonstationary in time. Therefore, generally, methods of classical data analysis are not applicable to the analysis of properties of estimates in such situations.

In many models, estimates can be presented as points (sets) of extrema of random functions that are constructed as additive functionals on the trajectories of observed systems.

Let us consider a rather general model of a time series for which data are constructed from observations in a trajectory of a stochastic system.

Let $S(t)$, $t \geq 0$, be a trajectory of some (random or determinate) system with values in a space X , and $t_1 < t_2 < \dots$ be the instants of observations on the interval $[0, T]$. They can be some determinate or random instants, for example, instants of changes of the surroundings, mode switchings, etc.

Denote $s_k = S(t_k + 0)$, $k \geq 0$. Assume that the following quantities are observed:

$$z_k = g(\theta_0, s_k) + \varepsilon_k, \quad 0 \leq k \leq \nu(T),$$

where the function $g(\theta, s)$ is given, θ_0 is an unknown parameter, the quantities $\varepsilon_k = \varepsilon(s_k)$ are random noise, $\mathbf{E}[\varepsilon_k / s_k] = 0$, $\mathbf{E}[\varepsilon_k \varepsilon_k^* / s_k] < \infty$, $k \geq 0$, and $\nu(T)$ specifies the number of observations on the interval $[0, T]$.

Put

$$F(\theta, T) = T^{-1} \sum_{k=1}^{\nu(T)} (z_k - g(\theta, s_k))^2. \quad (1)$$

Then the least-squares method estimate of an unknown parameter is the set of minima of the function $F(\theta, T)$ in θ

$$\{\theta_T\} = \arg \min_{\theta} F(\theta, T).$$

This example shows that an analysis of the problem of asymptotic study of observation-parameter estimates constructed in the trajectory of a stochastic system necessitates studies of the following classes of problems:

- (i) study of asymptotic properties of extremal sets of random functions;
- (ii) asymptotic analysis of additive functionals of special types in the trajectories of stochastic systems.

^aTaras Shevchenko University, Kiev, Ukraine. ^bBilkent University, Ankara, Turkey. Translated from *Kibernetika i Sistemnyi Analiz*, No. 2, pp. 62-72, March-April, 2000. Original article submitted February 1, 1999.

The results in the theory of asymptotic estimation are mainly devoted to the analysis of independent observations [1] or are based on martingale techniques [2, 3]. Various estimates of parameters of random processes were obtained in [4] using constructive approaches. A number of estimates of parameters in trajectories of random processes satisfying averaging-type conditions are presented in [5–9].

A new approach to statistical-parameter estimation from observations in trajectories of random processes is proposed in the present study. The essence of this approach is as follows: an estimate is represented by a point (set) of the extremum of some additive functional in the trajectory of a stochastic system, and then, using the solution of both problems, the behavior of the desired estimate is investigated.

Here, this approach is realized in analyzing nonlinear estimates of the parameters of nonlinear time series with dependent observations constructed in the trajectory of some stochastic system that satisfies certain general averaging-type assumptions. Note that the properties of nonlinear estimates of the least-squares method (consistency and asymptotic normality) were investigated in [10] in the case of independent observations.

Some results in this subject area were obtained in [2, 7] using the asymptotic properties of solutions of stochastic equations.

EXTREMAL PROBLEMS FOR RANDOM FUNCTIONS

First, let us present general results on the asymptotic behavior of extremal sets of random functions (see [11]). These results are used hereafter for analysis of the behavior of estimates.

Let for each $n \geq 0$ $F_n(\theta)$, $\theta \in \Theta \subset \mathcal{R}^r$ be a random function with values in R , Θ be a bounded closed set, and n be the parameter of a series.

Let us define $\underline{F}_n(\theta)$ as $\underline{F}_n(\theta) = \liminf_{\theta' \rightarrow \theta} F_n(\theta')$ for any function $F_n(\theta)$. If the function $F_n(\theta)$ is random, then the limit is being determined for each realization of $F_n(\theta)$. Assume that $\{\theta_n\} = \arg \min_{\theta \in \Theta} \underline{F}_n(\theta)$. Here, $\{\theta_n\}$ is the set of global minima of the function $F_n(\theta)$.

Let us study the conditions of convergence of $\{\theta_n\}$ when the sequence of functions $F_n(\theta)$ converges in some sense to a limiting (random or nonrandom) function $F_0(\theta)$ as $n \rightarrow \infty$, and let us also study the conditions of weak convergence of the sequence of vectors $v_n(\tilde{\theta}_n - \theta_0)$, where $\tilde{\theta}_n$ is a subsequence of local minima $F_n(\theta)$, and v_n is a normalizing factor.

For an arbitrary function $f(\theta)$, $\theta \in \Theta$, we introduce the modulus of continuity

$$\Delta_U(c, f(\cdot)) = \sup_{|\theta_1 - \theta_2| < c, \theta_1 \in \Theta, \theta_2 \in \Theta} |f(\theta_1) - f(\theta_2)|.$$

Let us formulate some necessary definitions.

Definition 1. Let G_n be a sequence of random sets in Θ . We say that the sequence G_n converges in probability to some (random or nonrandom) point g_0 if $\rho(g_0, G_n) \xrightarrow{P} 0$, where $\rho(g, G) = \sup_{z \in G} \|z - g\|$.

Denote this convergence by $G_n \xrightarrow{P} g_0$.

Definition 2. Let G_n be a sequence of random sets in Θ . We say that the sequence G_n weakly converges to some random variable γ_0 if g_n weakly converges to γ_0 for any sequence g_n such that $P\{g_n \in G_n\} = 1$.

Denote this convergence by $G_n \xrightarrow{w} \gamma_0$.

Definition 3. Let us assume that the sequence of functions $F_n(\theta)$ U -converges to a random or nonrandom function $F_0(\theta)$ in a set Θ if

(i) for any $k = 1, 2, \dots$ and for any $\theta_1, \theta_2, \dots, \theta_k \in \Theta$, the distribution of the vector $(F_n(\theta_i), i = \overline{1, k})$ weakly converges to that of the vector $(F_0(\theta_i), i = \overline{1, k})$;

(ii) for any $\varepsilon > 0$ $\lim_{c \rightarrow +0} \limsup_{n \rightarrow \infty} P(\Delta_U(c, F_n(\cdot)) > \varepsilon) = 0$.

Now, in accordance with [11], we present two theorems on convergence of a sequence of the sets $\{\theta_n\}$.

THEOREM 1. Let $F_n(\theta)$ be a sequence of random functions, and the following conditions hold:

(i) there exists a continuous random function $F_0(\theta)$ such that $F_n(\theta)$ U -converges to $F_0(\theta)$;

(ii) the following separability condition holds: $F_0(\theta_0) < F_0(\theta')$ with unit probability for any random variable θ' given on the same probability space and such that $\theta' \neq \theta_0$ with unit probability, where $\theta_0 = \arg \min_{\theta \in \Theta} F_0(\theta)$.

Then $\{\theta_n\} \xrightarrow{w} \theta_0$.

Further, let us study the behavior of the normalized deviation for the quantity $\{\theta_n\}$. Let us introduce the random function $A_n(z) = v_n^\alpha (F_n(\theta_0 + \frac{1}{v_n}z) - F_n(\theta_0))$ as a function of a new argument $z \in \mathcal{R}^r$.

THEOREM 2. Let the conditions of Theorem 1 be satisfied, and there be a nonrandom sequence $v_n \rightarrow \infty$ and $\alpha > 0$ such that for any $L > 0$ the sequence of functions $A_n(z)$ U -converges to some random function $A_0(z)$ in the domain $\|z\| \leq L$. Assume also that the point $\kappa_0 = \arg \min_z A_0(z)$ is a proper random variable and satisfies the separability condition with unit probability.

Then there exists a subsequence of local minima $\tilde{\theta}_n$ of the function $F_n(\theta)$ such that $v_n(\tilde{\theta}_n - \theta_0) \xrightarrow{w} \kappa_0$.

Example 1. Let the function $A_0(z)$ have the form $A_0(z) = \eta + (\mathcal{X}(a, B^2), z) + Cz, z)$, where η is an arbitrary random variable, the vector $\mathcal{X}(a, B^2)$ has a normal distribution with mean a and covariance matrix B^2 , and C is some matrix such that the matrix $C + C^*$ is invertible.

Then $\kappa_0 = -(C + C^*)^{-1} \mathcal{X}(a, B^2)$.

THE ASYMPTOTIC PROPERTIES OF ESTIMATES OF THE LEAST-SQUARES METHOD

Let us consider applications of Theorems 1, 2 in the analysis of the asymptotic behavior of estimates of the least-squares method constructed from observations in the trajectory of some stochastic system.

Homogeneous Case. We will first consider the homogeneous case as an illustration of the technique of the analysis. Let $g(\theta)$, $\theta \in \Theta \subset \mathcal{R}^r$, be some vector-valued function with values in \mathcal{R}^m , and ξ_1, ξ_2, \dots be independent equally distributed random vectors in \mathcal{R}^m such that

$$\mathbf{E} \xi_1 = 0, \quad \mathbf{E} \xi_1 \xi_1^* = G^2. \quad (2)$$

Assume that we have the quantities

$$z_k = g(\theta_0) + \xi_k, \quad 0 \leq k \leq n. \quad (3)$$

Let us examine the asymptotic properties of estimates of the least-squares method. Denote

$$F_n(\theta) = \frac{1}{n} \sum_{k=1}^n (z_k - g(\theta))^2. \quad (4)$$

As is well known, an estimate of the least-squares method is determined by the relation $\{\theta_n\} = \arg \min_{\theta} F_n(\theta)$.

THEOREM 3. Let a function $g(\theta)$ be continuous and bounded in the domain Θ and

$$g(\theta) \neq g(\theta_0) \text{ for } \theta \neq \theta_0. \quad (5)$$

Then

$$\{\theta_n\} \xrightarrow{P} \theta_0. \quad (6)$$

Further, assume that there exists $\beta > 0$ such that as $h \rightarrow 0$ it is uniform in each bounded domain $\|z\| \leq L$

$$h^{-\beta} (g(\theta_0 + hz) - g(\theta_0)) \rightarrow a(z), \quad (7)$$

where the function $a(z)$ is continuous, and the equation

$$a(z) = y \quad (8)$$

has a unique solution for any $y \in \mathcal{R}^m$.

Then there exists a sequence of local minima $\tilde{\theta}_n$ of the function $F_n(\theta)$ such that

$$n^{1/2\beta}(\tilde{\theta}_n - \theta_0) \overset{w}{\Rightarrow} a^{-1}(\mathcal{N}(0, G^2)), \quad (9)$$

where $a^{-1}(z)$ is an inverse function, and the vector $\mathcal{N}(0, G^2)$ has a multidimensional normal distribution with mean 0 and covariance matrix G^2 .

Proof. Using (3), we present the function $F_n(\theta)$ in the form

$$F_n(\theta) = \|g(\theta) - g(\theta_0)\|^2 - \frac{2}{n} \sum_{k=0}^n (g(\theta) - g(\theta_0), \xi_k) + \frac{1}{n} \sum_{k=0}^n \|\xi_k\|^2. \quad (10)$$

According to the law of large numbers, we have

$$\frac{1}{n} \sum_{k=0}^n \xi_k \xrightarrow{P} 0, \quad \frac{1}{n} \sum_{k=0}^n \|\xi_k\|^2 \xrightarrow{P} \mathbf{E}\|\xi_1\|^2. \quad (11)$$

Since the function $g(\theta)$ is bounded, it follows from relations (10), (11), that the sequence of random functions $F_n(\theta)$ U -converges to the function $\|g(\theta) - g(\theta_0)\|^2$.

Condition (5) is the separability condition for this function; therefore, the first part of the theorem follows from Theorem 1.

Further, we use Theorem 2. Put $v_n = n^{1/2\beta}$ and $\alpha = 2\beta$. Then the function $A_n(z)$ can be presented as

$$A_n(z) = v_n^{2\beta} \|g(\theta_0 + \frac{1}{v_n}z) - g(\theta_0)\|^2 - \frac{2}{n} v_n^{2\beta} \sum_{k=0}^n (g(\theta_0 + \frac{1}{v_n}z) - g(\theta_0), \xi_k). \quad (12)$$

According to (7), the first term converges to the quantity $\|a(z)\|^2$ uniformly in each bounded domain $\|z\| \leq L$, and for $v_n^\beta = \sqrt{n}$ the second term has the form

$$-2(v_n^\beta (g(\theta_0 + \frac{1}{v_n}z) - g(\theta_0))), \frac{1}{\sqrt{n}} \sum_{k=0}^n \xi_k).$$

According to the central limiting theorem and condition (7), this term, as the function of z , U -converges in each bounded domain $\|z\| \leq L$ to the function $-2(a(z), \mathcal{N}(0, G^2))$.

Finally, the sequence of the functions $A_n(z)$ U -converges to the function $q(z) = \|a(z)\|^2 - 2(a(z), \mathcal{N}(0, G^2))$ in each bounded domain $\|z\| \leq L$. Since for each y , there is the representation $\|a\|^2 - 2(a, y) = \|a - y\|^2 - \|y\|^2$, and this function has a minimum for $a = y$, the function $q(z)$ has a minimum when $a(z) = \mathcal{N}(0, G^2)$, i.e., $z = a^{-1}(\mathcal{N}(0, G^2))$. This finally proves Theorem 3.

Let us consider some examples.

Example 2. Let $\Theta = [a, b]$, $\theta_0 \in (a, b)$,

$$g(\theta) = \begin{cases} \alpha_1(\theta - \theta_0), & \text{if } \theta_0 < \theta \leq b, \\ \alpha_2(\theta - \theta_0), & \text{if } a \leq \theta \leq \theta_0, \end{cases}$$

where $\alpha_1, \alpha_2 > 0$. Let also the quantities ξ_1, ξ_2, \dots take values in \mathcal{R} , $\mathbf{E}\xi_1 = 0$, $\mathbf{E}\xi_1^2 = \sigma^2$.

Let us consider the same model of observations (3).

As can be easily seen, in our case condition (5) is satisfied, whence relation (6) follows.

Further, note that despite the fact that the function $g(\theta)$ is nondifferentiable, relation (7) is satisfied for $\beta=1$ with the function

$$a(z) = \begin{cases} \alpha_1 z & \text{for } z > 0, \\ \alpha_2 z & \text{for } z < 0. \end{cases}$$

Clearly, a solution of Eq. (8) exists and is unique.

Then $\sqrt{n}(\tilde{\theta}_n - \theta_0) \xrightarrow{w} \zeta$, where $\zeta = \alpha_1 \chi(\chi(0, \sigma^2) > 0) + \alpha_2 \chi(\chi(0, \sigma^2) < 0)$. In this example, there is a nonclassical limiting law even in the case of a homogeneous model of observations.

INHOMOGENEOUS CASE OF TRAJECTORY OBSERVATIONS

Let us consider a more general model of observations. Assume that a random or nonrandom sequence x_{nk} , $k \geq 0$, is given with values in X , which corresponds to the sequence of states of a system. Let a family of functions $g(\theta, x)$, $\theta \in \Theta$, $x \in X$, be given, with values in \mathcal{R}^m and families of random vectors $\{\xi_k(x)$, $x \in X\}$, $k \geq 0$, independent in the aggregate and of the sequence x_{nk} , with values in \mathcal{R}^m .

Assume for the sake of simplicity that the distributions of the quantities $\xi_k(x)$ do not depend on the index $k \geq 0$. Then the model of observations has the form

$$z_{nk} = g(\theta_0, x_{nk}) + \xi_k(x_{nk}), \quad k = 0, 1, \dots, n. \quad (13)$$

Denote

$$F_n(\theta) = \frac{1}{n} \sum_{k=0}^n \|z_{nk} - g(\theta, x_{nk})\|^2. \quad (14)$$

Let us assume that the sequence x_{nk} satisfies the averaging condition: (A) there exists a continuous function $x(u)$ such that for any continuous bounded function $f(x)$, $x \in X$,

$$\frac{1}{n} \sum_{k=0}^n f(x_{nk}) \xrightarrow{P} \int_0^1 f(x(u)) du. \quad (15)$$

Remark 1. Note that Condition (A) is oriented to nonstationary situations. It can be verified, for example, in the case where x_{nk} is a recurrent semi-Markovian-type process (see [12, 13]).

THEOREM 4. Let the function $g(\theta, x)$ be uniformly continuous in $\Theta \times X$. For any $x \in X$

$$g(\theta, x) \neq g(\theta_0, x) \quad \text{for } \theta \neq \theta_0, \quad (16)$$

$$\mathbf{E} \xi_1(x) \equiv 0, \quad \mathbf{E} \xi_1(x) \xi_1(x)^* = R(x)^2, \quad (17)$$

condition (A) holds and the Lindeberg condition in the following form is satisfied:

$$\lim_{L \rightarrow \infty} \sup_{x \in X} \mathbf{E} \|\xi_1(x)\|^2 \chi(\|\xi_1(x)\| > L) = 0. \quad (18)$$

Then

$$\{\theta_n\} \xrightarrow{P} \theta_0. \quad (19)$$

Proof. We will present the function $F_n(\theta)$ in the form

$$F_n(\theta) = \frac{1}{n} \sum_{k=0}^n \|g(\theta, x_{nk}) - g(\theta_0, x_{nk})\|^2 - \frac{2}{n} \sum_{k=0}^n (g(\theta, x_{nk}) - g(\theta_0, x_{nk}), \xi_k(x_{nk})) + \frac{1}{n} \sum_{k=0}^n \|\xi_k(x_{nk})\|^2. \quad (20)$$

The quantities $\xi_k(x_{nk})$, $k = 0, 1, \dots$, are conditionally independent for a fixed trajectory x_{nk}

$$\mathbf{E}\xi_k(x_{nk}) = 0, \quad \mathbf{E}[\xi_k(x_{nk})\xi_k(x_{nk})^* / x_{nk}] = R(x_{nk})^2$$

and the Lindeberg condition is satisfied. Then, by virtue of the boundedness of the function $g(\theta, x)$, the quantities $\eta_k(x_{nk}) = (g(\theta, x_{nk}) - g(\theta_0, x_{nk}), \xi_k(x_{nk}))$ also satisfy the Lindeberg condition and

$$\mathbf{E}[\eta(x_{nk}) / x_{nk}]^2 = (R(x_{nk})^2)(g(\theta, x_{nk}) - g(\theta_0, x_{nk}), g(\theta, x_{nk}) - g(\theta_0, x_{nk})).$$

Denote $R_n^2 = \sum_{k=0}^n \mathbf{E}\eta(x_{nk})^2$. By the central limit theorem $\frac{1}{R_n} \sum_{k=0}^n \eta_k(x_{nk}) \xrightarrow{w} \mathcal{N}(0, 1)$. Since by virtue of Condition (A)

$$\begin{aligned} \frac{1}{n} R_n^2 &\xrightarrow{P} \int_0^1 (R(x(u))^2)(g(\theta, x(u)) - g(\theta_0, x(u))), \\ &g(\theta, x(u)) - g(\theta_0, x(u))) du = \sigma^2, \end{aligned} \quad (21)$$

we have $\frac{1}{\sqrt{n}} \sum_{k=0}^n \eta_k(x_{nk}) \xrightarrow{w} \mathcal{N}(0, \sigma^2)$. Then for any fixed θ

$$\frac{1}{n} \sum_{k=0}^n (g(\theta, x_{nk}) - g(\theta_0, x_{nk}), \xi_k(x_{nk})) \xrightarrow{P} 0.$$

Moreover, by virtue of the law of large numbers

$$\frac{1}{n} \sum_{k=0}^n \|\xi_k(x_{nk})\|^2 \xrightarrow{P} \int_0^1 \beta(x(u)) du, \quad (22)$$

where $\beta(x) = \mathbf{E}\|\xi_1(x)\|^2$, and in accordance with Condition (A)

$$\frac{1}{n} \sum_{k=0}^n \|g(\theta, x_{nk}) - g(\theta_0, x_{nk})\|^2 \xrightarrow{P} \int_0^1 \|g(\theta, x(u)) - g(\theta_0, x(u))\|^2 du.$$

Finally, according to relations (20), (22), we obtain that

$$F_n(\theta) \xrightarrow{P} \int_0^1 \|g(\theta, x(u)) - g(\theta_0, x(u))\|^2 du + \int_0^1 \beta(x(u)) du = F(\theta)$$

and the point θ_0 is the unique point of minimum of the function $F(\theta)$.

Let us now prove the U -convergence of the sequence of functions $F_n(\theta)$ to the function $F(\theta)$. Note that

$$\begin{aligned} &\|g(\theta_1, x_{nk}) - g(\theta_0, x_{nk})\|^2 - \|g(\theta_2, x_{nk}) - g(\theta_0, x_{nk})\|^2 = (g(\theta_1, x_{nk}) \\ &- g(\theta_2, x_{nk}), g(\theta_1, x_{nk}) + g(\theta_2, x_{nk}) - 2g(\theta_0, x_{nk})). \end{aligned}$$

Denote $D = \sup_{\theta, x} \|g(\theta, x)\|$

$$\Delta_U(c) = \sup_x \sup_{\|\theta_1 - \theta_2\| < c} \|g(\theta_1, x) - g(\theta_2, x)\|$$

Then

$$\Delta_U(c, F_n(\cdot)) \leq 4D\Delta_U(c) + 2\Delta_U(c) \frac{1}{n} \sum_{k=0}^n \|\xi_k(x_{nk})\|. \quad (23)$$

Since

$$\frac{1}{n} \sum_{k=0}^n \|\xi_k(x_{nk})\| \xrightarrow{P} \int_0^1 \alpha(x(u)) du,$$

where $\alpha(x) = \mathbf{E}\|\xi_1(x)\|$, from (23) it follows that for any $\varepsilon > 0$

$$\lim_{c \rightarrow +0} \limsup_{n \rightarrow \infty} \mathbf{P}\{\Delta_U(c, F_n(\cdot)) > \varepsilon\} = 0,$$

and since the function $g(\theta, x)$ is uniformly continuous in a closed bounded domain, $\Delta_U(c) \rightarrow 0$ as $c \rightarrow 0$. This finally proves statement (19).

Now, we will examine the behavior of the normalized deviation $\theta_n - \theta_0$. Let us consider the case where the function $g(\theta, x)$ is twice continuously differentiable with respect to θ . Let us introduce the matrices

$$Q^2 = \int_0^1 g'_\theta(\theta_0, x(u)) g'_\theta(\theta_0, x(u)) du,$$

$$B^2 = \int_0^1 g'_\theta(\theta_0, x(u))^* R(x(u))^2 g'_\theta(\theta_0, x(u)) du,$$

where g'_θ is a matrix whose elements are partial derivatives with respect to the components of the vector θ of the elements of the vector $g(\theta, x)$.

THEOREM 5. Let Condition (A) hold, the function $g(\theta, x)$ be twice continuously differentiable with respect to θ , the function $g'_\theta(\theta, x)$ be uniformly continuous in (θ, x) , and the second partial derivatives of the elements of the function $g(\theta, x)$ be bounded.

Then there exists a sequence $\tilde{\theta}_n$ of local minima of the function $F_n(\theta)$ such that

$$\sqrt{n}(\tilde{\theta}_n - \theta_0) \xrightarrow{w} 2(Q^2 + Q^{2*})^{-1} B \mathcal{N}(0, 1). \quad (24)$$

Proof. Let us consider the random function $A_n(z) = n(F_n(\theta_0 + \frac{1}{\sqrt{n}}z) - F_n(\theta_0))$. Then

$$A_n(z) = \frac{1}{n} \sum_{k=0}^n n \|g(\theta_0 + \frac{1}{\sqrt{n}}z, x_{nk}) - g(\theta_0, x_{nk})\|^2 - \frac{2}{\sqrt{n}} \sum_{k=0}^n \sqrt{n} (g(\theta_0 + \frac{1}{\sqrt{n}}z, x_{nk}) - g(\theta_0, x_{nk}), \xi_k(x_{nk}))$$

$$= \frac{1}{n} \sum_{k=0}^n (g'_\theta(\theta_0, x_{nk})^* (g'_\theta(\theta_0, x_{nk})z, z) + O\left(\frac{1}{\sqrt{n}}\right))$$

$$- 2 \frac{1}{\sqrt{n}} \sum_{k=0}^n (g'_\theta(\theta_0, x_{nk})z, \xi_k(x_{nk})) - \frac{2}{n} \sum_{k=0}^n (\tilde{g}(\theta_0, z, x_{nk}), \xi_k(x_{nk})),$$

where the vector $\tilde{g}(\theta, z, x_{nk})$ can be written in terms of the second derivatives of the function $g(\theta, x)$ with respect to θ and is bounded by condition, and

$$\|\tilde{g}(\theta_0, z_1, x) - \tilde{g}(\theta_0, z_2, x)\| \leq \|z_1 - z_2\|. \quad (25)$$

By analogy with the proof of Theorem 3, it follows from these relations that the function $A_n(z)$ U -converges to the function $A(z) = (Q^2 z, z) - 2(B \mathcal{N}(0, 1), z)$ in each bounded domain $\|z\| \leq L$. From here, according to Example 1, we obtain the statement of Theorem 5.

As an illustration, we will consider the case where

$$g(\theta, x) = (g(\theta), f(x)), \quad (26)$$

where $g(\theta)$ and $f(x)$ are some vector-valued functions.

Assume for the sake of simplicity that the quantities $\xi_k(x)$ are one-dimensional random variables, $\mathbf{E}\xi_k(x) = 0$, $D\xi_k(x) = R(x)^2$.

THEOREM 6. Let condition (A) hold, the functions $g(\theta)$, $f(x)$, and $R(x)$ be continuous and bounded

$$g(\theta) \neq g(\theta_0) \text{ for } \theta \neq \theta_0, \quad (27)$$

the Lindeberg condition be satisfied and, for some $\beta > 0$ and any fixed $L > 0$, be uniformly in $\|z\| \leq L$,

$$\lim_{h \rightarrow +0} h^{-\beta} (g(\theta_0 + zh) - g(\theta_0)) = a(z), \quad (28)$$

where the function $a(z)$ is continuous and such that the equation

$$a(z) = y \quad (29)$$

has a unique solution $z = a^{-1}(y)$ for any $y \in \mathcal{R}^m$.

Then there exists a sequence $\tilde{\theta}_n$ of local minima of the function $F_n(\theta)$ such that

$$n^{1/2\beta} (\tilde{\theta}_n - \theta_0) \xrightarrow{w} k_0 = a^{-1}(\zeta), \quad (30)$$

where the vector ζ has normal distribution with the parameters

$$(0, C^{-1} \int_0^1 f(x(u)) R(x(u))^2 f(x(u))^* du C^{*-1}),$$

$$C = \int_0^1 f(x(u)) f(x(u))^* du.$$

Proof. Denote $v_n = n^{1/2\beta}$ and $\alpha = 2\beta$. Then the function $A_n(z)$ can be presented in the form

$$A_n(z) = \frac{1}{n} v_n^{2\beta} \sum_{k=0}^n (g(\theta_0 + \frac{1}{v_n} z) - g(\theta_0), f(x_{nk}))^2 - 2(v_n^\beta (g(\theta_0 + \frac{1}{v_n} z) - g(\theta_0)),$$

$$\frac{1}{\sqrt{n}} \sum_{k=0}^n f(x_{nk}) \xi_k(x_{nk})). \quad (31)$$

Since for any vectors $a, b, (a, b)^2 = (bb^* a, a)$, the first term on the right-hand side of (31) can be presented as

$$\left(\frac{1}{n} \sum_{k=0}^n f(x_{nk}) f(x_{nk})^* v_n^\beta (g(\theta_0 + \frac{1}{v_n} z) - g(\theta_0), v_n^\beta (g(\theta_0 + \frac{1}{v_n} z) - g(\theta_0)) \right)$$

and according to conditions (A) and (28), this term uniformly converges to the function $(C(a(z), a(z)))$ in each bounded domain $\|z\| \leq L$.

Further, let us note that the random vectors $f(x) \xi_k(x)$ have the mean value 0 and

$$\mathbf{E} f(x) \xi_k(x) (f(x) \xi_k(x))^* = f(x) f(x)^* R(x).$$

Then the process $\frac{1}{\sqrt{n}} \sum_{k=0}^n f(x_k) \xi(x_{nk})$ U -converges to the process $\int_0^1 f(x(u)) R(x(u)) dw(u)$. It follows from these

relations that the function $A_n(z)$ in domain $\|u\| \leq L$ U -converges to the function

$$A(z) = (Ca(z), a(z)) - 2(a(z), \int_0^1 f(x(u)) R(x(u)) dw(u)).$$

According to Example 1, the minimum of this function can be presented in the form

$$a(z) = 2(C(t)C(t)^*)^{-1} \int_0^1 f(x(u)) R(x(u)) dw(u).$$

Since the function $f(x)f(x)^*$ is self-adjoint, the statement of Theorem 6 follows from (29).

Example 3. Let the function $x(t)$, $t \in [0, T]$, be continuous, and observations be fulfilled at the points $t_{nk} = k/n$, $k = 0, 1, \dots, [nT]$, and have the form

$$y_k = (g(\theta_0), f(x(k/n))) + \xi_k(x(k/n)), \quad (32)$$

where $E\xi_k(x) = 0$, $E\xi_k(x)\xi_k(x)^* = R^2(x)$. Let the vector-valued function $g(\theta)$ be continuously differentiable, $G(\theta) = g'_\theta(\theta)$, and the matrices $G(\theta)$ and C be nondegenerate.

If the function $f(x)$ is continuous, and Conditions (A) and (27) are satisfied, then the statement of Theorem 5 holds, where $v_n = \sqrt{n}$, $\beta = 1$, $\alpha = 2$.

$$\kappa_0 = G(\theta_0)^{-1} C^{-1} \int_0^1 f(x(u))R(x(u))dw(u).$$

Proof. Indeed, in this case,

$$\frac{1}{n} \sum_{k=0}^n f(x(k/n)) \rightarrow \int_0^1 f(x(u))du,$$

in relation (28) $a(z) = G(\theta_0)z$, and this statement follows from the result of Theorem 6.

Remark 2. The results of Theorems 4–6 can be extended in the same way to the case where the sequence x_{nk} is ergodic in the following sense: there exists a probability measure $\pi(A)$, $A \in B_X$, such that for any measurable function $\varphi(x)$, $x \in X$,

$$\frac{1}{n} \sum_{k=0}^n \varphi(x_{nk}) \xrightarrow{P} \int_X \varphi(x)\pi(dx). \quad (33)$$

Note that condition (33) is satisfied for a wide class of Markovian and stationary sequences.

REFERENCES

1. I. A. Ibragimov and R. Z. Has'minskii, *Statistical Estimation — Asymptotic Theory*, Springer-Verlag, New York (1981).
2. Bo M. Bibby and M. Sorencen, "Martingale estimation functions for discretely observed diffusion processes," *Bernoulli*, **1** (1/2), 17-39 (1995).
3. R. L. Liptser and A. N. Shiryaev, *Statistics of Random Processes*, Springer, New York (1977).
4. Yu. A. Kutoyants, *Parameter Estimation for Stochastic Processes*, Heldermann Verlag, Berlin (1984).
5. V. V. Anisimov, "Parameter estimation of switching Poisson processes," *Teor. Veroyatn. Mat. Stat.*, No 31, 3-13 (1985).
6. V. V. Anisimov, "The asymptotic behavior of solutions of some classes of stochastic equations and application in problems of statistics," *Vych. Prikl. Mat.*, No. 57, 103-112 (1985).
7. V. V. Anisimov and A. O. Orzakychev, "Asymptotic parameter estimation of recurrent semi-Markovian-type processes," *Teor. Veroyatn. Mat. Stat.*, No. 49, 1-13 (1993).
8. V. V. Anisimov, "Extremal problems for random functions and applications in statistical parameter estimation," in: *Proc. of the 1st Ukrainian-Scandinavian Conf. "Stochastic Dynamical Systems: Theory and Applications"* (Uzhgorod, Ukraine, Sept. 30-Oct. 5, 1995), *Theory Stoch. Process.*, **2** (18), No. 1-2 (1996), pp. 28-43.
9. V. V. Anisimov and Hussein Kaibah, "Asymptotic behavior of solutions of stochastic equations and applications in statistics," *J. Turk. Stat. Assoc. Istatistik*, **1**, No. 3, 31-42 (1998).
10. B. L. S. Rao Prakasa, *Asymptotic Theory of Statistical Inference*, John Wiley & Sons (1987).
11. V. V. Anisimov and A. V. Seilkhamer, "The asymptotic properties of extremal sets of random fields," *Teor. Veroyatn. Mat. Stat.*, No. 51, 3-11 (1994).
12. V. V. Anisimov and A. O. Aliev, "The limiting theorems for recurrent semi-Markovian-type processes," *Teor. Veroyatn. Mat. Stat.*, No. 41, 7-13 (1990).
13. V. V. Anisimov, "Switching processes: averaging principle, diffusion approximation and applications," *Acta Appl. Math.*, **40**, 95-141 (1995).