Theory of Stochastic Processes
Vol.13 (29), no.1-2, 2007, pp.86-97

ALEXANDER V. IVANOV AND IGOR V. ORLOVSKY

# CONSISTENCY OF M-ESTIMATES IN GENERAL NONLINEAR REGRESSION MODELS

Nonlinear regression model with continuous time and weak dependent or long-range dependent stationary noise is considered. Strong consistency sufficient conditions of $M$-estimates of regression parameters are obtained.

## 1. Introduction

Consider a regression model

$$X(t) = g(y(t), \theta) + \varepsilon(t),\ t \geq 0, \tag{1}$$

where $g(y,\tau)$ is a non random function defined on $Y \times \Theta^c$, $\Theta^c$ is the closure in $\mathbf{R}^q$ of an open set $\Theta$, $Y \subset \mathbf{R}^m$ is a compact region of regression experiment design. Borel function $y(t) : [0,\infty) \to Y$ is a design of regression experiment, $\theta = (\theta_1, ..., \theta_q) \in \Theta^c$ is an unknown parameter. Let $\varepsilon(t),\ t \in \mathbf{R}^1$ be a random process satisfying the assumption

**A1.** $\varepsilon(t), t \in \mathbf{R}^1$ is a real valued mean-square continuous measurable stationary process with zero mean on a complete probability space $(\Omega, \Im, P)$.

We do not assume function $g(y,\theta)$ to be a linear form of coordinates of the vector $\theta$.

**Definition 1.** *$M$-estimate of unknown parameter $\theta$ obtained by the observations $X(t), t \in [0,T)$, of the type (1), is said to be any random vector $\widehat{\theta}_T$ that minimizes in $\tau \in \Theta^c$ the functional $M_T(\tau) = \frac{1}{T}\int\limits_0^T \rho(X(t) - g(y(t),\tau))dt$ with continuous risk function $\rho : \mathbf{R}^1 \to \mathbf{R}^1$.*

The consistency property of $M$-estimates for nonlinear regression model with independent identically distributed observation errors is considered in [1]. Some facts on consistency of the least squares estimates and least moduli estimates can be found in [2].

Sufficient conditions for strong consistency of $M$-estimates of an unknown parameter $\theta$ of the model (1) with random noise that satisfies weak or long-range dependence conditions are presented in this paper.

2000 *Mathematics Subject Classifications.* Primary 62J02; Secondary 62J99
*Key words and phrases.* Consistency, M-estimates, nonlinear regression model.

## 2. Assumptions and the main results

Let us impose some restriction on the random process $\varepsilon(t),\ t\in\mathbf{R}^1$.

**A2.** $\varepsilon(t),\ t\in\mathbf{R}^1$ is a strictly stationary process, such that for some $\delta>0$ $\mu_{2+\delta}=E|\varepsilon(0)|^{2+\delta}<\infty$ and

$$\int\limits_0^\infty (\alpha(r))^{\frac{\delta}{2+\delta}}\,dr<\infty,$$

where

$$\alpha(r)=\sup_{A\in\sigma(-\infty,s],\ B\in\sigma[s+r,\infty)}|P(AB)-P(A)P(B)|,$$

$\sigma(a,b]$ is $\sigma$-algebra generated by random variables (r.v.) $\{\varepsilon(t),\ t\in(a,b]\}$.

**Definition 2.** *If for symmetric r.v. $\xi$ the probabilities $P\{|\xi-b|<x\}$, $x\in[0,\infty)$ are nonincreasing functions of the variable $b\in[0,\infty)$, then we say that $\xi$ is a symmetric and unimodal r.v..*

**A3.** $\varepsilon(0)$ is a symmetric and unimodal r.v. with the distribution function (d.f.) $F(x)$.

Let $\mathcal{B}$ be a $\sigma$-algebra of Borel subsets of $Y$. For any $A\in\mathcal{B}$

$$\mu_T(A)=T^{-1}m\{t\in[0,T]:\ y(t)\in A\},$$

where $m$ is Lebesgue measure on $[0,\infty)$.

Let $\Delta g(y,\tau)=g(y,\theta)-g(y,\tau)$ and $v_\theta(\varepsilon)=\{\tau\in\mathbf{R}^q:\|\tau-\theta\|<\varepsilon\}$.

**B1.** The measures $\mu_T$ are weakly converge, as $T\to\infty$, to some measure $\mu$: $\mu_T\Longrightarrow\mu$ and for any $\varepsilon>0$ $\mu\{y\in Y:\Delta g(y,\tau)=0\}<1$ for each $\tau\notin v_\theta(\varepsilon)$.

**Example.** Assume $\{y_i\}_{i\ge1}\subset Y$ to be some sequence and $y(t)=y_i,\ t\in[i-1;i),\ i=1,2,\dots$ . Introduce the measure

$$\mu_T=\frac{1}{T}\sum_{i=1}^{[T]}\delta_{y_i}+\frac{\{T\}}{T}\delta_{y_{[T]+1}},$$

where $[T]$ and $\{T\}$ are integer and fractional parts of $T$. Then, if $\frac{1}{n}\sum\limits_{i=1}^{n}\delta_{y_i}\Rightarrow\mu$ as $n\to\infty$, then $\mu_T\Rightarrow\mu$ as $T\to\infty$.

Requirement on the measure $\mu$ in the condition **B1** can be written as follows: for any $\varepsilon>0$ $\mu\{y\in Y:\ g(y,\tau)\ne g(y,\theta)\}>0$ for each $\tau\notin v_\theta(\varepsilon)$.

Suppose that the measure $\mu$ is absolutely continuous with respect to Lebesgue measure $l$ on $Y$, furthermore $l(Y)>0$ and $\mu$ has the density $f(y)$ separated from zero: $\inf_{y\in Y}f(y)\ge f_*>0$. Then

$$\mu\{y\in Y:\ g(y,\tau)\ne g(y,\theta)\}=\int\limits_{\{y\in Y:\ g(y,\tau)\ne g(y,\theta)\}}f(y)dy\ge$$

$$\geq f_* l\{y \in Y : \; g(y,\tau) \neq g(y,\theta)\} > 0,$$

if $l\{y \in Y : \; g(y,\tau) \neq g(y,\theta)\} > 0$. But the last inequality is the property of the regression function to distinguish parameters.

**Definition 3.** *Function $J(\cdot) : \mathbf{R}^1 \to \mathbf{R}^1$ is called symmetric, if there exists some point $b_0 \in \mathbf{R}^1$(which is called the center of symmetry) and some function $\varphi(\cdot) : [0,\infty) \to \mathbf{R}^1$ such that $J(b) = \varphi(|b - b_0|)$. If $\varphi$ is a monotonically nondecreasing function and $\varphi(x) > \varphi(0)$ for $x > 0$, then $J$ is called unimodal and the center of symmetry is called the mode.*

Impose some restriction on risk function. Let $E\rho(\varepsilon(t)) = E\rho(\varepsilon(0)) < \infty$.

**C1.** $\rho(x)$ is continuous unimodal, with mode in zero, function such that $\rho(0) = 0$.

**C2.** There exists $c > 0$ such that $|\rho(x_1) - \rho(x_2)| \leq c|x_1 - x_2|$ for any $x_1,\ x_2 \in \mathbf{R}^1$.

Assume also

**A4.** $\int\limits_0^\infty [P\{|\varepsilon(0)| < z\} - P\{|\varepsilon(0) - b| < z\}]\, d\rho(z) > 0,\ b > 0.$

Note that from **C1** it follows that $\rho(x)$ is monotonically nondecreasing function in the region $x \geq 0$. It means that Lesbegue-Stilties integral in **A4** exists. Moreover, from **A3** it follows that the difference in square brackets **A4** is nonnegative.

**Theorem 1.** *Suppose that assumptions* **A1-A4***,* **B1***,* **C1** *and* **C2** *are fulfilled. Then $M$-estimate $\widehat{\theta}_T \to \theta$ a.s. as $T \to \infty$.*

To state the second result of the paper we need to introduce additional condition on $\varepsilon(t)$.

**Definition 4.** *Stationary process $\varepsilon(t),\ t \in \mathbf{R}^1 \quad E\varepsilon(t) = 0$ is called a process with long-range dependence if*

$$E\varepsilon(0)\varepsilon(t) = B(t) = \frac{L(|t|)}{|t|^\alpha},\ \ 0 < \alpha < 1, \tag{2}$$

*where $L(t) : [0,\infty) \to [0,\infty)$ is a slowly varying function (at infinity).*

**A5.** Gaussian random process $\varepsilon(t), t \in \mathbf{R}^1$ is a process with long-range dependence, $B(0) = 1$.

**Theorem 2.** *Suppose that assumptions* **A1, A4, A5, B1, C1** *and* **C2** *are fulfilled. Then $M$-estimate $\widehat{\theta}_T \to \theta$ a.s. as $T \to \infty$.*

## 3. Auxiliary assertions

Set

$$\delta_T(\tau) = Q_T(\tau) - EQ_T(\tau),\ \Delta_T(\tau) = Q_T(\tau) - Q_T(\theta).$$

**Definition 5.** *An unknown parameter $\theta$ is said to be identifiable, if for any $\varepsilon > 0$ there exist the numbers $T_0 = T_0(\varepsilon)$ and $\delta = \delta(\varepsilon) > 0$ such that $E\Delta_T(\tau) > \delta$ when $T > T_0$ and $\tau \notin v_\theta(\varepsilon)$.*

**Lemma 1.** *Assume that $\theta$ is identifiable parameter and*

$$\sup_{\tau\in\Theta^c} |\delta_T(\tau)| \underset{T\to\infty}{\longrightarrow} 0 \quad \text{a.s.,} \tag{3}$$

*then $\widehat{\theta}_T \longrightarrow \theta$ a.s. as $T \to \infty$.*

*Proof.* Let us denote by $\Omega_1$ the event of the probability 1, for which the condition (3) is fulfilled. For elementary events $\omega \in \Omega_1$ from the definition of the estimate $\widehat{\theta}_T$ we have

$$\Delta_T(\widehat{\theta}_T) \le 0. \tag{4}$$

Suppose that for some fixed $\omega \in \Omega_1$ $\widehat{\theta}_T \not\longrightarrow \theta$ as $T \to \infty$. It means that there exists some number $\varepsilon_0 > 0$ and the sequence of numbers $T_n \uparrow \infty$ as $n \to \infty$ such that for $n > n(\varepsilon_0)$ $\left\| \widehat{\theta}_{T_n} - \theta \right\| \ge \varepsilon_0$. As for these $T_n$ (4) also holds, then $\inf_{\tau\notin v_\theta(\varepsilon_0)} \Delta_{T_n}(\tau) \le 0$.

Let $T_n \ge T_0(\varepsilon_0)$ and for $n > n(\varepsilon_0)$ $\sup_{\tau\in\Theta^c} |\delta_{T_n}(\tau)| < \frac{\delta_0}{4}$, where $\delta_0 = \delta(\varepsilon_0)$. Then for $n > n(\varepsilon_0)$

$$\begin{aligned} 0 \ge \inf_{\tau\notin v_\theta(\varepsilon_0)} \Delta_{T_n}(\tau) &= \inf_{\tau\notin v_\theta(\varepsilon_0)} (\delta_{T_n}(\tau) + E\Delta_{T_n}(\tau)) - \delta_{T_n}(\theta) \\ &\ge \inf_{\tau\notin v_\theta(\varepsilon_0)} \delta_{T_n}(\tau) + \inf_{\tau\notin v_\theta(\varepsilon_0)} E\Delta_{T_n}(\tau) - \delta_{T_n}(\theta) \\ &\ge \inf_{\tau\in\Theta^c} \delta_{T_n}(\tau) + \inf_{\tau\notin v_\theta(\varepsilon_0)} E\Delta_{T_n}(\tau) - \delta_{T_n}(\theta) > \frac{\delta_0}{2}. \end{aligned}$$

We obtain contradiction. Hence, for $\omega \in \Omega_1$ $\widehat{\theta}_T \longrightarrow \theta$ as $T \to \infty$. □

Introduce function $J(b) = E\rho(\varepsilon(t) - b) = E\rho(\varepsilon(0) - b),\ b \in \mathbf{R}^1$.

The next lemma states sufficient conditions of identifiability of parameter $\theta$.

**Lemma 2.** *An unknown parameter $\theta$ is identifiable if*

(i) *for any $\varepsilon > 0$ there exist $T_0 = T_0(\varepsilon)$ and $x = x(\varepsilon) > 0$ such that for any $T > T_0$ and any $\tau \notin v_\theta(\varepsilon)$ $\mu_T \{y \in Y : |\Delta g(y, \tau)| > x\} > x$;*

(ii) *$J(b)$ is unimodal;*

(iii) *$J(b) > J(0)$ for any $b \ne 0$.*

*Proof.* It is easily seen that under the conditions (ii) and (iii) the mode is in $b = 0$. Furthermore,

$$E\Delta_T(\tau) = \frac{1}{T} \int_0^T [J(\Delta g(y(t), \tau)) - J(0)]\, dt. \tag{5}$$

Fix some $\varepsilon > 0$ and consider numbers $T_0$ and $x$ taken from the condition (i) of the Lemma. From the condition (ii) it follows that the right hand side of the relation (5) permits the lower bound

$$\frac{1}{T}\int_0^T [J(\Delta g(y(t),\tau)) - J(0)]\,dt \geq (J(x)-J(0))\,\frac{1}{T}\int_0^T \chi_{(x,\infty)}\,(|\Delta g(y(t),\tau)|)\,dt$$

$$= (J(x)-J(0))\,\mu_T\left\{y \in Y : |\Delta g(y,\tau)| > x\right\},$$

where $\chi_A(x)$ is the indicator of the set $A$.

From (i) and (ii) it follows that in the definition of the identifiability of parameter one can set $\delta = x(J(x)-J(0))$. □

Further we formulate some sufficient conditions on the validity of Lemmas 1 and 2.

**Lemma 3.** *If the assumption* **B1** *holds, then the condition (i) of the Lemma 2 fulfiles.*

*Proof.* Let $\varepsilon > 0$ be an arbitrary number. It is necessary to show that there exists some numbers $T_0$ and $x > 0$ such that

$$\mu_T\left\{y \in Y : |\Delta g(y,\tau)| > x\right\} > x,\ T > T_0,\ \tau \notin v_\theta(\varepsilon). \tag{6}$$

Assume that (6) does not hold. Then there exist some sequences $T_n \uparrow \infty$ as $n \to \infty$ and $\tau_n \in \Theta^c \backslash v_\theta(\varepsilon)$ such that

$$\mu_{T_n}\left\{y \in Y : |\Delta g(y,\tau_n)| > n^{-1}\right\} \leq n^{-1},\ n \geq 1. \tag{7}$$

As the set $\Theta^c \backslash v_\theta(\varepsilon)$ is compact, there exists some point $\tau^* \in \Theta^c \backslash v_\theta(\varepsilon)$ and the sequence $n_k,\ k \geq 1$ such that $\tau_{n_k} \to \tau^*$ as $k \to \infty$.

Let $\delta > 0$ be an arbitrary fixed number. Then there exists some number $k_\delta$ such that for $k > k_\delta$, uniformly in $y \in Y$,

$$|\Delta g(y,\tau_{n_k}) - \Delta g(y,\tau^*)| \leq \frac{\delta}{2}. \tag{8}$$

Thanks to (8), for $k > k_\delta$

$$\{|\Delta g(y,\tau^*)| > \delta\} \subset \{|\Delta g(y,\tau^*) - \Delta g(y,\tau_{n_k})| + |\Delta g(y,\tau_{n_k})| > \delta\}$$

$$\subset \left\{|\Delta g(y,\tau_{n_k})| > \frac{\delta}{2}\right\}. \tag{9}$$

Taking into account the inequality (7) for $n_k > \frac{2}{\delta}$ one has

$$\mu_{T_{n_k}}\left\{y \in Y : |\Delta g(y,\tau_{n_k})| > \frac{\delta}{2}\right\} \leq \frac{1}{n_k}. \tag{10}$$

Then, from (9) and (10) it follows that

$$\mu_{T_{n_k}}\left\{y \in Y : |\Delta g(y,\tau^*)| > \delta\right\} \leq n_k^{-1}, \tag{11}$$

which is true for any $k > k'_\delta = \max\left(k_\delta,\ \min\left\{k : n_k > \frac{2}{\delta}\right\}\right)$.

Denote by $Y_\delta = \{y \in Y : |\Delta g(y, \tau^*)| \le \delta\}$. From (11) it follows that $\mu_{T_{n_k}}(Y_\delta) > 1 - n_k^{-1}$ for all $k > k'_\delta$.

As $Y_\delta$ is a closed set, then thanks to weak convergence of $\mu_T$ to the measure $\mu$, we obtain (see, for example, [3], p. 21)

$$\overline{\lim_{k\to\infty}}\, \mu_{T_{n_k}}(Y_\delta) \le \mu(Y_\delta),\ \delta > 0.$$

For $\delta \downarrow 0$, from the continuity of the measure $\mu$ it follows that

$$\mu\{y \in Y : \Delta g(y, \tau) = 0\} = 1. \qquad (12)$$

But the relation (12) contradicts to the condition **B1**. □

**Lemma 4.** *If the assumptions* **A3, A4** *and* **C1** *hold, then the conditions (ii) and (iii) of the Lemma 2 are fulfilled.*

*Proof.* Without loss of generality, assume that $\rho(x),\ x \ge 0$ is strictly monotonically increasing function. From the formula for the mean of the nonnegative r.v. (see, for example, [4], p. 190) one has

$$J(b) - J(0) = \int_0^\infty \left(P\{\rho(\varepsilon(0)) < x\} - P\{\rho(\varepsilon(0) - b) < x\}\right) dx =$$

$$\int_0^\infty \left(P\left\{-\rho^{-1}(x) < \varepsilon(0) < \rho^{-1}(x)\right\} - P\left\{-\rho^{-1}(x) < \varepsilon(0) - b < \rho^{-1}(x)\right\}\right) dx,$$

where $\rho^{-1}(x)$ is the inverse of the function $\rho(x),\ x \ge 0$.

By the change of variable $x = \rho(z),\ z \ge 0$ in the last integral,

$$J(b) - J(0) = \int_0^\infty \left(P\left\{|\varepsilon(0)| < z\right\} - P\left\{|\varepsilon(0) - b| < z\right\}\right) d\rho(z) =$$

$$= \int_0^\infty \left(F(z) - F(z-b) - F(z+b) + F(z)\right) d\rho(z), \qquad (13)$$

where $F(x)$ is the d.f. of the r.v. $\varepsilon(0)$.

The integral in the first equality of the relations (13) coincides with the expression of **A4**, and the condition (iii) of Lemma 2 is fulfilled.

From the symmetry of $\rho$ and r.v. $\varepsilon(0)$ it follows the symmetry of $J(b)$.

Denote by $\Delta_b^2 F(z) = \left(F(z) - F(z-b)\right) - \left(F(z+b) - F(z)\right),\ b, z \ge 0$. Then **A4** can be rewritten in the form

$$\int_0^\infty \Delta_b^2 F(z) d\rho(z) > 0,\ b > 0.$$

From (13) it follows that

$$\Delta_b^2 F(z) = P\{|\varepsilon(0)| < z\} - P\{|\varepsilon(0) - b| < z\}.$$

Consider for $b_2 > b_1$ the difference

$$J(b_2) - J(b_1) = \int\limits_0^\infty \left(\Delta_{b_2}^2 F(z) - \Delta_{b_1}^2 F(z)\right) d\rho(z).$$

It is easily seen that

$$\Delta_{b_2}^2 F(z) - \Delta_{b_1}^2 F(z) = P\{|\varepsilon(0) - b_1| < z\} - P\{|\varepsilon(0) - b_2| < z\} \geq 0$$

from the unimodality of the r.v. $\varepsilon(0)$. It means that $J(b_2) - J(b_1) \geq 0$, and the condition (ii) of Lemma 2 is a corollary of **A3** and **C1**. □

Assume that the d.f. $F(x)$ is continuously differentiable and the density of the distribution $p(x)$ is an even strictly decreasing for $x \geq 0$ function. Suppose that a continuous even function $\rho(x)$ is such that $\rho(0) = 0$ and strictly monotonically increasing for $x \geq 0$. Then one can use Lemma 10.2 of the book [3], p. 217-218, and for any $b \neq 0$

$$J(b) - J(0) = E\rho(\varepsilon(0) - b) - E\rho(\varepsilon(0)) > 0,$$

and the integral in **A4** is strictly positive.

Consider next sufficient conditions of the uniform convergence in (3) of Lemma 1.

**Lemma 5.** *Suppose the condition* **C2** *fulfiles and*

$$\delta_T(\tau) \underset{T\to\infty}{\longrightarrow} 0 \quad \text{a.s.}, \tau \in \Theta^c, \tag{14}$$

*then (3) holds.*

*Proof.* From **C2** it follows that for $\tau_1,\ \tau_2 \in \Theta^c$

$$|Q_T(\tau_1) - Q_T(\tau_2)| \leq \frac{c}{T} \int\limits_0^T |g(y(t), \tau_1) - g(y(t), \tau_2)| dt.$$

Similarly, from **C2** for $\tau_1,\ \tau_2 \in \Theta^c$ one has

$$|\delta_T(\tau_1) - \delta_T(\tau_2)| \leq \frac{2c}{T} \int\limits_0^T |g(y(t), \tau_1) - g(y(t), \tau_2)| dt.$$

Hence, the family of functions $\{\delta_T(\tau) : \omega \in \Omega,\ T > 0\}$ is an equicontinuous on the set $\Theta^c$. So for any $\delta > 0$ there exists a finite number of points $\tau_1, ..., \tau_k \in \Theta^c$ such that

$$\sup_{\tau\in\Theta^c} |\delta_T(\tau)| \leq \max_{1\leq j\leq k} |\delta_T(\tau_j)| + \delta,\ \omega \in \Omega,\ T > 0.$$

From (14) it follows that $\max_{1\leq j\leq k} |\delta_T(\tau_j)| \longrightarrow 0$ a.s. as $T \to \infty$, and, hence, $\sup_{\tau\in\Theta^c} |\delta_T(\tau)| \longrightarrow 0$ a.s. as $T \to \infty$. □

## 4. Proof of Theorem 1

We shall prove that (14) holds under the assumptions of Theorem 1. Using the notation

$$\xi(t) = \rho(\varepsilon(t) - \Delta g(y(t),\tau)) - E\rho(\varepsilon(t) - \Delta g(y(t),\tau)),\ \tau \in \Theta^c$$

one has

$$\delta_T(\tau) = \frac{1}{T}\int_0^T \xi(t)dt, \quad E\delta_T^2(\tau) = \frac{1}{T^2}\int_0^T\int_0^T E\xi(t)\xi(s)dtds \leq$$

$$\leq \frac{10}{T^2}\int_0^T\int_0^T \left[E\rho^{2+\delta}(\varepsilon(t) - \Delta g(y(t),\tau))\right]^{\frac{1}{2+\delta}} \times$$

$$\times \left[E\rho^{2+\delta}(\varepsilon(s) - \Delta g(y(s),\tau))\right]^{\frac{1}{2+\delta}} \alpha^{\frac{\delta}{2+\delta}}(|t-s|)dtds. \tag{15}$$

To obtain (15) the Davidov inequality has been used with $p = q = 2+\delta,\ r = 1 + \frac{2}{\delta}$ (see [5], and also Lemma 1.6.2 of the book [6]).

As $\rho(0) = 0$, then from the condition **C2** one obtains

$$E\rho^{2+\delta}\left(\varepsilon(t) - \Delta g(y(t),\tau)\right) \leq c^{2+\delta} E\left|\varepsilon(0) - \Delta g(y(t),\tau)\right|^{2+\delta}.$$

By obvious inequalities

$$|a+b|^{\kappa} \leq 2^{\kappa-1}\left(|a|^{\kappa} + |b|^{\kappa}\right), \qquad |a+b|^{\frac{1}{\kappa}} \leq |a|^{\frac{1}{\kappa}} + |b|^{\frac{1}{\kappa}},\ \kappa = 2+\delta, \tag{16}$$

$$\left[E\rho^{2+\delta}\left(\varepsilon(t) - \Delta g(y(t),\tau)\right)\right]^{\frac{1}{2+\delta}} \leq 2^{\frac{1+\delta}{2+\delta}} c\left(\mu_{2+\delta}^{\frac{1}{2+\delta}} + |\Delta g(y(t),\tau)|\right),$$

i.e.

$$E\delta_T^2(\tau) \leq 2^{\frac{\delta}{2+\delta}} c^2 \frac{20}{T^2}\int_0^T\int_0^T \alpha^{\frac{\delta}{2+\delta}}(|t-s|)\left[\mu_{2+\delta}^{\frac{1}{2+\delta}} + |\Delta g(y(t),\tau)|\right] \times$$

$$\times \left[\mu_{2+\delta}^{\frac{1}{2+\delta}} + |\Delta g(y(s),\tau)|\right] dtds \leq$$

$$\leq 2^{\frac{\delta}{2+\delta}} c^2 \frac{20}{T^2}\int_0^T\int_0^T \alpha^{\frac{\delta}{2+\delta}}(|t-s|)\left[\mu_{2+\delta}^{\frac{1}{2+\delta}} + |\Delta g(y(t),\tau)|\right]^2 dtds.$$

Using the first inequality of (16) with $\kappa = 2$,

$$E\delta_T^2(\tau) \leq 2^{\frac{\delta}{2+\delta}} c^2 \frac{40}{T^2}\int_0^T\int_0^T \alpha^{\frac{\delta}{2+\delta}}(|t-s|)\left[\mu_{2+\delta}^{\frac{2}{2+\delta}} + |\Delta g(y(t),\tau)|^2\right] dtds.$$

It remains to estimate two integrals, namely:

$$I_1 = \frac{1}{T^2}\int_0^T\int_0^T \alpha^{\frac{\delta}{2+\delta}}(|t-s|)dtds \le \frac{1}{T^2}\int_0^T ds \int_{-T}^T \alpha^{\frac{\delta}{2+\delta}}(|t|)dt = O(T^{-1})$$

as $T \to \infty$, under assumption **A2**. On the other hand,

$$I_2 = \frac{1}{T^2}\int_0^T\int_0^T \alpha^{\frac{\delta}{2+\delta}}(|t-s|)|\Delta g(y(t),\tau)|^2 dtds$$

$$\le \left(2\int_0^\infty \alpha^{\frac{\delta}{2+\delta}}(s)ds\right)\frac{1}{T^2}\int_0^T |\Delta g(y(t),\tau)|^2 dt. \qquad (17)$$

As $g(y,\tau)$ is continuous function on the compact $Y \times \Theta^c$ , the right hand side of the inequality (17) is of the order $O(T^{-1})$ as $T \to \infty$.

Thus, $E\delta_T^2(\tau) = O(T^{-1})$ as $T \to \infty$, and $\delta_T(\tau) \longrightarrow 0$ in probability as $T \to \infty$ .

Note that for the sequence $T_n = n^2$, $n \ge 1$ $\sum\limits_{n=1}^{\infty} E\delta_{T_n}^2(\tau) < \infty$, i.e. $\delta_{T_n}(\tau) \longrightarrow_{n\to\infty} 0$ a.s.

If $T \in [T_n, T_{n+1}]$, then

$$|\delta_T(\tau)| \le \sup_{T_n \le T \le T_{n+1}} |\delta_T(\tau) - \delta_{T_n}(\tau)| + |\delta_{T_n}(\tau)|,$$

and the Theorem will be proved, if $\sup\limits_{T_n \le T \le T_{n+1}} |\delta_T(\tau) - \delta_{T_n}(\tau)| \longrightarrow_{n\to\infty} 0$ a.s.

Obviously

$$\delta_T(\tau) - \delta_{T_n}(\tau) = \frac{1}{T}\int_0^T \xi(t)dt - \frac{1}{T_n}\int_0^{T_n} \xi(t)dt =$$

$$= \left(\frac{1}{T} - \frac{1}{T_n}\right)\int_0^{T_n} \xi(t)dt + \frac{1}{T}\int_{T_n}^T \xi(t)dt = I_3 + I_4.$$

Furthermore, for $T \in [T_n, T_{n+1}]$

$$|I_3| \le \frac{T_{n+1} - T_n}{T_n} |\delta_{T_n}(\tau)| \underset{n\to\infty}{\longrightarrow} 0 \text{ a.s.};$$

$$|I_4| \le \frac{1}{T_n}\int_{T_n}^{T_{n+1}} |\xi(t)|dt \le \frac{1}{T_n}\int_{T_n}^{T_{n+1}} \rho\left(\varepsilon(t) - \Delta g(y(t),\tau)\right)dt +$$

$$+\frac{1}{T_n}\int\limits_{T_n}^{T_{n+1}} E\rho\left(\varepsilon(t)-\Delta g(y(t),\tau)\right)dt = I_5 + I_6.$$

As under the Lipshits condition **C2**

$$\rho\left(\varepsilon(t)-\Delta g(y(t),\tau)\right) \le c\left(|\varepsilon(t)| + |\Delta g(y(t),\tau)|\right),$$

then

$$I_5 \le \frac{c}{T_n}\int\limits_{T_n}^{T_{n+1}} |\varepsilon(t)|dt + \frac{c}{T_n}\int\limits_{T_n}^{T_{n+1}} |\Delta g(y(t),\tau)|dt = I_7 + I_8,$$

$$I_8 = c\left(\frac{T_{n+1}}{T_n}\cdot\frac{1}{T_{n+1}}\int\limits_0^{T_{n+1}} |\Delta g(y(t),\tau)|dt - \frac{1}{T_n}\int\limits_0^{T_n} |\Delta g(y(t),\tau)|dt\right).$$

From the assumption **B1** of the Theorem it follows

$$\frac{1}{T_n}\int\limits_0^{T_n} |\Delta g(y(t),\tau)|dt = \int\limits_Y |\Delta g(y,\tau)|\mu_{T_n}(dy) \underset{n\to\infty}{\longrightarrow} \int\limits_Y |\Delta g(y,\tau)|\mu(dy),$$

then $I_8 \longrightarrow 0$ as $n\to\infty$.

On the other hand,

$$I_7 = c\left(\frac{1}{T_n}\int\limits_{T_n}^{T_{n+1}} \left(|\varepsilon(t)| - E|\varepsilon(t)|\right)dt + E|\varepsilon(0)|\frac{T_{n+1}-T_n}{T_n}\right) \underset{n\to\infty}{\longrightarrow} 0 \quad \text{a.s.}$$

by Davidov inequality.

Similarly, it can be shown that $I_6 \longrightarrow 0$ as $n\to\infty$.

Consequently, (14) is fulfilled. The validity of Theorem 1 follows now from the Lemmas 1-5 proved above. □

## 5. Proof of Theorem 2

Similarly to proof of Theorem 1 we need to proof that (14) holds. Then the result of Theorem 2 will follow from the Lemmas 1-5.

Consider a random process

$$G(\varepsilon(t),t) = \rho(\varepsilon(t)-\Delta g(y(t),\tau)). \tag{18}$$

From **C2** and **A5**

$$EG^2(\varepsilon(t),t) \le c^2E\left|\varepsilon(t)-\Delta g(y(t),\tau)\right|^2 = c^2\left(1+|\Delta g(y(t),\tau)|^2\right) \le C < \infty \tag{19}$$

uniformly in $t \geq 0$ and $\tau \in \Theta^c$. Therefore in Gilbert space $L_2(\mathbf{R}^1, \varphi(u)du)$, where $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ is a standard Gaussian density, there exists an expansion (see, for example, [6])

$$G(u,t) = \sum_{m=0}^{\infty} \frac{C_m(t)}{m!} H_m(u), \; C_m(t) = \int\limits_{\mathbf{R}^1} G(u,t) H_m(u) \varphi(u) du, \; m \geq 0$$

by Chebyshev-Hermite polynomials

$$H_m(u) = (-1)^m e^{\frac{u^2}{2}} \frac{d^m}{du^m} e^{-\frac{u^2}{2}}, \; m \geq 0. \tag{20}$$

Note that $C_0(t) = E\rho(\varepsilon(0) - \Delta g(y(t), \tau)) = J\left(\Delta g(y(t), \tau)\right)$.

Thanks to relations

$$EH_m(\varepsilon(t)) H_k(\varepsilon(s)) = \delta_m^k m! B^m(t-s), \tag{21}$$

where $\delta_m^k$ is Kroneker delta we have

$$E\xi(t)\xi(s) = \operatorname{cov}\left(G(\varepsilon(t), t), G(\varepsilon(s), s)\right) = \sum_{m=1}^{\infty} \frac{C_m(t) C_m(s)}{m!} B^m(t-s).$$

Hence, taking into account that $B(0) = 1$, we obtain

$$\begin{aligned} E\delta_T^2(\tau) &= \sum_{m=1}^{\infty} \frac{1}{m!} \frac{1}{T^2} \int\limits_0^T \int\limits_0^T C_m(t) C_m(s) B^m(t-s) dt ds \\ &\leq \sum_{m=1}^{\infty} \frac{1}{m!} \frac{1}{T^2} \int\limits_0^T \int\limits_0^T C_m^2(t) B^m(t-s) dt ds \\ &\leq \frac{1}{T^2} \int\limits_0^T \int\limits_0^T \left( \sum_{m=1}^{\infty} \frac{C_m^2(t)}{m!} \right) B(t-s) dt ds. \end{aligned}$$

Note that, thanks to (19),

$$\sum_{m=1}^{\infty} \frac{C_m^2(t)}{m!} = EG^2(\varepsilon(0), t) - \left(EG(\varepsilon(0), t)\right)^2 = DG(\varepsilon(0), t) \leq C < \infty,$$

and

$$E\delta_T^2(\tau) \leq C \frac{1}{T^2} \int\limits_0^T \int\limits_0^T B(t-s) dt ds.$$

On the other hand, as $T \to \infty$,

$$\frac{1}{T^2}\int_0^T\int_0^T B(t-s)dtds = \int_0^1\int_0^1 B\left(T(t-s)\right)dtds = \frac{1}{T^\alpha}\int_0^1\int_0^1 \frac{L\left(T|t-s|\right)}{|t-s|^\alpha}dtds$$

$$\sim \left(\int_0^1\int_0^1 \frac{dtds}{|t-s|^\alpha}\right)\frac{L(T)}{T^\alpha} = \frac{2}{(1-\alpha)(2-\alpha)}\frac{L(T)}{T^\alpha}$$

by the properties of the slowly varying function (see, for example [7],[8]).

For the sequence $T_n = n^{\frac{1}{\alpha}+\nu}$, where $\nu > 0$ is some number, $\sum\limits_{n=1}^{\infty}\frac{L(T_n)}{T_n^\alpha} < \infty$, and so $\delta_{T_n}(\tau) \longrightarrow 0$ a.s., as $n \to \infty$.

Taking into account the proof of Theorem 1, it remains to show that

$$\frac{1}{T_n}\int_0^{T_n}\left(|\varepsilon(t)| - E|\varepsilon(t)|\right)dt \underset{n\to\infty}{\longrightarrow} 0 \text{ a.s.} \tag{22}$$

But the proof of (22) is similar to the previous reasoning for $G(\varepsilon(t),t)$.□

## Bibliography


1. F. Liese and I. Vajda *Consistency of M-estimates in general regression models,* Journal of Multivariate Analysis, **50**, 1, (1994), 93-114.
2. A. V. Ivanov *Asymptotic Theory of Nonlinear Regression,* Kluwer Academic Publishers, Dordrecht, (1997).
3. Ibrahimov I.A. and Hasminsky R.Z. *Asymptotic theory of estimation*, Nauka, Moskow, (1979). (in Russian)
4. Feller V. *Introduction to probability theory and its applications*, Mir, Moskow, volume 2, (1967). (in Russian)
5. Davidov U.A. *About convergence of the distributions, generated by stationary random processes*, TVIP, Vol. 15, 3, (1970), 498-509.(in Russian)
6. Ivanov A.V. and Leonenko N.N. *Statistical analysis of random fields*, Vyscha Shkola, Kyiv,(1986). (in Russian)
7. Seneta E. *Regularly varying function*, Nauka, Moskow, (1985). (in Russian)
8. Ivanov A.V. and Orlovsky I.V. *$L_p$-Estimates in Nonlinear Regression with Long-Range Dependence*, Theory of Stochastic Processes, Vol. 7(23), . 3-4, (2002), 38-49.



National technical university of Ukraine, "KPI". Peremogi avenue 37, Kiev

*E-mail address*: ivanov@paligora.kiev.ua

National technical university of Ukraine, "KPI". Peremogi avenue 37, Kiev

*E-mail address*: avalon@ln.ua