Теорія Ймовір. та Матем. Статист.
Вип. 93, 2015

Theor. Probability and Math. Statist.
No. 93, 2016, Pages 33–49
http://dx.doi.org/10.1090/tpms/993
Article electronically published on February 7, 2017

ASYMPTOTIC PROPERTIES OF M-ESTIMATORS OF PARAMETERS OF A NONLINEAR REGRESSION MODEL WITH A RANDOM NOISE WHOSE SPECTRUM IS SINGULAR

UDC 519.21

A. V. IVANOV AND I. V. ORLOVSKYI

Abstract. Time continuous nonlinear regression model with a noise being a nonlinearly transformed Gaussian stationary process with a singular spectrum is considered in the paper. Sufficient conditions for the asymptotic normality of the M-estimator are found for the vector parameter in this model.

1. Introduction

Sufficient conditions for the asymptotic normality of M-estimators of an unknown parameter of a nonlinear regression model with continuous time and random noise with a singular spectrum are obtained.

Properties of M-estimators for linear regression models with independent errors of observation are considered by Huber [1, 2], Hampel et al. [3] and many other authors thereafter.

Asymptotic properties of M-estimators of parameters for linear as well as for nonlinear regression models with a long range dependent random noise are studied by Koul [4, 5], Koul and Mukherjee [6], Giraitis et al. [7], Koul and Surgailis [8, 9, 10], Giraitis and Koul [11], Koul et al. [12] in the case of discrete time, and by Ivanov and Leonenko [13, 14], Ivanov [15], Ivanov and Orlovskyi [16, 17, 18], Savich [19] in the case of continuous time.

Orlovskyi [20], Ivanov and Orlovskyi [17, 18, 21], Ivanov [15] consider asymptotic properties of M-estimators of parameters for nonlinear regression models with continuous time and a weakly dependent random noise.

In the current paper, we study M-estimators constructed with the help of smooth loss functions. Smooth loss functions as well as their nondifferentiable analogs are widely used when solving various problems in data analysis (see, for example, [22]).

Note that the key tools in the proof of the asymptotic normality are the central limit theorem for weighted nonlinear transformations of a Gaussian stationary stochastic process with a singular spectrum obtained by Ivanov et al. in [23] and the Brouwer fixed point theorem [24, 2]. The latter theorem requires that a solution of the system of normal equations that determines the M-estimator is unique in a certain asymptotic sense. The asymptotic uniqueness of M-estimators is considered by Ivanov [15] and Orlovskyi [18] for nonlinear regression models.

2010 *Mathematics Subject Classification.* Primary 62J02; Secondary 62J99.

Key words and phrases. Asymptotic uniqueness of an estimator, asymptotic normality, M-estimators, nonlinear regression models, singular spectrum.

The paper was prepared following the talk at the International Conference “Probability, Reliability and Stochastic Optimization (PRESTO-2015)” held in Kyiv, Ukraine, April 7–10, 2015.

2. Asymptotic normality of M-estimators

2.1. **Conditions and statement of the main result.** Consider the regression model

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

where $g\colon [0,+\infty) \times \Theta_\beta \to \mathbb{R}$ is a continuous function, $\Theta_\beta = \bigcup_{\|a\|\leq 1}(\Theta + \beta a)$, $\beta > 0$ a certain number, $\Theta \subset \mathbb{R}^q$ a bounded convex open set, and $\theta \in \Theta$ the true value of the parameter.

Throughout the paper we consider the derivatives of the regression function in the set Θ^c, where Θ^c denotes the closure of Θ. Therefore, we need to define the regression function in Θ_β.

We further assume that the noise $\varepsilon(t)$ satisfies the following conditions.

A1. $\varepsilon(t)$, $t \in \mathbb{R}$, is a local functional of a Gaussian stationary stochastic process $\xi(t)$, that is, $\varepsilon(t) = G(\xi(t))$, where $G(x)$, $x \in \mathbb{R}$, is a Borel function. Moreover, $\mathsf{E}\,\varepsilon(0) = 0$ and $\mathsf{E}\,\varepsilon^4(0) < \infty$.

A2. $\xi(t)$, $t \in \mathbb{R}$, is a mean square continuous measurable stationary Gaussian stochastic process with zero mean and covariance function

$$B(t) = \sum_{j=0}^{r} A_j B_{\alpha_j,\chi_j}(t), \qquad r \geq 0, \tag{2}$$

where

$$B_{\alpha_j,\chi_j}(t) = \frac{\cos(\chi_j t)}{(1+t^2)^{\alpha_j/2}},$$

$0 \leq \chi_0 < \chi_1 < \cdots < \chi_r$, $0 < \alpha_j < 1$, $j = 0,\ldots,r$, and $\sum_{j=0}^{r} A_j = 1$, $A_j > 0$.

Such a correlation function is introduced in the paper [25] in order to construct an example of a spectral density with nonzero singularities in contrast to the case of strongly dependent processes where singularities are at zero. Condition **A2** has been used in papers [23, 26] for the same reason.

The spectral density f of the stochastic process ξ is given by

$$f(\lambda) = \sum_{j=0}^{r} A_j f_{\alpha_j,\chi_j}(\lambda), \qquad \lambda \in \mathbb{R},$$

where

$$f_{\alpha_j,\chi_j}(\lambda) = \frac{C_1(\alpha_j)}{2}\left[K_{\frac{\alpha_j-1}{2}}(|\lambda+\chi_j|)|\lambda+\chi_j|^{\frac{\alpha_j-1}{2}} + K_{\frac{\alpha_j-1}{2}}(|\lambda-\chi_j|)|\lambda-\chi_j|^{\frac{\alpha_j-1}{2}}\right],$$

$$j = 0,\ldots,r, \qquad C_1(\alpha) = 2^{\frac{1-\alpha}{2}}/\sqrt{\pi}\,\Gamma\left(\frac{\alpha}{2}\right);$$

$$K_\nu(z) = \frac{1}{2}\int_0^\infty s^{\nu-1}\exp\left\{-\frac{1}{2}\left(s+\frac{1}{s}\right)z\right\}ds, \qquad z \geq 0,\ \nu \in \mathbb{R},$$

is a modified Bessel function of the third kind and of order ν.

Note that $K_{-\nu}(z) = K_\nu(z)$ and

$$K_\nu(z) \sim \Gamma(\nu)2^{\nu-1}z^{-\nu}, \qquad \nu > 0,$$

as $z \downarrow 0$. Thus,

$$f_{\alpha_j,\chi_j}(\lambda) \sim \frac{C_2(\alpha_j)}{2}|\lambda \pm \chi_j|^{\alpha_j-1}\big(1 - h_j(|\lambda \pm \chi_j|)\big)$$

as $\lambda \to \pm\chi_j$, $j = 0, \dots, r$, where $C_2(\alpha) = [2\Gamma(\alpha)\cos(\alpha\pi/2)]^{-1}$ and

$$h_j(|\lambda|) = \frac{\Gamma\left(\frac{\alpha_j+1}{2}\right)}{\Gamma\left(\frac{3-\alpha_j}{2}\right)} \cdot \left|\frac{\lambda}{2}\right|^{1-\alpha_j} + \frac{\Gamma\left(\frac{\alpha_j+1}{2}\right)}{4\Gamma\left(\frac{3+\alpha_j}{2}\right)} \cdot \left|\frac{\lambda}{2}\right|^2 + o\left(|\lambda|^2\right), \qquad \lambda \to 0,\ j = 0, \dots, r.$$

Hence, condition **A2** implies that the spectral density f has $2r+2$ different points of singularity

$$\{-\chi_r, -\chi_{r-1}, \dots, -\chi_1, -\chi_0, \chi_0, \chi_1, \dots, \chi_r\}$$

if $\chi_0 \neq 0$ and $0 < \alpha_j < 1$, $j = 0, \dots, r$. Otherwise, if $\chi_0 = 0$, then there are $2r+1$ points of singularity of the spectral density f.

Definition 2.1. Any random vector $\hat{\theta}_T = \hat{\theta}_T(X(t), t \in [0,T]) \in \Theta^c$, such that

$$(3) \qquad Q_T(\hat{\theta}_T) = \min_{\tau \in \Theta^c} Q_T(\tau), \qquad Q_T(\tau) = \int_0^T \rho(X(t) - g(t,\tau))\, dt, \qquad \tau \in \Theta^c,$$

is called an M-estimator of the unknown parameter $\theta \in \Theta$ constructed from observations $X(t)$, $t \in [0,T]$, for model (1) with a continuous loss function $\rho(x) \geq 0$, $x \in \mathbb{R}$.

Below we list some assumptions imposed on the regression function $g(t,\tau)$ and loss function $\rho(x)$. Let $g(t,\tau)$ be a twice continuously differentiable function with respect to $\tau \in \Theta^c$. Put

$$g_i(t,\tau) = \frac{\partial}{\partial \tau_i} g(t,\tau), \qquad g_{il}(t,\tau) = \frac{\partial}{\partial \tau_i \partial \tau_l} g(t,\tau), \qquad \tau \in \Theta^c,\ i,l = 1, \dots, q;$$

$$d_T^2(\theta) = \operatorname{diag}\left(d_{iT}^2(\theta)\right)_{i=1}^q, \qquad d_{iT}^2(\theta) = \int_0^T g_i^2(t,\theta)\, dt, \qquad i = 1, \dots, q,$$

$$\liminf_{T\to\infty} T^{-1} d_{iT}^2(\theta) > 0, \qquad i = 1, \dots, q,$$

$$d_{il,T}^2(\theta) = \int_0^T g_{il}^2(t,\theta)\, dt, \qquad i,l = 1, \dots, q.$$

The letters k (with subscripts) denote positive constants. Assume that, for all sufficiently large T ($T > T_0$),

B1. $g(t,\cdot) \in C^2(\Theta^c)$ for all $t \geq 0$, and

(i) $\displaystyle\sup_{t\in[0,T]} \sup_{\tau\in\Theta^c} \frac{|g_i(t,\tau)|}{d_{iT}(\theta)} \leq k^i T^{-1/2}$;

(ii) $\displaystyle\sup_{t\in[0,T]} \sup_{\tau\in\Theta^c} \frac{|g_{il}(t,\tau)|}{d_{il,T}(\theta)} \leq k^{il} T^{-1/2}$;

(iii) $\displaystyle\sup_{\tau\in\Theta^c} \frac{d_{il,T}(\theta)}{d_{iT}(\theta) d_{lT}(\theta)} \leq \tilde{k}^{il} T^{-1/2}$, $i,l = 1, \dots, q$.

C1. The function $\rho(x)$ is nonnegative, even, twice continuously differentiable,

$$\rho(0) = 0,$$

and its derivatives $\rho'(x) = \psi(x)$ and $\rho''(x) = \psi'(x)$ are such that

(i) $\mathsf{E}\,\psi\big(G(\xi(0))\big) = 0$;

(ii) $\mathsf{E}\,\psi'\big(G(\xi(0))\big) > 0$;

(iii) for all $x, h \in \mathbb{R}$ and some constant L,

$$|\psi'(x+h) - \psi'(x)| \leq L|h|.$$

Condition **C1**(iii) implies that

$$|\psi(x) - \psi(0)| = |\psi'(\eta x)| \cdot |x| \leq (|\psi'(0)| + L|\eta x|)|x| \leq |\psi'(0)| \cdot |x| + Lx^2$$

for all x and some $\eta = \eta(x) \in (0,1)$, whence

$$|\psi(x)| \le |\psi(0)| + |\psi'(0)| \cdot |x| + Lx^2.$$

Moreover,

$$|\psi'(x)| \le |\psi'(0)| + L|x|.$$

Therefore, the stochastic processes $\psi(G(\xi(t)))$ and $\psi'(G(\xi(t)))$, $t \in \mathbb{R}$, possess finite second moments under assumptions **A1**, **A2**, and **C1**.

Let $K \in L_2(\mathbb{R}, \varphi(x)dx)$, where $\varphi(x) = (2\pi)^{-1/2} e^{-x^2/2}$. Then one can expand the function K into the Fourier series in the space $L_2(\mathbb{R}, \varphi(x)\,dx)$,

$$K(x) = \sum_{n=0}^{\infty} \frac{C_n(K)}{n!} H_n(x), \qquad C_n(K) = \int_{-\infty}^{\infty} K(x) H_n(x) \varphi(x)\,dx, \qquad n \ge 0,$$

with respect to the Chebyshev–Hermite polynomials

$$H_n(x) = (-1)^n e^{\frac{x^2}{2}} \frac{d^n}{dx^n} e^{-\frac{x^2}{2}}, \qquad n \ge 0.$$

Definition 2.2. We say that a function $K \in L_2(\mathbb{R}, \varphi(x)\,dx)$ has Hermite rank m and write $\operatorname{Hrank}(K) = m$ if either $C_1(K) \ne 0$ and $m = 1$ or, for some $m \ge 2$,

$$C_1(K) = \cdots = C_{m-1}(K) = 0, \qquad C_m(K) \ne 0.$$

Under assumption **C1**(i), the functions $\psi \circ G$ and $\psi' \circ G$ can be expanded into the Fourier series with respect to the Chebyshev–Hermite polynomials in the Hilbert space $L_2(\mathbb{R}, \varphi(x)\,dx)$, that is,

$$\psi(G(x)) = \sum_{n=m}^{\infty} \frac{C_n(\psi \circ G)}{n!} H_n(x),$$

$$\psi'(G(x)) = C_0(\psi' \circ G) + \sum_{n=m'}^{\infty} \frac{C_n(\psi' \circ G)}{n!} H_n(x),$$

where $m = \operatorname{Hrank}(\psi \circ G)$, $m' = \operatorname{Hrank}(\psi' \circ G)$, and $C_0(\psi \circ G) = \mathsf{E}\,\psi\big(G(\xi(0))\big) = 0$.

C2. Either

(i) $\operatorname{Hrank}(\psi \circ G) = 1$, $\alpha > \frac{1}{2}$,

(ii) or $\operatorname{Hrank}(\psi \circ G) = m$, $\alpha m > 1$, where $\alpha = \min_{j=0,\dots,r} \alpha_j$ and α_j, $j = 0, \dots, r$, are the numbers involved in condition **A2**.

Put

$$J_T(\theta) = (J_{il,T}(\theta))_{i,l=1}^{q},$$

$$J_{il,T}(\theta) = d_{iT}^{-1}(\theta) d_{lT}^{-1}(\theta) \int_0^T g_i(t,\theta) g_l(t,\theta)\,dt, \qquad i,l = 1,\dots,q.$$

Let $\lambda_{\min}(A)$ ($\lambda_{\max}(A)$) be the minimal (maximal) eigenvalue of a positive definite matrix A.

B2. For some $\lambda_* > 0$ and $T > T_0$, $\lambda_{\min}(J_T(\theta)) \ge \lambda_*$.

Put $\Lambda_T(\theta) = J_T^{-1}(\theta)$ and consider a matrix measure $\mu_T(dx;\theta)$ in $(\mathbb{R}, \mathfrak{B})$ with the matrix of densities

$$\left(\mu_T^{jl}(x;\theta)\right)_{j,l=1}^{q},$$

where $\mathfrak{B}$ is the Borel σ-algebra in $\mathbb{R}$ and

$$\mu_T^{jl}(x;\theta) = g_T^j(x,\theta)\overline{g_T^l(x,\theta)} \left(\int_{\mathbb{R}} \left| g_T^j(x,\theta) \right|^2 dx \int_{\mathbb{R}} \left| g_T^l(x,\theta) \right|^2 dx \right)^{-1/2},$$

$$g_T^j(x,\theta) = \int_0^T e^{ixt} g_j(t,\theta)\,dt, \qquad j,l = 1,\dots,q.$$

Note that $d^2_{jT}(\theta) = (2\pi)^{-1} \int_{\mathbb{R}} \left|g^j_T(x,\theta)\right|^2 dx$.

B3. The family of measures $\mu_T(\cdot;\theta)$ weakly converges as $T \to \infty$ to the measure $\mu(\cdot;\theta)$ such that $\mu(\mathbb{R};\theta)$ is a positive definite matrix.

Definition 2.3 ([28, 27])**.** The matrix measure $\mu(\cdot;\theta) = (\mu^{jl}(\cdot;\theta))^q_{j,l=1}$ is called the spectral measure of the regression function $g(t,\theta)$.

Conditions **B2** and **B3** imply that

$$J_T(\theta) = \int_{\mathbb{R}} \mu_T(dx;\theta) \to \int_{\mathbb{R}} \mu(dx;\theta) = \mu(\mathbb{R};\theta) = J(\theta)$$

as $T \to \infty$. Put $\Lambda(\theta) = J^{-1}(\theta)$.

Next we recall the notion of a μ-admissible spectral density $f(\lambda)$ (more detail is given in [28, 29]).

Definition 2.4. A spectral density f is called μ-admissible if f is integrable with respect to the measure μ, that is, all entries of the matrix

$$\int_{\mathbb{R}} f(\lambda)\, \mu(d\lambda)$$

are finite, and

$$\int_{\mathbb{R}} f(\lambda)\, \mu_T(d\lambda) \to \int_{\mathbb{R}} f(\lambda)\, \mu(d\lambda), \qquad T \to \infty.$$

Sufficient conditions for the μ-admissibility of the spectral density of a stationary process can be found in [23, 30]. These conditions are satisfied, for example, for the spectral density f of the stochastic process ξ with covariance function (2). The key condition is that the family of points of singularity of f and the family of atoms of the spectral measure μ are disjoint. It is worth mentioning that μ is atomic for all examples known up to now.

Let $f^{(*1)}(\lambda) = f(\lambda)$ and

$$f^{(*j)}(\lambda) = \int_{\mathbb{R}^{j-1}} f(\lambda - \lambda_2 - \cdots - \lambda_j) \prod_{i=2}^{j} f(\lambda_i)\, d\lambda_2 \ldots d\lambda_j$$

for $j \geq 2$ be the j-fold convolution of the spectral density $f(\lambda)$ of the stochastic process ξ with itself,

$$\gamma = \left(\mathsf{E}\, \psi'\left(G(\xi(0))\right)\right)^{-1}.$$

A3. $\int_{\mathbb{R}} f^{(*j)}(\lambda)\, \mu(d\lambda)$, $j \geq 1$, are positive definite matrices.

Below we use the following assumption.

D1. For all $\varepsilon > 0$, there exists $T_0 = T_0(\varepsilon)$ such that the system of equations

$$\nabla Q_T(\tau) = 0$$

possesses a unique solution for all $T > T_0$ with probability that is not less than $1 - \varepsilon$.

In Section 3, we provide sufficient conditions for **D1** that hold simultaneously with assumptions of Theorem 2.1 provided that $d_{iT}(\theta)$, $d_{il,T}(\theta) = O(T^{1/2})$, $i, l = 1, \ldots, q$.

Theorem 2.1. *Let assumptions* **A1-A3, B1-B3, C1, C2, D1** *hold and the spectral density f of the stochastic process ξ is μ-admissible. Then the distribution of the random vector $\hat{u}_T(\theta) = d_T(\theta)(\hat{\theta}_T - \theta)$ converges as $T \to \infty$ to the Gaussian distribution $N(0, \sigma(\theta))$, where*

$$\sigma(\theta) = 2\pi\gamma^2 \Lambda(\theta) \cdot \left(\sum_{j=m}^{\infty} \frac{C_j^2(\psi \circ G)}{j!} \int_{-\infty}^{\infty} f^{(*j)}(\lambda)\, \mu(d\lambda, \theta) \right) \cdot \Lambda(\theta). \tag{4}$$

2.2. **Auxiliary results.** Consider the normalized M-estimator

(5) $$\hat{u}_T = \hat{u}_T(\theta) = d_T(\theta)(\hat{\theta}_T - \theta).$$

Now we change the variables in the regression function and its derivatives in a way that corresponds to normalization (5), that is,

$$g(t,\tau) = g\left(t, \theta + d_T^{-1}(\theta)u\right) = h(t,u), \qquad g_i(t,\tau) = g_i\left(t, \theta + d_T^{-1}(\theta)u\right) = h_i(t,u),$$
$$g_{il}(t,\tau) = g_{il}\left(t, \theta + d_T^{-1}(\theta)u\right) = h_{il}(t,u), \qquad i,l = 1,\dots,q.$$

Also put

$$H(t;u_1,u_2) = h(t,u_1) - h(t,u_2), \qquad H_i(t;u_1,u_2) = h_i(t,u_1) - h_i(t,u_2), \quad i = 1,\dots,q.$$

Consider the vectors

$$M_T(u) = \left(M_T^i(u)\right)_{i=1}^q = \left(\gamma \int_0^T \psi(X(t) - h(t,u)) \frac{h_i(t,u)}{d_{iT}(\theta)}\, dt\right)_{i=1}^q$$

and

$$\Psi_T(u) = \left(\Psi_T^i(u)\right)_{i=1}^q = \left(\gamma \int_0^T \psi\big(G(\xi(t))\big) \frac{h_i(t,u)}{d_{iT}(\theta)}\, dt + \int_0^T H(t;0,u) \frac{h_i(t,u)}{d_{iT}(\theta)}\, dt\right)_{i=1}^q.$$

The vectors $M_T(u)$ and $\Psi_T(u)$ are defined for $u \in U_T^c(\theta)$ and $U_T(\theta) = d_T(\theta)(\Theta - \theta)$.

According to the assumptions imposed above, the sets $U_T(\theta)$ are expanding to $\mathbb{R}^q$ as $T \to \infty$. Then $v(R) = \{u \in \mathbb{R}^q \colon \|u\| < R\} \subset U_T(\theta)$ for an arbitrary $R > 0$ and $T > T_0(R)$.

The statistical meaning of the vectors $M_T(u)$ and $\Psi_T(u)$ is easy to understand. Consider the functional $\gamma Q_T(\theta + d_T^{-1}(\theta)u)$. Then the normalized M-estimator $\hat{u}_T$ satisfies the system of equations

(6) $$M_T(u) = 0.$$

Let

$$\eta(t) = \gamma\psi\big(G(\xi(t))\big), \qquad t \in \mathbb{R},$$

and let the observations be of the following form:

(7) $$Y(t) = g(t,\theta) + \eta(t), \qquad t \in [0,T].$$

Then

$$\Psi_T(u) = 0$$

is the system of normal equations used to determine the normalized least squares estimator

$$\breve{u}_T = \breve{u}_T(\theta) = d_T(\theta)(\breve{\theta}_T - \theta)$$

of the unknown parameter θ of a virtual regression model (7).

Lemma 2.1. *Let assumptions* **A1**, **A2**, **B1**, *and* **C1** *hold. Then*

(8) $$\mathsf{P}\left\{\sup_{u \in v^c(\mathbb{R})} \|M_T(u) - \Psi_T(u)\| > r\right\} \to 0, \qquad T \to \infty,$$

for all $R > 0$ and $r > 0$.

Proof. For a fixed i,

$$M_T^i(u) - \Psi_T^i(u) = \gamma \int_0^T \frac{h_i(t,u)}{d_{iT}(\theta)} \left[\psi(G(\xi(t)) + H(t;0,u)) - \psi\big(G(\xi(t))\big) - \psi'\big(G(\xi(t))\big)H(t;0,u)\right] dt + \gamma \int_0^T H(t;0,u)\frac{h_i(t,u)}{d_{iT}(\theta)}\zeta(t)\,dt = I_1(u) + I_2(u),$$

$$\zeta(t) = \psi'\big(G(\xi(t))\big) - \mathsf{E}\,\psi'\big(G(\xi(t))\big), \qquad t \in \mathbb{R}.$$

Now we prove that $I_1(u)$ and $I_2(u)$ converge to zero in probability uniformly with respect to $u \in v^c(\mathbb{R})$. Let $u \in v^c(\mathbb{R})$ be fixed. Then $\mathsf{E}\, I_2(u) = 0$ and

(9) $$\mathsf{E}\, I_2^2(u) = \gamma^2 \int_0^T\int_0^T H(t;0,u)H(s;0,u)\frac{h_i(t,u)}{d_{iT}(\theta)}\frac{h_i(s,u)}{d_{iT}(\theta)}\operatorname{cov}(\zeta(t),\zeta(s))\,dt\,ds.$$

The Taylor formula and Cauchy–Bunyakovskiĭ inequality imply that

$$\sup_{t\in[0,T]} |H(t;0,u)| = \sup_{t\in[0,T]}\left|\sum_{i=1}^q \frac{h_i(t,u_t^*)}{d_{iT}(\theta)}u_i\right| \le \|u\| \sup_{t\in[0,T]}\left(\sum_{i=1}^q \left[\frac{h_i(t,u_t^*)}{d_{iT}(\theta)}\right]^2\right)^{1/2},$$

where $\|u_t^*\| \le \|u\|$.

In view of **B1**(i), we derive from the latter inequality that

(10) $$\sup_{t\in[0,T]} |H(t;0,u)| \le T^{-1/2}\|k\|\cdot\|u\|,$$

where $k = (k^1,\dots,k^q)$ is the vector of constants in assumption **B1**(i).

Applying inequality (10) and condition **B1**(i) to integral (9), we get

$$\mathsf{E}\, I_2^2(u) \le \gamma^2\|k\|^2\left(k^i\right)^2 R^2T^{-2}\int_0^T\int_0^T \operatorname{cov}(\zeta(t),\zeta(s))\,dt\,ds.$$

Then we show that

(11) $$T^{-2}\int_0^T\int_0^T \operatorname{cov}(\zeta(t),\zeta(s))\,dt\,ds \to 0, \qquad T\to\infty.$$

Using the following relation (see, for example, [31])

$$\mathsf{E}\, H_l(\xi(t))H_n(\xi(s)) = \delta_l^n l!\, B^n(t-s),$$

where δ_l^n denotes the Kronecker symbol, we obtain

$$\operatorname{cov}\big(\psi'\big(G(\xi(t))\big),\psi'\big(G(\xi(s))\big)\big) = \sum_{n=m'}^\infty \frac{C_n^2(\psi'\circ G)}{n!}B^n(t-s).$$

Since $|B(t)|\le 1$, $t\in\mathbb{R}$, we conclude that

(12) $$\left|\operatorname{cov}\left(\psi'\big(G(\xi(t))\big),\psi'(G(\xi(s)))\right)\right| \le \sum_{n=m'}^\infty \frac{C_n^2(\psi'\circ G)}{n!}|B(t-s)| \le \mathbf{D}\,\psi'\big(G(\xi(0))\big)\cdot|B(t-s)|$$

and

$$T^{-2}\int_0^T\int_0^T \operatorname{cov}(\zeta(t),\zeta(s))\,dt\,ds \le \mathbf{D}\,\psi'\big(G(\xi(0))\big)\cdot T^{-2}\int_0^T\int_0^T |B(t-s)|\,dt\,ds.$$

On the other hand,

$$T^{-2}\int_0^T\int_0^T |B(t-s)|\,dt\,ds \le T^{-2}\int_0^T\int_0^T \frac{dt\,ds}{|t-s|^{\alpha}} = T^{-\alpha}\int_0^1\int_0^1 \frac{dt'\,ds'}{|t'-s'|^{\alpha}} = O(T^{-\alpha}), \tag{13}$$

where $\alpha = \min_{j=0,\dots,r}\alpha_j$, that is, relation (11) holds. Thus, $I_2(u) \xrightarrow{\mathsf{P}} 0$ as $T\to\infty$ pointwise for $u\in v^c(\mathbb{R})$.

If $u_1,u_2\in v^c(\mathbb{R})$, then

$$I_2(u_1)-I_2(u_2) = \gamma\int_0^T H(t;0,u_1)\frac{H_i(t;u_1,u_2)}{d_{iT}(\theta)}\zeta(t)\,dt - \gamma\int_0^T H(t;u_1,u_2)\frac{h_i(t;u_2)}{d_{iT}(\theta)}\zeta(t)\,dt = I_3(u_1,u_2)+I_4(u_1,u_2).$$

Consider the probability

$$\begin{aligned}\mathsf{P}\left\{\sup_{\|u_1-u_2\|\le h}|I_3(u_1,u_2)|>r\right\} &\le r^{-1}\,\mathsf{E}\sup_{\|u_1-u_2\|\le h}|I_3(u_1,u_2)| \\ &\le 2r^{-1}\gamma\,\mathsf{E}\big|\psi'\big(G(\xi(0))\big)\big|T\sup_{u\in v^c(\mathbb{R})}\sup_{t\in[0,T]}|H(t;0,u)| \\ &\quad\times\sup_{\|u_1-u_2\|\le h}\sup_{t\in[0,T]}\frac{|H_i(t;u_1,u_2)|}{d_{iT}(\theta)};\end{aligned} \tag{14}$$

$$\begin{aligned}&\sup_{\|u_1-u_2\|\le h}\sup_{t\in[0,T]}\frac{|H_i(t;u_1,u_2)|}{d_{iT}(\theta)} \\ &\quad\le h\sup_{t\in[0,T]}\sum_{l=1}^{q}\left(\sup_{u\in v^c(\mathbb{R})}\frac{|h_{il}(t,u)|}{d_{il,T}(\theta)}\right)\frac{d_{il,T}(\theta)}{d_{iT}(\theta)d_{lT}(\theta)} \le \sum_{l=1}^{q}k^{il}\tilde{k}^{il}hT^{-1},\end{aligned} \tag{15}$$

for all $h>0$ and $r>0$ (we used conditions **B1**(ii) and **B1**(iii)).

Then we apply inequalities (10) and (15) to (14). As a result, we conclude that

$$\mathsf{P}\left\{\sup_{\|u_1-u_2\|\le h}|I_3(u_1,u_2)|>r\right\} \le k_1 r^{-1}T^{-1/2}h, \tag{16}$$

where $k_1 = 2\gamma\,\mathsf{E}\big|\psi'\big(G(\xi(0))\big)\big|R\|k\|\big(\sum_{i,l=1}^{q}k^{il}\tilde{k}^{il}\big)$.

Similarly, taking into account **B1**(i),

$$\begin{aligned}\mathsf{P}\left\{\sup_{\|u_1-u_2\|\le h}|I_4(u_1,u_2)|>r\right\} &\le r^{-1}\,\mathsf{E}\sup_{\|u_1-u_2\|\le h}|I_4(u_1,u_2)| \\ &\le 2r^{-1}\gamma\,\mathsf{E}\big|\psi'\big(G(\xi(0))\big)\big|T \\ &\quad\times\sup_{u\in v^c(\mathbb{R})}\sup_{t\in[0,T]}\frac{|h_i(t,u)|}{d_{iT}(\theta)}\sup_{\|u_1-u_2\|\le h}\sup_{t\in[0,T]}|H(t;u_1,u_2)| \\ &\le k_2r^{-1}h,\end{aligned} \tag{17}$$

where $k_2 = 2\gamma\,\mathsf{E}\big|\psi'\big(G(\xi(0))\big)\big|k^i\|k\|$.

Now (16) and (17) imply that

$$\mathsf{P}\left\{\sup_{\|u_1-u_2\|\le h}|I_2(u_1)-I_2(u_2)|>r\right\} \le 2r^{-1}h\left(k_1T^{-1/2}+k_2\right) \le k_3r^{-1}h. \tag{18}$$

Denote by N_h a finite h-net of the ball $v^c(\mathbb{R})$. Then

$$\sup_{u\in v^c(\mathbb{R})}|I_2(u)| \le \sup_{\|u_1-u_2\|\le h}|I_2(u_1)-I_2(u_2)| + \max_{u\in N_h}|I_2(u)|. \tag{19}$$

It follows from (18) and (19) that

$$\mathsf{P}\left\{\sup_{u\in v^c(\mathbb{R})}|I_2(u)|>r\right\}\le 2k_3r^{-1}h+\mathsf{P}\left\{\max_{u\in N_h}|I_2(u)|>r/2\right\}$$

for all $r>0$.

For $\varepsilon>0$ let $h=\varepsilon r/(4k_3)$. Then

$$\mathsf{P}\left\{\max_{u\in N_{\frac{\varepsilon r}{4k_3}}}|I_2(u)|>\frac{r}{2}\right\}\le\frac{\varepsilon}{2}$$

for $T>T_0$ by the pointwise convergence of $I_2(u)$ to zero in probability. Therefore,

$$\mathsf{P}\left\{\max_{u\in v^c(\mathbb{R})}|I_2(u)|>r\right\}\le\varepsilon.$$

On the other hand, if $t\in[0,T]$ and $u\in v^c(R)$ are fixed, then almost surely there exists a number $\delta\in(0,1)$ such that

$$\begin{aligned}
&\big|\psi(G(\xi(t))+H(t;0,u))-\psi\big(G(\xi(t))\big)-\psi'\big(G(\xi(t))\big)H(t;0,u)\big|\\
&\quad=\big|\psi'(G(\xi(t))+\delta H(t;0,u))-\psi'\big(G(\xi(t))\big)\big|\cdot|H(t;0,u)|\\
&\quad\le L\cdot|H(t;0,u)|^2\le L\|k\|^2R^2T^{-1}.
\end{aligned}\tag{20}$$

By **B1**(i) and (20),

$$\sup_{u\in v^c(\mathbb{R})}|I_1(u)|\le L\gamma k^i\|k\|^2R^2T^{-1/2}\quad\text{almost surely,}$$

and thus Lemma 2.1 is proved. □

Consider the random vector

$$L_T(u)=(L_T^i(u))_{i=1}^q=\left(\int_0^T\left(\eta(t)-\sum_{l=1}^q\frac{g_l(t,\theta)}{d_{lT}(\theta)}u_l\right)\cdot\frac{g_i(t,\theta)}{d_{iT}(\theta)}\,dt\right)_{i=1}^q \tag{21}$$

that corresponds to the virtual linear regression model

$$Z(t)=\sum_{i=1}^q g_i(t,\theta)\beta_i+\eta(t),\qquad t\in[0,T].$$

The system of normal equations

$$L_T(u)=0 \tag{22}$$

determines a normalized linear least squares estimator $\tilde\beta_T$ of the parameter $\beta\in\mathbb{R}^q$, that is, the estimator

$$\tilde u_T=\tilde u_T(\theta)=d_T(\theta)(\tilde\beta_T-\beta). \tag{23}$$

Lemma 2.2. *Let all the assumptions of Lemma* 2.1 *hold. Then*

$$\mathsf{P}\left\{\sup_{u\in v^c(\mathbb{R})}\|\Psi_T(u)-L_T(u)\|>r\right\}\to 0,\qquad T\to\infty, \tag{24}$$

for all $R>0$ and $r>0$.

Proof. It is clear that

$$\begin{aligned}\Psi_T^i(u) - L_T^i(u) &= \int_0^T \eta(t)\frac{h_i(t,u)}{d_{iT}(\theta)}\,dt + \int_0^T H(t;0,u)\frac{h_i(t,u)}{d_{iT}(\theta)}\,dt \\ &\quad - \int_0^T \eta(t)\frac{h_i(t,0)}{d_{iT}(\theta)}\,dt + \int_0^T \frac{h_i(t,0)}{d_{iT}(\theta)}\sum_{l=1}^q \frac{h_l(t,0)}{d_{lT}(\theta)}u_l\,dt \\ &= \int_0^T \eta(t)\frac{H_i(t;u,0)}{d_{iT}(\theta)}\,dt + \int_0^T H(t;0,u)\frac{H_i(t;u,0)}{d_{iT}(\theta)}\,dt \\ &\quad + \int_0^T \frac{h_i(t,0)}{d_{iT}(\theta)}\left[H(t;0,u) + \sum_{l=1}^q \frac{h_l(t,0)}{d_{lT}(\theta)}u_l\right]\,dt \\ &= I_5(u) + I_6(u) + I_7(u).\end{aligned}$$

If $u \in v^c(\mathbb{R})$ is fixed, then inequality (15) implies that

$$\begin{aligned}\mathsf{E}\, I_5^2(u) &= \int_0^T\int_0^T \operatorname{cov}(\eta(t),\eta(s))\frac{H_i(t;u,0)}{d_{iT}(\theta)}\frac{H_i(s;u,0)}{d_{iT}(\theta)}\,dt\,ds \\ &\le \left(\sum_{l=1}^q k^{il}\tilde{k}^{il}\right)^2 R^2\cdot T^{-2}\int_0^T\int_0^T |\operatorname{cov}(\eta(t),\eta(s))|\,dt\,ds.\end{aligned}$$

Now we show that

$$T^{-2}\int_0^T\int_0^T \operatorname{cov}(\eta(t),\eta(s))\,dt\,ds \to 0, \qquad T\to\infty. \tag{25}$$

Similarly to (12),

$$\left|\operatorname{cov}\big(\psi\big(G(\xi(t))\big),\psi\big(G(\xi(s))\big)\big)\right| \le \mathsf{E}\,\psi^2\big(G(\xi(0))\big)|B(t-s)|. \tag{26}$$

Using (13) and (26), we get

$$\begin{aligned}&T^{-2}\int_0^T\int_0^T |\operatorname{cov}(\eta(t),\eta(s))|\,dt\,ds \\ &\qquad \le \gamma^2\,\mathsf{E}\,\psi^2\big(G(\xi(0))\big)\cdot T^{-2}\int_0^T\int_0^T |B(t-s)|\,dt\,ds = O(T^{-\alpha}),\end{aligned}$$

and thus (25) holds; that is, $I_5(u)\xrightarrow{\mathsf{P}} 0$ as $T\to\infty$ pointwise for $u\in v^c(\mathbb{R})$.

On the other hand, it follows from (15) that

$$\mathsf{E}\sup_{\|u_1-u_2\|\le h}|I_5(u_1)-I_5(u_2)| \le |\gamma|\,\mathsf{E}\left|\psi\big(G(\xi(0))\big)\right|\left(\sum_{l=1}^q k^{il}\tilde{k}^{il}\right)h.$$

Similarly to the case of $I_2(u)$ in the proof of Lemma 2.1, one can prove that $I_5(u)$ converges to zero in probability uniformly with respect to $u\in v^c(\mathbb{R})$.

Taking into account inequalities (10) and (15), we get

$$\sup_{u\in v^c(\mathbb{R})}|I_6(u)| \le \|k\|\left(\sum_{l=1}^q k^{il}\tilde{k}^{il}\right)R^2T^{-1/2}\to 0, \qquad T\to\infty.$$

Note that $I_7(u)$ can be written as

$$I_7(u) = -\frac{1}{2}\sum_{j,l=1}^q\left(\int_0^T \frac{h_{jl}(t,u_T^*)}{d_{jT}(\theta)d_{lT}(\theta)}\frac{h_i(t,0)}{d_{iT}(\theta)}\,dt\right)u_ju_l$$

for some $u_T^* \in v(R)$. Then

$$|I_7(u)| \leq \frac{k^i}{2} \sum_{j,l=1}^{q} \left(k^{jl}\tilde{k}^{jl}|u_j| \cdot |u_l|\right) T^{-1/2} \leq \frac{qk^i}{2} \max_{j,l=1,\dots,q} \left[k^{jl}\tilde{k}^{jl}\right] \|u\|^2 T^{-1/2}$$

by **B1**, and thus $\sup_{u\in v^c(\mathbb{R})} |I_7(u)| \to 0$ as $T \to \infty$. Lemma 2.2 is proved. □

Applying (8) and (24) we obtain the following result.

Corollary 2.1. *Let all the assumptions of Lemma* 2.1 *hold. Then*

$$\mathsf{P}\left\{\sup_{u\in v^c(\mathbb{R})} \|M_T(u) - L_T(u)\| > r\right\} \to 0 \qquad T \to \infty.$$

for all $R > 0$ and $r > 0$.

If condition **B2** holds, then relations (21) and (22) yield

$$\tilde{u}_T = \Lambda_T(\theta) d_T^{-1}(\theta) \int_0^T \eta(t) \nabla g(t,\theta)\, dt \tag{27}$$

(see (23)).

Now we are ready to state the theorem on the asymptotic normality of the weighted integral of a nonlinear transformation of a Gaussian stationary stochastic process with a singular spectrum [23].

Theorem 2.2. *Assume that conditions* **A1***,* **A2***,* **B1***(i),* **B2***,* **B3** *hold. Further, let at least one of the following conditions hold for the function $K \in L_2(\mathbb{R}, \varphi(x)\, dx)$:*

(i) Hrank$(K) = 1$ *and the spectral density f of the stochastic process ξ is μ-admissible;*
(ii) Hrank$(K) = m$ *and $\alpha m > 1$, where $\alpha = \min_{j=0,\dots,r} \alpha_j$.*

Then the random vector

$$\zeta_T = d_T^{-1}(\theta) \int_0^T K(\xi(t)) \nabla g(t,\theta)\, dt$$

is asymptotically normal as $T \to \infty$ with parameters $N(0, \bar{\sigma})$, where

$$\bar{\sigma} = 2\pi \sum_{j=m}^{\infty} \frac{C_j^2(K)}{j!} \int_{\mathbb{R}} f^{(*j)}(\lambda)\, \mu(d\lambda, \theta).$$

In what follows, we need the following well-known Brouwer fixed point theorem (see, for example, [24]).

Theorem 2.3. *Let $F\colon v^c(R) \to v^c(R)$ be a continuous mapping. Then there exists a point $x_0 \in v^c(R)$ such that $F(x_0) = x_0$.*

Let $\mathfrak{B}^q$ denote the σ-algebra of Borel subsets of $\mathbb{R}^q$. For $A \in \mathfrak{B}^q$ and $\varepsilon > 0$, let

$$A_\varepsilon = \left\{x \in \mathbb{R}^q \colon \inf_{y\in A} \|x - y\| < \varepsilon\right\}, \qquad A_{-\varepsilon} = \mathbb{R}^q \setminus (\mathbb{R}^q \setminus A)_\varepsilon.$$

The following result is proved in §3 of the book [32].

Theorem 2.4. *Let ν be a nonnegative differentiable function on $[0,+\infty)$ such that*

$$b=\int_0^{\infty}|\nu'(\lambda)|\lambda^{q-1}\,d\lambda<+\infty,\qquad \lim_{\lambda\to\infty}\nu(\lambda)=0.$$

Then

$$\int_{C_\varepsilon\setminus C_{-\delta}}\nu(\|\lambda\|)\,d\lambda\le b\left(\frac{2\pi^{q/2}}{\Gamma(\frac{q}{2})}\right)(\varepsilon+\delta)$$

for an arbitrary convex set $C\in\mathfrak{B}^q$ and all $\varepsilon>0$ and $\delta>0$.

2.3. **Proof of Theorem 2.1.** We prove that the distribution function $F_T(y,\theta)$ of the random vector $\hat u_T(\theta)=d_T(\theta)(\hat\theta_T-\theta)$ converges pointwise as $T\to\infty$ to the Gaussian distribution function $\Phi_{0,\sigma(\theta)}(y)$.

First we show that

$$\Delta_T(r)=\mathsf{P}\{\|\hat u_T-\tilde u_T\|>r\}\to 0,\qquad T\to\infty, \tag{28}$$

for an arbitrary $r>0$.

Consider the random event

$$A_T=\{\|\tilde u_T\|\in v^c(R-r)\},$$

where R is such that $\mathsf{P}(\bar A_T)\le\varepsilon/3$ for a fixed number $\varepsilon>0$ if $T>T_0$. This property holds in view of the asymptotic normality of $\tilde u_T$.

We also introduce the random event $B_T=\{\sup_{u\in v^c(\mathbb{R})}\|\Lambda_T(\theta)(M_T(u)-L_T(u))\|\le r\}$. Condition **B2** and Corollary 2.1 imply that

$$\mathsf{P}(\bar B_T)\le\mathsf{P}\left\{\sup_{u\in v^c(\mathbb{R})}\|M_T(u)-L_T(u)\|>\lambda_* r\right\}\le\frac{\varepsilon}{3}$$

for $T>T_0$.

Taking into account condition **D1**, consider the event C_T consisting in those elementary random events for which the M-estimator $\hat u_T$ is a unique solution of the system of equations (6) such that $\mathsf{P}(\bar C_T)\le\varepsilon/3$ for $T>T_0$. Thus, if $T>T_0$, then

$$\mathsf{P}(A_T\cap B_T\cap C_T)\ge 1-\varepsilon. \tag{29}$$

We derive from (21) and (27) that $\Lambda_T(\theta)L_T(u)=\tilde u_T-u$. If the event $A_T\cap B_T\cap C_T$ occurs, then

$$\|u+\Lambda_T(\theta)M_T(\theta)\|\le\|\tilde u_T\|+\|\Lambda_T(\theta)(M_T(u)-L_T(u))\|\le(R-r)+r=R$$

for $u\in v^c(\mathbb{R})$; that is, $F_T(u)=u+\Lambda_T(\theta)M_T(u)$ is a continuous mapping acting from $v^c(R)$ to $v^c(R)$.

Now we apply the Brouwer fixed point theorem (Theorem 2.3) to $F_T(u)$. Thus, we prove the existence of a point $u_T^0\in v^c(\mathbb{R})$ such that $F(u_T^0)=u_T^0$. In other words, $M_T(u_T^0)=0$, since $\Lambda_T(\theta)$ is nondegenerate. Since the event C_T occurs, the normalized M-estimator $\hat u_T$ is a unique solution of the system of equations (6) in the ball $v^c(R)$.

Therefore, $A_T\cap B_T\cap C_T\subset\{\hat u_T\in v^c(\mathbb{R})\}$ and $\mathsf{P}\{\hat u_T\in v^c(\mathbb{R})\}\ge 1-\varepsilon$.

Note that

$$\begin{aligned}1-\varepsilon&\le\mathsf{P}\left\{\{\hat u_T\in v^c(R)\}\cap B_T\right\}\\&\le\mathsf{P}\{\|\Lambda_T(\theta)(M_T(\hat u_T)-L_T(\hat u_T))\|\le r\}=\mathsf{P}\{\|\tilde u_T-\hat u_T\|\le r\}\end{aligned} \tag{30}$$

for $T>T_0$, since inequality (29) holds. This means that relation (28) holds, as well.

Put

$$\Pi(-\infty;y\pm\vec\varepsilon)=(-\infty;y_1\pm\varepsilon)\times\cdots\times(-\infty;y_q\pm\varepsilon),\qquad\varepsilon\ge 0.$$

Taking into account relation (28), we obtain for the distribution function

$$F_T(y,\theta)=\mathsf{P}\{\hat u_T\in\Pi(-\infty,y)\}$$

that

$$\mathsf{P}\{\tilde{u}_T \in \Pi(-\infty; y - \vec{\varepsilon})\} - \Delta_T(\varepsilon) \le F_T(y,\theta) \le \mathsf{P}\{\tilde{u}_T \in \Pi(-\infty; y + \vec{\varepsilon})\} + \Delta_T(\varepsilon) \tag{31}$$

for all $y \in \mathbb{R}^q$ and an arbitrary $\varepsilon > 0$.

Now Theorem 2.2 implies that the random vector $\tilde{u}_T$ is asymptotically normal as $T \to \infty$ with parameters 0 and $\sigma(\theta)$, where $\sigma(\theta)$ is defined by equality (4). Thus, we obtain that

$$\left|\mathsf{P}\{\tilde{u}_T \in \Pi(-\infty; y + \vec{\varepsilon})\} - \Phi_{0,\sigma(\theta)}(y \pm \vec{\varepsilon})\right| \to 0, \qquad T \to \infty. \tag{32}$$

Let $\varphi(y,\theta)$ be the Gaussian probability density corresponding to the distribution function $\Phi_{0,\sigma(\theta)}(y)$.

Since $\lambda_{\min}(\sigma(\theta)) = \underline{\lambda} > 0$ and $\lambda_{\max}(\sigma(\theta)) = \overline{\lambda} < +\infty$, we get

$$\varphi(y,\theta) \le (2\pi\underline{\lambda})^{-q/2} \exp\{-\|y\|^2/2\overline{\lambda}\} = \nu(\|y\|).$$

If $A = \Pi(-\infty, y)$, then $A_{-\varepsilon} = \Pi(-\infty, y - \vec{\varepsilon}]$ and $(\Pi(-\infty, y + \vec{\varepsilon}])_{-\varepsilon} = \Pi(-\infty, y] = A^c$.

We apply Theorem 2.4 to $\nu(\|y\|)$. Then

$$\left|\Phi_{0,\sigma(\theta)}(y) - \Phi_{0,\sigma(\theta)}(y + \vec{\omega})\right| = \int_{\Pi} \varphi(y,\theta)\, dy \le b \left(\frac{2\pi^{q/2}}{\Gamma(\frac{q}{2})}\right) \cdot |\omega|$$

for all $\omega \ne 0$, where

$$\Pi = \begin{cases} \Pi(-\infty, y + \overrightarrow{\omega}) \setminus A^c, & \text{if } \omega > 0, \\ A \setminus A_\omega, & \text{if } \omega < 0. \end{cases}$$

For all $y \in \mathbb{R}^q$ and an arbitrary $\varepsilon > 0$,

$$\begin{aligned} F_T(y,\theta) - \Phi_{0,\sigma(\theta)}(y) \le \Delta_T(\varepsilon) &+ \left|\mathsf{P}\{\widetilde{u}_T \in \Pi(-\infty, y + \overrightarrow{\varepsilon})\} - \Phi_{0,\sigma(\theta)}(y + \overrightarrow{\varepsilon})\right| \\ &+ \left|\Phi_{0,\sigma(\theta)}(y + \overrightarrow{\varepsilon}) - \Phi_{0,\sigma(\theta)}(y)\right|; \end{aligned} \tag{33}$$

$$\begin{aligned} \Phi_{0,\sigma(\theta)}(y) - F_T(y,\theta) \le \Delta_T(\varepsilon) &+ \left|\Phi_{0,\sigma(\theta)}(y - \overrightarrow{\varepsilon}) - \mathsf{P}\{\widetilde{u}_T \in \Pi(-\infty, y - \overrightarrow{\varepsilon})\}\right| \\ &+ \left|\Phi_{0,\sigma(\theta)}(y) - \Phi_{0,\sigma(\theta)}(y - \overrightarrow{\varepsilon})\right|. \end{aligned} \tag{34}$$

Relations (28)–(34) imply

$$|F_T(y,\theta) - \Phi_{0,\sigma(\theta)}(y)| \to 0, \qquad T \to \infty.$$

Theorem 2.1 is proved.

3. Asymptotic uniqueness of M-estimators

We find sufficient conditions for **D1**, that is, for the asymptotical uniqueness in probability of M-estimators for parameters of models (1). If the regression function as well as the loss function is differentiable, an M-estimator $\hat{\theta}_T$ satisfies the system of equations

$$\nabla Q_T(\tau) = 0. \tag{35}$$

Some of further conditions are, in fact, certain modifications of assumptions listed in Section 2.1.

We have

$$\widetilde{J}_T(\theta) = (\widetilde{J}_{il,T}(\theta))_{i,l=1}^q,$$

$$\widetilde{J}_{il,T}(\theta) = T^{-1} \int_0^T g_i(t,\theta) g_l(t,\theta)\, dt, \qquad i,l = 1,\dots,q.$$

B2′. $\lambda_{\min}(\widetilde{J}_T(\theta)) \ge \widetilde{\lambda}_* $ for some $\widetilde{\lambda}_* > 0$ and $T > T_0$.

B4. (i) $\sup_{t\ge 0}\sup_{\tau\in\Theta^c}|g_i(t,\tau)|\le k(i)<\infty$;

(ii) $\sup_{t\ge 0}\sup_{\tau\in\Theta^c}|g_{il}(t,\tau)|\le k(i,l)<\infty$;

(iii) for $T>T_0$, $i,l=1,\dots,q$, and $\tau_1,\tau_2\in\Theta^c$,

$$T^{-1}\Phi_T^{il}(\tau_1,\tau_2)=T^{-1}\int_0^T (g_{il}(t,\tau_1)-g_{il}(t,\tau_2))^2\,dt\le k_{il}\|\tau_1-\tau_2\|^2.$$

We also assume that

C3. $\sup_{x\in\mathbb{R}}|\psi(x)|=k_\psi<\infty$ and $\sup_{x\in\mathbb{R}}|\psi'(x)|=k_{\psi'}<\infty$.

Finally, we assume that an M-estimator is such that:

D2. $\hat\theta_T$ is a weakly consistent estimator of θ; that is,

$$\mathsf{P}\left\{\|\hat\theta_T-\theta\|\ge r\right\}\to 0,\qquad T\to\infty,$$

for an arbitrary $r>0$.

Some sufficient conditions for the consistency of M-estimators of parameters for nonlinear regression models are obtained in [17, 15, 21, 19].

Theorem 3.1. *Let conditions* **A1***,* **A2***,* **B2**$'$*,* **B4***,* **C1***,* **C3***, and* **D2** *hold. Then, given an arbitrary $\varepsilon>0$, there exists a number $T_0=T_0(\varepsilon)$ such that the system of equations* (35) *possesses a unique solution for $T>T_0$ with probability that is not less than $1-\varepsilon$.*

Proof. Put

$$H(t;\tau,\theta)=g(t,\tau)-g(t,\theta),\qquad H_i(t;\tau,\theta)=g_i(t,\tau)-g_i(t,\theta),$$
$$H_{il}(t;\tau,\theta)=g_{il}(t,\tau)-g_{il}(t,\theta),\qquad i,l=1,\dots,q,$$
$$G_T(\tau)=\left(G_T^{il}(\tau)\right)_{i,l=1}^q=\left(\gamma\frac{\partial^2}{\partial\tau_i\partial\tau_l}Q_T(\tau)\right)_{i,l=1}^q.$$

The proof of the theorem follows if the Hesse matrix $G_T(\tau)$ of the functional $\gamma Q_T(\tau)$ is positive definite in some neighborhood of the true value of the parameter θ with probability that approaches unity as $T\to\infty$.

For all $i,l=1,\dots,q$,

$$\begin{aligned}G_T^{il}(\tau)&=\gamma T^{-1}\int_0^T\psi'(X(t)-g(t,\tau))g_i(t,\tau)g_l(t,\tau)\,dt\\&\quad-\gamma T^{-1}\int_0^T\psi(X(t)-g(t,\tau))g_{il}(t,\tau)\,dt\\&=G_1^{il}(\tau)+G_2^{il}(\tau).\end{aligned}\tag{36}$$

Consider the second term in (36),

$$\begin{aligned}G_2^{il}(\tau)&=-\gamma T^{-1}\int_0^T\left[\psi(G(\xi(t))-H(t;\tau,\theta))-\psi\big(G(\xi(t))\big)\right]g_{il}(t,\tau)\,dt\\&\quad-\gamma T^{-1}\int_0^T\psi\big(G(\xi(t))\big)H_{il}(t;\tau,\theta)\,dt-\gamma T^{-1}\int_0^T\psi\big(G(\xi(t))\big)g_{il}(t,\theta)\,dt\\&=G_3^{il}(\tau)+G_4^{il}(\tau)+G_5^{il}.\end{aligned}$$

In view of condition **B4**(i),

$$|H(t;\tau,\theta)|=\left|\sum_{i=1}^q g_i(t,\tau_t^*)(\tau_i-\theta_i)\right|\le\|\overline{k}\|\cdot\|\tau-\theta\|,\tag{37}$$

where $\tau_t^* = \theta + \eta(\tau - \theta)$ and $\eta = \eta_t \in (0,1)$, $\overline{k} = (k(1), \dots, k(q))$. Moreover,

$$\psi(G(\xi(t)) - H(t;\tau,\theta)) - \psi\big(G(\xi(t))\big) = \psi'(G(\xi(t)) - \delta_t H(t;\tau,\theta))H(t;\tau,\theta) \tag{38}$$

for some $\delta_t \in (0,1)$.

Taking into account conditions **C3**, **B4**(ii) and relations (37) and (38), we conclude that

$$|G_3^{il}(\tau)| \le k_{\psi'} k(i,l)\gamma \|\overline{k}\| \cdot \|\tau - \theta\|. \tag{39}$$

Then

$$|G_4^{il}(\tau)| \le \gamma \left(T^{-1} \int_0^T \psi^2\big(G(\xi(t))\big)\,dt \right)^{1/2} \cdot \big(T^{-1}\Phi_T^{il}(\tau,\theta)\big)^{1/2} \le \gamma k_\psi k_{il}^{1/2} \|\tau - \theta\| \tag{40}$$

by condition **B4**(iii).

Now (13) and (26) imply that

$$\mathsf{E}\left(G_5^{il}\right)^2 \le \gamma^2 k^2(i,l)\,\mathsf{E}\,\psi^2\big(G(\xi(0))\big) \cdot T^{-2} \int_0^T\!\!\int_0^T |B(t-s)|\,dt\,ds = O(T^{-\alpha}),$$

whence

$$\left|G_5^{il}\right| \xrightarrow{\mathsf{P}} 0, \qquad T \to \infty. \tag{41}$$

Inequalities (39)–(41) show that

$$\left|G_2^{il}(\tau)\right| \le \gamma(k_{\psi'} k(i,l)\|\overline{k}\| + k_\psi k^{1/2}(i,l))\|\tau - \theta\| + \left|G_5^{il}\right| = K_{il}^{(2)}\|\tau - \theta\| + \left|G_5^{il}\right|. \tag{42}$$

On the other hand,

$$\begin{aligned} G_1^{il}(\tau) &= \gamma T^{-1} \int_0^T \left[\psi'(G(\xi(t)) - H(t;\tau,\theta)) - \psi'\big(G(\xi(t))\big)\right] g_i(t,\tau) g_l(t,\tau)\,dt \\ &\quad + \gamma T^{-1} \int_0^T \psi'\big(G(\xi(t))\big) \cdot [g_i(t,\tau)H_l(t;\tau,\theta) + g_l(t,\theta)H_i(t;\tau,\theta)]\,dt \\ &\quad + \gamma T^{-1} \int_0^T \left[\psi'\big(G(\xi(t))\big) - \mathsf{E}\,\psi'\big(G(\xi(t))\big)\right] g_i(t,\theta) g_l(t,\theta)\,dt + \widetilde{J}_T^{il}(\theta) \\ &= G_6^{il}(\tau) + G_7^{il}(\tau) + G_8^{il} + \widetilde{J}_T^{il}(\theta). \end{aligned}$$

We obtain from **C1**(iii) and (37) that

$$|G_6^{il}(\tau)| \le Lk(i)k(l)\gamma\|\overline{k}\| \cdot \|\tau - \theta\|. \tag{43}$$

Similarly to the proof of inequality (40),

$$|G_7^{il}(\tau)| \le \gamma k_{\psi'}(k(i)\|\overline{k}_l\| + k(l))\|\overline{k}_i\| \cdot \|\tau - \theta\|, \tag{44}$$

where $\overline{k}_i = (k(i,1), \dots, k(i,q))$, $i = 1, \dots, q$.

Finally, we derive from (12) that

$$\mathsf{E}\left(G_8^{il}\right)^2 \le \gamma^2 k^2(i) k^2(l) D\psi'\big(G(\xi(0))\big) \cdot T^{-2} \int_0^T\!\!\int_0^T |B(t-s)|\,dt\,ds = O(T^{-\alpha}), \tag{45}$$

whence

$$\left|G_8^{il}\right| \xrightarrow{\mathsf{P}} 0, \qquad T \to \infty. \tag{46}$$

Now we find from (43)–(46) that

$$\begin{aligned} \left|G_1^{il}(\tau) - J_T^{il}(\theta)\right| &\le \gamma\left(Lk(i)k(l)\|\overline{k}\| + k_{\psi'}\left(k(i)\|\overline{k}_l\| + k(l)\|\overline{k}_i\|\right)\right)\|\tau - \theta\| + \left|G_8^{il}\right| \\ &= K_{il}^{(1)}\|\tau - \theta\| + \left|G_8^{il}\right|. \end{aligned} \tag{47}$$

Using a property of eigenvalues of the sum of two symmetric matrices (see, for example, [33, pp. 101–103]), we conclude that

$$\begin{aligned}\left|\lambda_{\min}(G(\tau)) - \lambda_{\min}(\widetilde{J}_T(\theta))\right| &\le q \max_{1\le i,l\le q}\left|G_T^{il}(\tau) - \widetilde{J}_T^{il}(\theta)\right| \\ &\le q\left(\max_{1\le i,l\le q}\left|G_1^{il}(\tau) - \widetilde{J}_T^{il}(\theta)\right| + \max_{1\le i,l\le q}\left|G_2^{il}(\tau)\right|\right).\end{aligned} \tag{48}$$

Let $r = \widetilde{\lambda}_*/4q$, where $\widetilde{\lambda}_*$ is the number in condition **B2′**. If the random event

$$\Omega_r = \left\{\max_{1\le i,l\le q}\left(\left|G_5^{il}\right| + \left|G_8^{il}\right|\right) < r; \|\hat{\theta}_T - \theta\| \le \frac{r}{R}\right\}$$

occurs, where $R = \max_{1\le i,l\le q}\left(K_{il}^{(1)} + K_{il}^{(2)}\right)$, and $K_{il}^{(1)}$ and $K_{il}^{(2)}$ are the constants involved in relations (47) and (42), respectively, then inequality (48) yields

$$\begin{aligned}\mathsf{P}(\Omega_r) &\le \mathsf{P}\{|\lambda_{\min}(G(\hat{\theta}_T)) - \lambda_{\min}(\widetilde{J}_T(\theta))| \le 2qr\} \\ &\le \mathsf{P}\left\{\lambda_{\min}(G(\hat{\theta}_T)) - \lambda_{\min}(\widetilde{J}_T(\theta)) \ge -\frac{\lambda_*}{2}\right\} \le \mathsf{P}\left\{\lambda_{\min}(G(\hat{\theta}_T)) \ge \frac{\lambda_*}{2}\right\}\end{aligned}$$

for $T > T_0$ in accordance with condition **B2′**. For all $\varepsilon > 0$ and $T > T_0$, relations (42) and (47) and condition **D2** imply that $\mathsf{P}(\overline{\Omega}_r) < \varepsilon$. Hence, $\mathsf{P}(\Omega_r) \ge 1 - \varepsilon$ for $T > T_0$. This means that $\hat{\theta}_T$ is a unique solution of the system of equation (35) with probability that is not less than $(1 - \varepsilon)$, since the Hesse matrix $G_T(\tau)$ of the functional $\gamma Q_T(\tau)$ is positive definite in some neighborhood of the point θ with probability tending to unity as $T \to \infty$. □

Bibliography

1. P. J. Huber, *Robust regression: asymptotics, conjectures and Monte-Carlo*, Ann. Statist. **1** (1973), no. 5, 799–821. MR0356373
2. P. J. Huber, *Robust Statistics*, John Wiley & Sons, Inc., New York, 1981. MR606374
3. F. R. Hampel, E. M. Ronchetti, P. J. Rousseeuw, and W. A. Stahel, *Robust Statistics. The Approach Based on Influence Functions*, Wiley, New York, 1986. MR829458
4. H. L. Koul, *M-estimators in linear models with long range dependent errors*, Stat. Probab. Lett. **14** (1992), 153–164. MR1173413
5. H. L. Koul, *Asymptotics of M-estimations in non-linear regression with long-range dependence errors*, Proc. Athens Conf. Appl. Probabl. and Time Ser. Analysis (P. M. Robinson and M. Rosenblatt, eds.), Springer-Verlag Lecture Notes in Statistics, vol. II, 1996, pp. 272–291. MR1466752
6. H. L. Koul and K. Mukherjee, *Regression quantiles and related processes under long range dependent errors*, J. Multiv. Anal. **51** (1994), 318–337. MR1321301
7. L. Giraitis, H. L. Koul, and D. Surgailis, *Asymptotic normality of regression estimators with long memory errors*, Stat. Probab. Lett. **29** (1996), 317–335. MR1409327
8. H. L. Koul and D. Surgailis, *Asymptotic expansion of M-estimators with long memory errors*, Ann. Statist. **25** (1997), 818–850. MR1439325
9. H. L. Koul and D. Surgailis, *Second order behavior of M-estimators in linear regression with long-memory errors*, J. Statist. Plann. Inference **91** (2000), 399–412. MR1814792
10. H. L. Koul and D. Surgailis, *Robust estimators in regression models with long memory errors*, Theory and Application of Long-Range Dependence (P. Doukhan, G. Pantheism, and M. S. Taqqu, eds.), Birkhäuser, Boston, 2003, pp. 339–353. MR1957498
11. L. Giraitis and H. L. Koul, *Estimation of the dependence parameter in linear regression with long-range dependent errors*, Stat. Probab. Lett. **29** (1996), 317–335. MR1409327
12. R. T. Baillie, H. L. Koul, and D. Surgailis, *Regression model fitting with a long memory covariance process*, Economic Theory **20** (2004), 485–512. MR2061725
13. A. V. Ivanov and N. N. Leonenko, *Asymptotic behavior of M-estimators in continuous-time non-linear regression with long-range dependent errors*, Random Oper. Stoch. Equ. **10** (2002), no. 3, 201–222. MR1923424

14. A. V. Ivanov and N. N. Leonenko, *Robust estimators in nonlinear regression models with long-range dependence*, Optimal Design and Related Areas in Optimization and Statistics (L. Pronzato and A. Zhigljavsky, eds.), Springer, Berlin, 2009, pp. 193–221. MR2513352
15. A. V. Ivanov, *Asymptotic properties of L_p-estimators*, Theory Stoch. Process. **14(30)** (2008), no. 1, 60–68. MR2479706
16. A. V. Ivanov and I. V. Orlovsky, *L_p-estimates in nonlinear regression with long-range dependence*, Theory Stoch. Process. **7(23)** (2002), no. 3–4, 38–49.
17. A. V. Ivanov and I. V. Orlovskyi, *Consistency of M-estimators in nonlinear regression models with continuous time*, Naukovi visti NTUU "KPI" **4(42)** (2005), 140–147. (Ukrainian)
18. A. V. Ivanov and I. V. Orlovskyi, *The uniqueness of M-estimators of parameters in nonlinear regression models*, Naukovi Visti NTUU "KPI" **4(66)** (2009), 135–141. (Ukrainian)
19. I. N. Savych, *Consistency of quantile estimators in regression models with long-range dependent noise*, Teor. Ĭmovir. Mat. Stat. translation in **82** (2010), 128–136; English transl. in Theory Probab. Math. Statist. **82** (2011), 129–138. MR2790488
20. I. V. Orlovsky, *M-estimates in nonlinear regression with weak dependence*, Theory Stoch. Process. **9(25)** (2003), no. 1–2, 108–122. MR2080017
21. A. V. Ivanov and I. V. Orlovsky, *Consistency of M-estimates in general nonlinear model*, Theory Stoch. Process. **13(29)** (2007), no. 1–2, 86–97. MR2343813
22. I. V. Orlovsky, *Asymptotic properties of M-estimators of parameters in nonlinear regression models*, Abstract of candidate thesis (Physics and Mathematics), Kyiv National Taras Shevchenko University, Kyiv, 2007. (Ukrainian)
23. A. V. Ivanov, N. N. Leonenko, M. D. Ruiz-Medina, and I. N. Savych, *Limit theorems for weighted non-linear transformations of Gaussian processes with singular spectra*, Ann. Probab. **41** (2013), no. 2, 1088–1114. MR3077537
24. Yu. V. Goncharenko and S. I. Lyashko, *Brouwer's Theorem*, "KIĬ", Kiev, 2000. (Russian)
25. V. V. Anh, V. P. Knopova, and N. N. Leonenko, *Continuous-time stochastic processes with cyclical long-range dependence*, Aust. NZ J. Stat. **46** (2004), 275–296. MR2076396
26. A. V. Ivanov, N. N. Leonenko, M. D. Ruiz-Medina, B. M. Zhurakovsky, *Estimation of harmonic component in regression with cyclically dependent errors*, Statistics: A Journal of Theoretical and Applied Statistics, **49** (2015), no. 1, 156–186. MR3304373
27. U. Grenander, *On the estimation of regression coefficients in the case of an autocorrelated disturbance*, Ann. Statist **25** (1954), no. 2, 252–272. MR0062402
28. I. A. Ibragimov and Yu. A. Rozanov, *Gaussian Random Processes*, "Nauka", Moscow, 1970; English transl. Springer-Verlag, New York–Berlin, 1978. MR543837
29. P. Billingsley, *Convergence of probability measures*, John Wiley & Sons, Inc., New York–London–Sydney, 1968. MR0233396
30. A. V. Ivanov and I. M. Savych, *The μ-admissibility of a spectral density for long range dependent random noise in nonlinear regression models*, Naukovi visti NTUU "KPI" **1** (2009), 143–148. (Ukrainian)
31. N. N. Leonenko and A. V. Ivanov, *Statistical Analysis of Random Fields*, "Vyshcha Shkola", Kiev, 1986; English transl., Kluwer Academic Publishers Group, Dordrecht, 1989.
32. R. N. Bhattacharya and R. Ranga Rao, *Normal Approximation and Asymptotic Expansions*, John Wiley & Sons, New York–London–Sydney, 1976.
33. J. H. Wilkinson, *The Algebraic Eigen Value Problem*, Clarendon Press, Oxford, 1962.

Department of Mathematical Analysis and Probability Theory, National Technical University of Ukraine "KPI", Peremogy Avenue, 37, Kyiv 03056, Ukraine
E-mail address: alexntuu@gmail.com

Department of Mathematical Analysis and Probability Theory, National Technical University of Ukraine "KPI", Peremogy Avenue, 37, Kyiv 03056, Ukraine
E-mail address: i.v.orlovsky@gmail.com

Received 05/AUG/2015
Translated by S. KVASKO