ISSN 0005-1179, Automation and Remote Control, 2010, Vol. 71, No. 2, pp. 257–274. © Pleiades Publishing, Ltd., 2010.
Original Russian Text © A.V. Kitaeva, G.M. Koshkin, 2010, published in Avtomatika i Telemekhanika, 2010, No. 2, pp. 92–111.

ESTIMATION AND FILTERING

Semi-Recursive Nonparametric Identification in the General Sense of a Nonlinear Heteroscedastic Autoregression[1]

A. V. Kitaeva* and G. M. Koshkin**

*Tomsk Polytechnical University, Tomsk, Russia
**Tomsk State University, Tomsk, Russia

Received March 2, 2009

Abstract—We consider semi-recursive kernel estimates of conditional mean, volatility function, and sensitivity function for a nonlinear heteroscedastic autoregression. We find the principal parts of mean square errors for these estimates.

DOI: 10.1134/S0005117910020086

1. INTRODUCTION. PROBLEM SETTING

Suppose that a scalar sequence $(X_t)_{t=\ldots,-1,0,1,2,\ldots}$ is generated by a nonlinear heteroscedastic autoregression of order m

$$X_t = \Psi(U_t) + \Phi(U_t)\xi_t, \tag{1}$$

where $U_t = (X_{t-i_1}, \ldots, X_{t-i_m})$, $1 \le i_1 < i_2 < \ldots < i_m$ is a known subsequence of natural numbers, (ξ_t) is a sequence of independent identically distributed (with density positive on R^1) random variables with zero mean, unit variance, zero third, and finite fourth moments, Ψ and $\Phi > 0$ are unknown nonperiodic functions bounded on compacts, where the function $\Phi^2(x) = \mathsf{D}(X_t \mid U_t = x)$ is the conditional variance (volatility function). Note that i_m may be large, while m is small. Models of type (1)—CHARH (conditional heteroskedastic autoregressive nonlinear)—find many applications in financial time series analysis, e.g., in analyzing currency exchange rates [1–3], where the volatility function characterizes the corresponding financial risks.

By identifying a model (1) in a general sense [4, 5] we mean the problem of nonparametric estimation of the functions Ψ and Φ and the derivatives of Ψ that show how much input variables in (1) influence the output variable.

We will assume that function $\Psi(x)$ and $\Phi(x)$, $x \in \mathsf{R}^m$, satisfy, in addition, the following conditions (3.2 (c) [1]). First, there exist vectors $a = (a_1, \ldots, a_m)$ and $d = (d_1, \ldots, d_m)$, $d_i \ge 0$, $i = 1, \ldots, m$ (perhaps, both of them zero) such that for $\|x\| \to \infty$

$$\Psi(x) = \sum_{i=1}^{m} a_i x_i + o(\|x\|),$$

$$\Phi(x) = \sum_{i=1}^{m} d_i x_i^2 + o(\|x\|^2).$$

[1] This work was supported by the Russian Foundation for Basic Research, project no. 09-08-00595-a.

Besides, we define an i_m-dimensional square matrix A as the zero matrix if $a = 0$ and

$$A = \begin{bmatrix} 0 & 0 & \dots & 0 & a'_1 \\ 1 & 0 & \dots & 0 & a'_2 \\ 0 & 1 & \dots & 0 & a'_3 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & a'_{i_m} \end{bmatrix},$$

otherwise and assume that $\rho(AA^{\mathrm{T}}) + \max\limits_i(d_i) < 1$ if $d \neq 0$, and $\rho(A) < 1$ if $d = 0$. Here ρ denotes the spectral radius of the matrix, $a'_{i_r} = a_r$, $r = \overline{1,m}$, $a'_{i_r} = 0$ otherwise, $\|x\|$ is the Euclidean norm of x, A^{T} is the transpose of A.

In this case, as shown in [1] (Lemma 3.1, p. 264, these conditions are also discussed there), the process $(X_t)_{t=\dots,-1,0,1,2,\dots}$ will be a strictly stationary process and will satisfy the strong mixing (s.m.) condition with s.m. coefficient

$$\alpha(\tau) \sim e^{-\delta\tau}, \quad \delta > 0, \quad \tau \to \infty. \tag{2}$$

Here and below the notation $f(x) \sim g(x)$ for $x \to \infty$ means that $\lim\limits_{x\to\infty} f(x)/g(x) = 1$.

Let $X_{1-i_m}, \dots, X_n$ be observations generated by an autoregression process (1). The conditional expectation $\Psi(x) = \mathsf{E}(X_t \mid U_t = x) = \mathsf{E}(X_t \mid x)$, $x \in \mathsf{R}^m$, we estimate by the statistic which is a semi-recursive counterpart of the Nadaraya–Watson estimate [6, 7]

$$\Psi_{n,m}(x) = \sum_{t=1}^{n} \frac{X_t}{h_t^m} \mathbf{K}\left(\frac{x - U_t}{h_t}\right) \Big/ \sum_{t=1}^{n} \frac{1}{h_t^m} \mathbf{K}\left(\frac{x - U_t}{h_t}\right). \tag{3}$$

We estimate the conditional variance $\Phi^2(x)$ with a statistic similar to (3):

$$\Phi^2_{n,m}(x) = \sum_{t=1}^{n} \frac{X_t^2}{h_t^m} \mathbf{K}\left(\frac{x - U_t}{h_t}\right) \Big/ \sum_{t=1}^{n} \frac{1}{h_t^m} \mathbf{K}\left(\frac{x - U_t}{h_t}\right) - \Psi^2_{n,m}(x). \tag{4}$$

The estimate (4) is a semi-recursive counterpart of a widely used volatility function estimate [8].

Estimates with this structure are called semi-recursive [9], because only the numerator and the denominator are computed recursively.

One can study how some variables influence others in stochastic systems with the s.m. power of sensitivity functions [10]. In particular, for the conditional expectation $\mathsf{E}(X_t \mid x)$ in the model (1) the sensitivity function on the variable x_j is defined as

$$T_j(x) = \frac{\partial \mathsf{E}(X_t \mid x)}{\partial x_j} = \frac{\partial \Psi(x)}{\partial x_j}. \tag{5}$$

Thus, the problem of identifying in the general sense a model (1) is a special case of estimating functions of the form

$$H(A) = H\Big(a(x),\, a^{(1j)}(x)\Big), \tag{6}$$

where $a(x) = \big(a_0(x), a_1(x), a_2(x)\big)$, $a^{(1j)}(x) = \big(a_0^{(1j)}(x), a_1^{(1j)}(x), a_2^{(1j)}(x)\big)$, $j = \overline{1,m}$, $x \in \mathsf{R}^m$, $H(\cdot)\colon \mathsf{R}^{3(m+1)} \to \mathsf{R}^1$ is a given function. To unify notation in what follows, we introduce the following notation: $a(x) \equiv a^{(0j)}(x)$. Functionals $a_i(x)$ and $a_i^{(1j)}(x)$ are defined as

$$a_i(x) = \int y^i f(x,y)dy, \quad a_i^{(1j)}(x) = \frac{\partial a_i(x)}{\partial x_j}, \quad i = 0,1,2, \quad j = \overline{1,m},$$

where $f(x,y)$ is the unknown probability density of the random vector $Z=(U_{t-1},X_t)$ in stationary mode.

Since $a_0(x)=\int f(x,y)dy=p\,(x)$, where $p\,(x)$ is the probability density of U_t, we can represent in the form (6) any function of the conditional moments

$$b_i(x)=a_i(x)/p\,(x)=\int y^i f(y\mid x)dy$$

and their derivatives $b_i^{(1j)}(x)=\dfrac{\partial b_i(x)}{\partial x_j}$, $i=1,2$, which is sufficient for our needs: for $H(a_0,a_1)=a_1/a_0$ we get the conditional expectation $\Psi(x)$, for $H(a_0,a_1,a_2)=a_2/a_0-(a_1/a_0)^2$—the volatility function $\Phi^2(x)=\mathsf{D}\,(Y\mid x)=b_2(x)-b_1^2(x)$, and for $H(a_0,a_1,a_0^{(1j)},a_1^{(1j)})=\dfrac{a_1^{(1j)}}{a_0}-\dfrac{a_1a_0^{(1j)}}{a_0^2}=b_1^{(1j)}$—the sensitivity function $T_{t-j}(x)$.

As recursive nonparametric estimates for the functionals $a^{(0j)}(x)\equiv a(x)$ and the derivatives $a^{(1j)}(x)$ at the point x we take the statistics

$$\begin{aligned}a_n^{(rj)}(x)&=\frac{1}{n}\sum_{i=1}^{n}\frac{(1,X_i,X_i^2)}{h_i^{m+r}}\mathbf{K}^{(rj)}\left(\frac{x-U_{i-1}}{h_i}\right)\\&=a_{n-1}^{(rj)}(x)-\frac{1}{n}\left[a_{n-1}^{(rj)}(x)-\frac{(1,X_n,X_n^2)}{h_n^{m+r}}\mathbf{K}^{(rj)}\left(\frac{x-U_{n-1}}{h_n}\right)\right],\end{aligned}$$

where $r=0,1$, $Z_l=(U_{l-1},X_l)$, $l=\overline{1,n}$ are $(m+1)$-dimensional observables characterized by the density $f(x,y)$, $\mathbf{K}^{(0j)}(u)\equiv\mathbf{K}(u)=\prod\limits_{i=1}^{m}K(u_i)$ is the m-dimensional multiplicative kernel; $\mathbf{K}^{(1j)}(u)=\dfrac{\partial\mathbf{K}(u)}{\partial u_j}=K(u_1)\ldots K(u_{j-1})K^{(1)}(u_j)K(u_{j+1})\ldots K(u_m)$, $K^{(1)}(u_j)=\dfrac{dK(u_j)}{du_j}$; (h_n) is a sequence of numbers; $a_n^{(rj)}(x)=\left(a_{0n}^{(rj)}(x),a_{1n}^{(rj)}(x),a_{2n}^{(rj)}(x)\right)$.

Semi-recursive kernel estimates for substituting conditional functionals $b(x)=(b_1(x),b_2(x))$ at the point x are given by

$$b_n(x)=\frac{\sum\limits_{i=1}^{n}\dfrac{(X_i,X_i^2)}{h_i^m}\mathbf{K}\left(\dfrac{x-U_{i-1}}{h_i}\right)}{\sum\limits_{i=1}^{n}\dfrac{1}{h_i^m}\mathbf{K}\left(\dfrac{x-U_{i-1}}{h_i}\right)}.\tag{7}$$

A semi-recursive (functionals $a_i(x)$ and their derivatives are computed recursively) substitution estimate for (6) looks like

$$H(A_n)=H\left(\left\{a_n^{(rj)}(x)\right\},\ j=\overline{1,m},\ r=0,1\right).\tag{8}$$

The problems of the instability of estimates (7) and, perhaps, (8) and related obstacles to finding the mean squared error (MSE) of the estimates we propose to solve via a piecewise smooth approximation [11].

Since the observations are dependent, studying the estimates' properties becomes much harder. For example, the main part of the asymptotic MSE Nadaraya–Watson estimate for s.m. sequences was found only in 1999 [12]; the authors also proved that this estimate converges with probability one.

In this paper, we propose a unified approach to the identification problem in the general sense for the model (1) based on the estimates $H(A_n)$ and their piecewise smooth approximations. We find the main parts of the MSE and determine the (enhanced due to kernel selection) mean square convergence speed for the considered estimates to $H(A)$.

2. ASYMPTOTIC PROPERTIES OF ESTIMATES OF THE FUNCTIONALS $a_i(x)$ AND THEIR DERIVATIVES

Since when we consider asymptotical bias and convergence speed of the estimates $a_n^{(rj)}(x)$, the dependence of sample values is irrelevant, then results shown in Lemmas 1–3 in [13] still hold (in this case, in the notation of [13] the vector function $g(y) = (1, y, y^2)$).

Thus, in conditions of Lemma 3 from [13], the following relation holds for the convergence speed of the estimates' deviations:

$$\left|\mathsf{b}\left(a_n^{(rj)}(x)\right) - S_\nu \omega_\nu^{(rj)}(x) h_n^\nu\right| = o(h_n^\nu), \tag{9}$$

where

$$\omega_\nu^{(rj)}(x) = \left(\omega_{0\nu}^{(rj)}(x), \omega_{1\nu}^{(rj)}(x), \omega_{2\nu}^{(rj)}(x)\right) = \frac{T_\nu}{\nu!}\sum_{l=1}^{m} \frac{\partial^\nu a^{(rj)}(x)}{\partial x_l^\nu}, \quad r = 0, 1, \quad j = \overline{1, m},$$

$$T_\nu = \int u^\nu K(u) du, \quad \nu = 1, 2, \ldots.$$

The constant S_ν appears since the estimates are recursive; it reflects the constraints imposed upon the sequence (h_n) (cf. Definition 5 in [13]):

$$\frac{1}{n}\sum_{i=1}^{n} h_i^\alpha = S_\alpha h_n^\alpha + o(h_n^\alpha); \tag{10}$$

by (10), for (9) $\alpha = \nu$.

We introduce the following notation:

$$a_i^+(x) = \int |y^i|\, f(x, y)\, dy,$$

$$a^+_{1(1+\tau),tp}(x, y) = \int\limits_{\mathsf{R}^2} |v^t q^p| f_{1(1+\tau)}(x, v, y, q)\, dv\, dq,$$

where $f_{1(1+\tau)}$ is the $2(m + 1)$-dimensional density of the sample values distribution $(Z_1, Z_{1+\tau})$, $\tau \geqslant 1$. The event that $(X_j)_{j\geqslant 1}$ satisfies the s.m. condition with s.m. coefficient $\alpha(\tau)$ we denote by $(X_j)_{j\geqslant 1} \in \mathcal{S}(\alpha)$. We define the necessary classes of functions introduced in Definitions 1–4 in [13].

A function $H(\cdot)\colon \mathsf{R}^s \to \mathsf{R}^1$ belongs to the class $\mathcal{N}_\nu(x)$ ($H(\cdot) \in \mathcal{N}_\nu(x)$) if the function and all its partial derivatives (of order up to and including ν) are continuous at point x. The function $H(\cdot) \in \mathcal{N}_\nu(\mathsf{R})$ if the conditions for $H(\cdot)$ are satisfied for all $x \in \mathsf{R}^s$.

A function $K(u)$ belongs to the class of normalized kernels $\mathcal{A}^{(r)}$, $r = 0, 1$, if $\int |K^{(r)}(u)|\, du < \infty$, $\int K(u)du = 1$. A function $K(\cdot) \in \mathcal{A}_\nu^{(r)}$ if $K(\cdot) \in \mathcal{A}^{(r)}$ and $K(u)$ satisfies conditions $\int |u^\nu K(u)| du < \infty$, $T_j = 0$, $j = 1, \ldots, \nu - 1$, $T_\nu \neq 0$, $K(u) = K(-u)$.

For simplicity, we denote $\mathcal{A}^{(0)} = \mathcal{A}$, $\mathcal{A}_\nu^{(0)} = \mathcal{A}_\nu$.

Lemma 1 (covariance of the estimates $a_{tn}^{(rj)}(x)$ and $a_{pn}^{(qk)}(x)$). *Suppose that the index θ takes values t and p, the index γ takes values r and q, and the following conditions hold*:

(1) $(Z_j)_{j\geqslant 1} \in \mathcal{S}(\alpha)$, $\int\limits_0^\infty [\alpha(\tau)]^\lambda d\tau < \infty$ *for some* $\lambda \in (0,1)$;

(2) *the functions* $a_{t+p}(\cdot) \in \mathcal{N}_0(\mathsf{R})$, $a_\theta(\cdot) \in \mathcal{N}_0(\mathsf{R})$, $a^+_{\frac{2\theta}{1-\lambda}}(\cdot) \in \mathcal{N}_0(x)$;

(3) $\sup\limits_x a^+_\theta(x) < \infty$, $\sup\limits_x a^+_{\frac{2\theta}{1-\lambda}}(x) < \infty$; $\sup\limits_x a^+_{t+p}(x) < \infty$,

(4) $K(\cdot) \in \mathcal{A}^{(\gamma)}$, $\sup\limits_{u\in\mathsf{R}^1} |K^{(\gamma)}(u)| < \infty$; $\sup\limits_{u\in\mathsf{R}^1} |K(u)| < \infty$ *for* $m > 1$, $rq = 1$;

(5) *for a monotone nonincreasing sequence* (h_n) *it holds that*

$$(h_n + 1/(nh_n^{m(1+\lambda)+r+q})) \downarrow 0.$$

Then for $n \to \infty$

$$\left|\operatorname{cov}\left(a_{tn}^{(rj)}(x), a_{pn}^{(qk)}(x)\right)\right| \leqslant \frac{24}{nh_n^{m(1+\lambda)+r+q}} \left[a^+_{\frac{2t}{1-\lambda}}(x) a^+_{\frac{2p}{1-\lambda}}(x)\right]^{\frac{1-\lambda}{2}}$$

$$\times \left[\int\limits_{\mathsf{R}^m} \left|\mathbf{K}^{(r)}(u)\right|^{\frac{2}{1-\lambda}} du \int\limits_{\mathsf{R}^m} \left|\mathbf{K}^{(q)}(u)\right|^{\frac{2}{1-\lambda}} du\right]^{\frac{1-\lambda}{2}} \int\limits_0^\infty [\alpha(\tau)]^\lambda d\tau + o\left(\frac{1}{nh_n^{m(1+\lambda)+r+q}}\right). \tag{11}$$

If, moreover,

(6) $\lambda < 1/2$, *and for the sequence* (h_n) (10) *holds for* $\alpha = -(m+r+q)$;

(7) $\sup\limits_{x,y} a^+_{i(i+\tau),tp}(x,y) < C$, $a^+_{\frac{2\theta}{1-\lambda}}(\cdot) \in \mathcal{N}_0(x)$, $\sup\limits_x a^+_{\frac{2\theta}{1-\lambda}}(x) < \infty$, *then for* $n \to \infty$

$$\left|\operatorname{cov}\left(a_{tn}^{(rj)}(x), a_{pn}^{(qk)}(x)\right) - \frac{S_{-(m+r+q)}}{nh_n^{m+r+q}}\, a_{t+p}(x) \int K^{(r)}(u) K^{(q)}(u)\, du \left(\int K^2(u)\, du\right)^{m-1}\right|$$
$$= o\left(\frac{1}{nh_n^{m+r+q}}\right) \tag{12}$$

and, in particular, for $t = p$

$$\mathsf{D}\, a_{tn}^{(rj)}(x) \sim \frac{S_{-(m+2r)}}{nh_n^{m+2r}} a_{2t}(x) \int [K^{(r)}(u)]^2\, du \left(\int K^2(u)\, du\right)^{m-1}.$$

Theorem 1 of [13] and Lemma 1 of this paper imply Theorem 1.

Theorem 1 (MSE for optimal estimates of the functionals $a_t^{(rj)}(x)$). *Under the conditions of Lemma 3 in* [13], *conditions* (1)–(4) *and* (6), (7) *of Lemma 1 for* $\gamma = r$, $\theta = p = t$, *and, in addition,* $\omega_{t\nu}^{(rj)}(x) \neq 0$, *then for* $n \to \infty$

$$h_{tn}^{(rj)\,o} = \mathop{\arg\min}_{h_{tn}^{(rj)}>0} \mathsf{u}^2\left(a_{tn}^{(rj)}(x)\right)$$

$$\sim \left[\frac{(m+2r)(m+\nu+2r)\, a_{2t}(x)}{4n\nu(m+2\nu+2r)[\omega_{t\nu}^{(rj)}(x)]^2} \int [K^{(r)}(u)]^2\, du \left(\int K^2(u)\, du\right)^{m-1}\right]^{\frac{1}{m+2(\nu+r)}},$$

$$\mathsf{u}^2\left(a_{tn}^{(rj)}(x)\big|_{h_{tn}^{(rj)} = h_{tn}^{(rj)\,o}}\right) = \mathsf{u}^2\left(a_{tn}^{(rj)\,o}(x)\right) \sim (m+1+2r)\left[\frac{m+2(\nu+r)}{m+\nu+2r}\right]^{\frac{2(m+\nu+2r)}{m+2(\nu+r)}}$$

$$\times \left[\frac{a_{2t}(x)}{4n\nu} \int [K^{(r)}(u)]^2\, du \left(\int K^2(u)\, du\right)^{m-1}\right]^{\frac{2\nu}{m+2(\nu+r)}} \left[\frac{[\omega_{t\nu}^{(rj)}(x)]^2}{m+2r}\right]^{\frac{m+2r}{m+2(\nu+r)}}$$

$$= O\left(n^{-\frac{2\nu}{m+2(\nu+r)}}\right).$$

By Theorem 1, the convergence speed order for optimal non-parametric estimates $a_{tn}^{(rj)\,o}(x)$ for s.m. observations, which equals $\dfrac{2\nu}{m+2(\nu+r)}$, for large ν, just like the case of independent variables, tends to the usual convergence speed order of parametric estimates, which equals one.

We introduce the following notation:

$f_{1(i+1)(i+j+1)(i+j+k+1)}(z,s,u,w)$—distribution density for sample values $(Z_1, Z_{i+1}, Z_{i+j+1}, Z_{i+j+k+1})$,

$a^+_{1(i+1)(i+j+1)(i+j+k+1),\,t}(x,\,y,x',y') =$
$\int\limits_{\mathsf{R}^4} |vsv's'|^t f_{1(i+1)(i+j+1)(i+j+k+1)}(x,v,y,s,x',v',y',s')\,dvdsdv'ds'$, $1 \leqslant i,j,k < n$, $i+j+k \leqslant n-1$;

similarly, $a^+_{1(1+j)(1+j+k),\,t}(x,\,y,x') = \int\limits_{\mathsf{R}^3} |vsv'|^t f_{1(1+j)(1+j+k)}(x,v,y,s,x',v')\,dvdsdv'$,

$a^+_{1(i+1),\,t}(x,x') = \int\limits_{\mathsf{R}^2} |vs|^t f_{1(1+j)}(x,v,s,x')\,dvds$;

$M_4(a_{tn}) = \mathsf{E}\,[a_{tn}(x) - a_t(x)]^4$, $S_{tn} = a_{tn}(x) - \mathsf{E}a_{tn}(x)$.

Lemma 2 (convergence order for the fourth central moments of the estimates $a_{tn}^{(rj)}(x)$). *Suppose that for $r = 0$ (or 1):*

(1) $(Z_j) \in \mathcal{S}(\alpha)$ *and* $\int\limits_0^\infty \tau^2[\alpha(\tau)]^{\frac{\delta}{2+\delta}}d\tau < \infty$ *for some* $0 < \delta < 2$;

(2) $\sup\limits_{u\in\mathsf{R}^1} |K^{(r)}(u)| < \infty$, $\int |K^{(r)}(u)|\,du < \infty$;

(3) *for a monotone nonincreasing sequence* (h_n) *it holds that*

$$(h_n + 1/(nh_n^{m+2r})) \downarrow 0;$$

(4) $\sup\limits_x a^+_{t\beta}(x) < \infty$, $\beta = 0, 4$;

(5) $\sup\limits_x a^+_{1(i+1)(i+j+1)(i+j+k+1),t}(x,x,x,x) < \infty$, $\quad \sup\limits_x a^+_{1(1+j)(1+j+k),\,t(2+\delta)}(x,x,x) < \infty$, $\sup\limits_x a^+_{1(i+1),\,t(2+\delta)}(x,x) < \infty$.

Then for $n \to \infty$ correspondingly for $r = 0$ (or 1),

$$\mathsf{E}\left(S_{tn}^{(rj)}\right)^4 = O\left(\frac{1}{n^2h_n^{2(m+2r)}}\right).$$

Lemma 3 (convergence order for the fourth moments of deviations $M_4(a_{tn}^{(rj)})$). *If for $r = 0$ (or 1) conditions of Lemma 3 from* [13] *and Lemma 2 hold, then for $n \to \infty$*

$$M_4\left(a_{tn}^{(rj)}\right) = O\left(\frac{1}{n^2h_n^{2(m+2r)}} + h_n^{4\nu}\right).$$

The proofs of Lemmata 1 and 2 are given in the Appendix, Lemma 3 is proven just like Lemma 6 in [13].

3. MSE OF THE SUBSTITUTION ESTIMATES AND THEIR PIECEWISE SMOOTH APPROXIMATIONS

To find MSE, we use the results of [11].

For $r = 0, 1,\ j = \overline{1, m},\ i = 0, 1, 2$ we denote

$$H_{ijr} = \partial H(A)/\partial \left(a_i^{(rj)}\right),$$

$$Q = \begin{cases} \{0\}, & \text{if } \forall j \quad r = 0 \\ \{1\}, & \text{if } \forall j \quad r = 1 \\ \{0, 1\}, & \text{if } \exists j \quad r = 0 \bigwedge r = 1, \end{cases} \qquad \max(r) = \max_{r \in Q}(r).$$

For $H(A_n)$ we introduce the piecewise smooth approximation

$$\widetilde{H}(A_n, \delta_n) = \frac{H(A_n)}{(1 + \delta_n |H(A_n)|^{\tau})^{\rho}},$$

where $\tau > 0,\ \rho > 0,\ \rho\tau \geqslant 1,\ (\delta_n) \downarrow 0$ for $n \to \infty$.

Theorem 2 (MSE for the estimate of $H(A_n)$). *Suppose that for $t, p = 0, 1, 2$ and $r \in Q$:*

(1) $(Z_i) \in \mathcal{S}(\alpha)$ *and* $\int\limits_0^{\infty} \tau^2 [\alpha(\tau)]^{\frac{\delta}{2+\delta}} d\tau < \infty$ *for some* $0 < \delta < 2$;

(2) $a_{t+p}(\cdot) \in \mathcal{N}_0(\mathsf{R}),\ a^{+}_{t(2+\delta)}(\cdot) \in \mathcal{N}_0(x);\ \sup\limits_x a^{+}_{t+p}(x) < \infty,\ \sup\limits_x a^{+}_{t\beta}(x) < \infty,\ \beta = 0, 4$;

(3) $K(\cdot) \in \mathcal{A}_{\nu}^{(r)},\ \sup\limits_{u \in \mathsf{R}^1} |K^{(r)}(u)| < \infty$; *if* $1 \in Q$, *then* $\lim\limits_{|u| \to \infty} K(u) = 0$ *and for* $m > 1$ $\sup\limits_{u \in \mathsf{R}^1} |K(u)| < \infty$;

(4) $a_t^{(rj)}(\cdot) \in \mathcal{N}_{\nu}(\mathsf{R}),\ \sup\limits_x |a_t^{(rj)}(x)| < \infty;\ \sup\limits_x \left| \dfrac{\partial^{\nu} a_t^{(rj)}(x)}{\partial x_l \ldots \partial x_q} \right| < \infty,\ l, \ldots, q = \overline{1, m}$;

(5) *for a monotone nonincreasing sequence* (h_n) *it holds that*

$$(d_n) = \left(h_n + 1/\left(n h_n^{m + 2\max(r)}\right)\right) \downarrow 0,$$

and (10) *holds for* $\alpha = \nu,\ \alpha = -m - 2l,\ l = \overline{0, \max(r)}$;

(6) $\sup\limits_x a_{1(i+1)(i+j+1)(i+j+k+1),t}(x, x, x, x) < \infty, \quad \sup\limits_x a^{+}_{1(i+1)(i+j+1),t(2+\delta)}(x, x, x) < \infty,$
$\sup\limits_x a^{+}_{1(i+1),\,t(2+\delta)}(x, x) < \infty, \quad \sup\limits_{x,\,y} a^{+}_{1(i+1),\,tp}(x,\ y) < \infty$ *for all* $i, j, k \geqslant 1$;

(7) $H(\cdot) \in \mathcal{N}_2(A)$;

(8) *for all possible values of* $X_1, \ldots, X_n$ *the sequence* $\{|H(t_n)|\}$ *is dominated by the sequence of numbers* $(C_0 d_n^{\gamma}),\ (d_n) \uparrow \infty$, *where* C_0 *is a constant, and* $0 \leqslant \gamma \leqslant 1/4$.

Then

$$\mathsf{u}^2\left(H(A_n)\right) = \sum_{t,p=0}^{2} \sum_{r,q \in Q} \sum_{j,k} H_{tjr} H_{pkq} \left[\frac{S_{-(m+r+q)}}{n h_n^{m+r+q}} a_{t+p}(x) \int K^{(r)}(u) K^{(q)}(u) du \right.$$
$$\left. \times \left(\int K^2(u) du \right)^{m-1} + S_{\nu}^2 \omega_{t\nu}^{(rj)}(x) \omega_{p\nu}^{(qk)}(x) h_n^{2\nu} \right] + O\left(\left[\frac{1}{n h_n^{m + 2\max(r)}} + h_n^{2\nu} \right]^{\frac{3}{2}} \right).$$

Theorem 3 (MSE for the estimate of $\widetilde{H}(A_n, \delta_n)$). *Suppose that the conditions of Theorem 2 hold, except for condition* 8) *which is substituted by*

(8*) $H(A) \neq 0$ *or* $\tau = 4, 6, \ldots$.

Then for $n \to \infty$

$$\mathsf{u}^2\left(\widetilde{H}(A_n, \delta_n)\right) \sim \mathsf{u}^2\left(H(A_n)\right).$$

The optimal sequences of the fuzziness parameter and the corresponding MSE for the estimates of $H(A_n)$ and their piecewise smooth approximations $\widetilde{H}(A_n, \delta_n)$ for s.m. observations are found similar to Theorem 1.

The proof of Theorem 2 is given in the Appendix, Theorem 3 follows from Corollary 4 from [11] for $k = 2$ and $m = 4$.

4. IDENTIFYING SECOND ORDER AUTOREGRESSIONS IN THE GENERAL SENSE. REAL DATA ANALYSIS

Let us go back to the autoregression process (1). To clarify the presentation, we consider the special case $m = 2$.

It is easy to see that if (2) holds, then $\alpha(\tau)$ satisfies the conditions of Theorems 2 and 3 for every $0 < \delta < 2$.

Note also that the s.m. coefficient $\alpha(\tau) \le c_0\rho_0^\tau$, $0 < \rho_0 < 1$, $c_0 > 0$, [2] also satisfies the conditions of Theorems 2 and 3.

By $f(x, y)$ we denote the stationary distribution (X_{t-1}, X_{t-2}, X_t). The function $\Psi(x) = H(a(x)) = \dfrac{a_1(x)}{a_0(x)}$, $a(x) = (a_0(x),\, a_1(x))$, $a_1(x) = \int y f(x, y)\, dy$, $a_0(x) = p(x)$, and the estimate

$$\Psi_{n,2}(x) = \sum_{t=3}^{n+2} \frac{X_t}{h_t^2}\mathbf{K}\left(\frac{x - U_t}{h_t}\right) \Bigg/ \sum_{t=3}^{n+2} \frac{1}{h_t^2}\mathbf{K}\left(\frac{x - U_t}{h_t}\right) = \frac{a_{1n}^{(0j)}(x)}{a_{0n}^{(0j)}(x)} = \frac{a_{1n}(x)}{p_n(x)}, \tag{13}$$

where $U_t = (X_{t-1}, X_{t-2})$.

To find the MSE of the estimate $\Psi_{n,2}(x)$ we use Theorem 2. Let $K(u) \in \mathcal{A}_\nu$, $\mathbf{K}(u) = K(u_1)K(u_2)$, $\sup\limits_{u \in \mathsf{R}^1} |K(u)| < \infty$, $(h_n + 1/(nh_n^2)) \downarrow 0$, and suppose that (10) holds for the sequence (h_n) for $\alpha = \nu$ and $\alpha = -2$.

Suppose that functions $a_i(x)$, $i = 0, 1$, and their derivatives up to and including order ν are continuous and bounded on R^2, and functions $\int y^2 f(x, y)dy$ and $\int y^4 f(x, y)dy$ are bounded on R^2, and, moreover, $\int y^2 f(x, y)dy$ and $\int |y|^{2+\delta} f(x, y)dy$ are continuous at the point x. Then conditions (1)–(5) of Theorem 2 hold; we also suppose that (6) holds. If $p(x) > 0$, then condition (7) holds, too.

If the random variables X_t are uniformly bounded, and we select a nonnegative kernel, then it is easy to show that $\Psi_{n,2}(x)$ are bounded for $\nu = 2$. By condition (8), this is equivalent to the existence of a majorizing sequence with $\gamma = 0$. As a result, for $n \to \infty$ we get:

$$\mathsf{u}^2(\Psi_{n,2}(x)) = \sum_{i,p=0}^{1} H_i H_p \left(S_{-2} \frac{a_{i+p}(x) \displaystyle\int K^2(u)du}{nh_n^2} + S_2^2 \omega_{i2}^{(0)}(x)\omega_{p2}^{(0)}(x) h_n^4 \right) + O\left(\left[\frac{1}{nh_n^2} + h_n^4\right]^{3/2}\right),$$

where

$$H_1 = \frac{1}{p\,(x)}, \quad H_0 = -\frac{\Psi(x)}{p^2\,(x)}, \quad a_2(x) = \int y^2 f(x, y)\, dy;$$

$$\omega_{12}^{(0)}(x) = \frac{T_2}{2}\left(\frac{\partial^2 a_1(x)}{\partial x_1^2} + \frac{\partial^2 a_1(x)}{\partial x_2^2}\right), \quad \omega_{02}^{(0)}(x) = \frac{T_2}{2}\left(\frac{\partial^2 p\,(x)}{\partial x_1^2} + \frac{\partial^2 p(x)}{\partial x_2^2}\right).$$

For $\nu > 2$ we cannot compute the MSE by this method, because the denominator of (3) may become zero, and this does not let us find a majorizing sequence (d_n) in condition (8) of Theorem 2.

In this case, a piecewise smooth approximation solves the problem.

$$\widetilde{\Psi}_{n,2}(x) = \frac{\Psi_{n,2}(x)}{(1+\delta_{n,\nu}|\Psi_{n,2}(x)|^{\tau})^{\rho}}, \tag{14}$$

where $\tau > 0$, $\rho > 0$, $\rho\tau \geqslant 1$, $\delta_{n,\nu} = O\left(h_n^{2\nu} + 1/(nh_n^2)\right)$, $(\delta_{n,\nu}) \downarrow 0$ for $n \to \infty$. Thus, by Theorem 3 for (14) condition (8*) holds if we take an even $\tau \geqslant 4$. Then for $n \to \infty$

$$\mathsf{u}^2\left(\widetilde{\Psi}_{n,2}(x)\right) = \sum_{i,p=0}^{1} H_i H_p \left(S_{-2}\frac{a_{i+p}(x)\int K^2(u)du}{nh_n^2} + S_\nu^2 \omega_{i\nu}^{(0)}(x)\omega_{p\nu}^{(0)}(x)h_n^{2\nu} \right)$$
$$+ O\left(\left[\frac{1}{nh_n^2} + h_n^{2\nu}\right]^{3/2}\right),$$

where

$$\omega_{1\nu}^{(0)}(x) = \frac{T_\nu}{\nu!}\left(\frac{\partial^\nu a_1(x)}{\partial x_1^\nu} + \frac{\partial^\nu a_1(x)}{\partial x_2^\nu}\right),$$
$$\omega_{0\nu}^{(0)}(x) = \frac{T_\nu}{\nu!}\left(\frac{\partial^\nu p(x)}{\partial x_1^\nu} + \frac{\partial^\nu p(x)}{\partial x_2^\nu}\right).$$

The estimate for the volatility function will look like

$$\Phi_{n,2}^2(x) = \frac{\sum\limits_{t=3}^{n+2} \frac{X_t^2}{h_t^2}\mathbf{K}\left(\frac{x-U_t}{h_t}\right)}{\sum\limits_{t=3}^{n+2} \frac{1}{h_t^2}\mathbf{K}\left(\frac{x-U_t}{h_t}\right)} - \frac{\left[\sum\limits_{t=3}^{n+2} \frac{X_t}{h_t^2}\mathbf{K}\left(\frac{x-U_t}{h_t}\right)\right]^2}{\left[\sum\limits_{t=3}^{n+2} \frac{1}{h_t^2}\mathbf{K}\left(\frac{x-U_t}{h_t}\right)\right]^2}.$$

Consider the estimates of sensitivity functions

$$T_{jn}(x) = \left[\frac{\sum\limits_{t=3}^{n+2} \frac{X_t}{h_t^3}\mathbf{K}^{(1j)}\left(\frac{x-U_t}{h_t}\right)}{\sum\limits_{t=3}^{n+2} \frac{1}{h_t^2}\mathbf{K}\left(\frac{x-U_t}{h_t}\right)} - \frac{\sum\limits_{t=3}^{n+2} \frac{X_t}{h_t^2}\mathbf{K}\left(\frac{x-U_t}{h_t}\right)\sum\limits_{t=3}^{n+2} \frac{1}{h_t^3}\mathbf{K}^{(1j)}\left(\frac{x-U_t}{h_t}\right)}{\left[\sum\limits_{t=3}^{n+2} \frac{1}{h_t^2}\mathbf{K}\left(\frac{x-U_t}{h_t}\right)\right]^2}\right],$$

which show how much changes in the input variable x_{t-j} are related to the output variable x_t of the model (1), $j = 1, 2$.

For the sensitivity function, the obstacles we have encountered for finding the majorizing sequence force us to use a piecewise smooth approximation

$$\widetilde{T}_{jn}(x) = \frac{T_{jn}(x)}{(1+\delta_{n,\nu} \mid T_{jn}(x)|^{\tau})^{\rho}}.$$

Here the kernel $K(u)$, in addition to the conditions stated above in this section, should also conform to $\sup\limits_{u\in\mathsf{R}^1}|K^{(1)}(u)| < \infty$ and $\lim\limits_{|u|\to\infty} K(u)$, and for (h_n) it should hold that $\lim\limits_{n\to\infty}\left(h_n + \frac{1}{nh_n^4}\right) = 0$. To use the result of Theorem 3 and find $\mathsf{u}^2\left(\widetilde{T}_{jn}(x)\right)$, we also have to demand that functions $a_0(x)$ and $a_1(x)$ have continuous and bounded on R^2 derivatives for orders up to and including $(\nu+1)$.

Let us apply the estimates above to the data about the prices of SJC "Gazprom" stock for the period from January 15, 2008 till March 24, 2009. The data are shown on Fig. 1, the number observations is $n = 292$. Figure 2 shows the estimate (13) that reflects how "today" price depends on the prices in the previous two days. Figures 3 and 4 show the estimate for the volatility function $\Phi_{n,2}^2(x)$ and level lines of this surface; they allow to estimate the risks of forecasts that

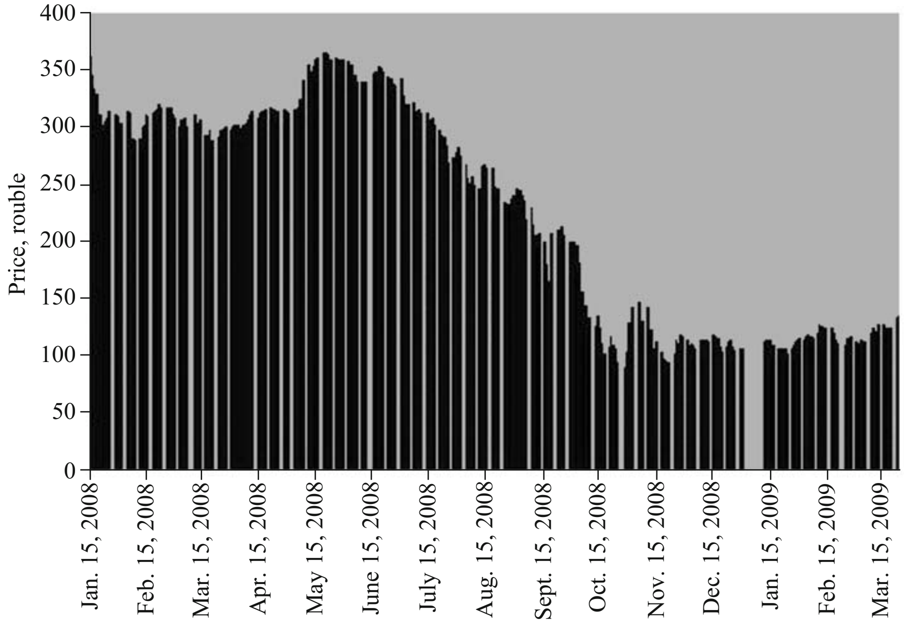

Fig. 1.

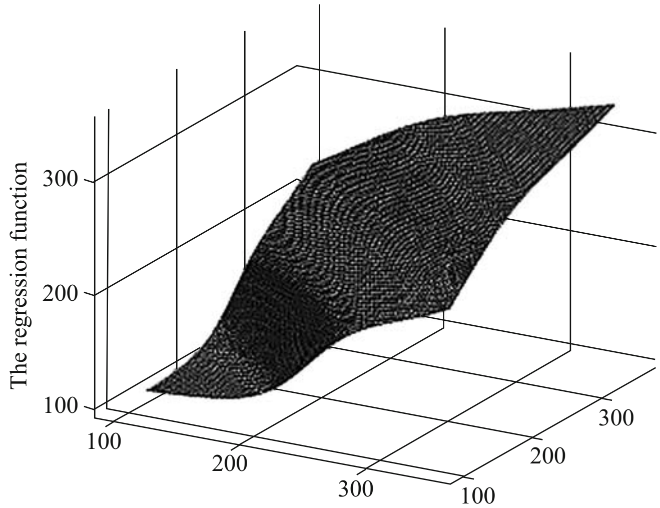

Fig. 2.

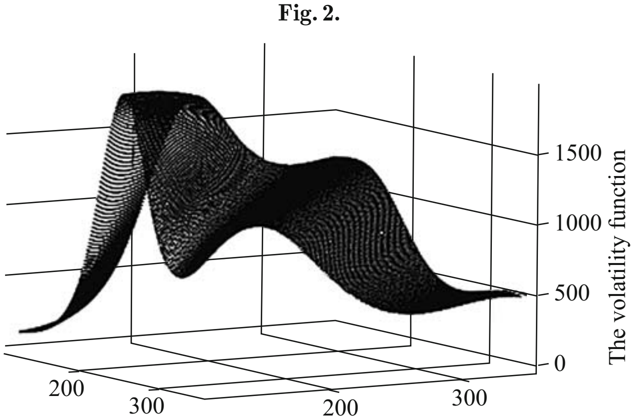

Fig. 3.

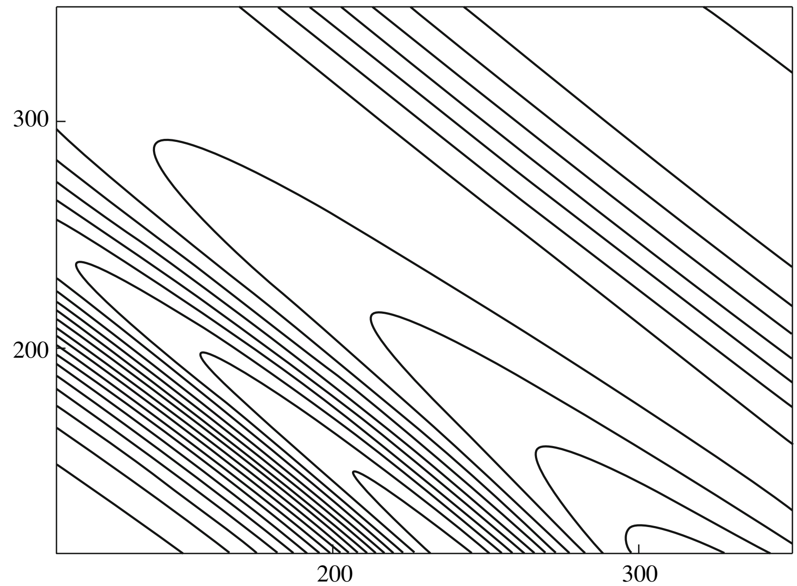

Fig. 4.

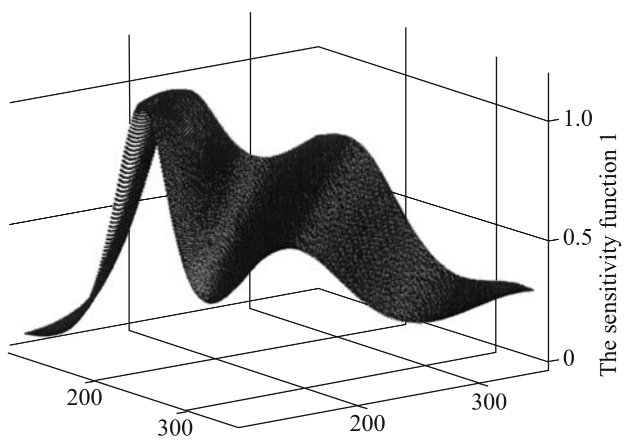

Fig. 5.

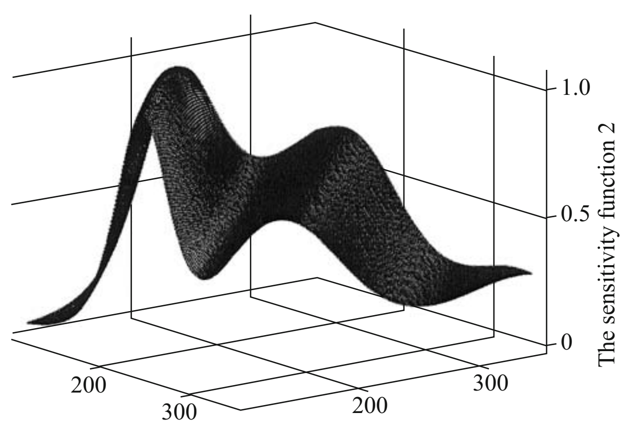

Fig. 6.

use the function $\Psi_{n,2}(x)$. Figures 5 and 6 show the estimates of the sensitivity functions $T_{1n}(x)$ and $T_{2n}(x)$, respectively. It is clear from these figures that the prices of "yesterday" and "the day before yesterday" have approximately the same influence on the "today" prices.

Twenty of the last values were used to estimate the quality of a forecast obtained with $\Psi_{n,2}(x)$. The mean absolute error of this forecast is 3.28 roubles, the relative error—2.7%. We have used the standard Gaussian kernel everywhere and have subjectively adjusted the fuzziness parameters. All surfaces were constructed for the ranges of both variables from 110 to 350; the computations were conducted with the Mathcad 14.0 package.

5. CONCLUSION

This work presents a unified approach to estimating the characteristics (conditional expectation, conditional variance, and sensitivity functions) of the process $(X_t)_{t=\ldots,-1,0,1,2,\ldots}$ of a nonlinear heteroscedastic autoregression of an arbitrary order m. The approach is based on estimating the substitutions of functions depending on functionals of the joint stationary distribution of $X_t, X_{t-1}, \ldots, X_{t-m}$. We have shown that in the considered case we can restrict our attention to functionals of the form

$$a_i(x) \equiv a_i^{(0j)}(x) = \int y^i f(x,y)dy, \quad a_i^{(1j)}(x) = \frac{\partial a_i(x)}{\partial x_j}, \quad i = 0,1,2, \quad j = \overline{1,m}.$$

As substitution elements we propose to use recursive estimates of the kernel type $a_{in}^{(rj)}(x)$. Under the assumption that the autoregression process satisfies the strong mixing condition with exponentially decreasing the strong mixing coefficient, we have considered asymptotic properties of the estimates $a_{in}^{(rj)}(x)$. These estimates allow for finding the main parts of asymptotic MSE of the estimates for the characteristics we are studying due to convergence theorems [11] and their piecewise smooth approximations. The substitution estimates are semirecursive, i.e., we recursively compute only the estimates $a_n^{(rj)}(x)$ appearing in them. By using piecewise smooth approximations of the estimates, we have managed to avoid the problems that occur while we consider MSE due to the existence of the majorant sequence (see condition (8) of Theorem 2).

The proposed algorithms have been illustrated on the example of a second order autoregression and used to forecast the stock prices for SJC "Gazprom."

APPENDIX

By $\bigtriangledown$ we mark the end of a proof.

Informally speaking, Lemmas 1 and 2 are proven very similarly to Lemmas 4 and 5 in [14]. Technical problems arising from the recursive structure of the estimates can be overcome with the help of condition (10) and the monotonicity of (h_n).

Proof of Lemma 1. We denote $\xi_{ti}(x) = \dfrac{1}{h_n^m} X_i^t \mathbf{K}\left(\dfrac{x - U_{i-1}}{h_n}\right)$.

We rewrite the covariance as

$$\operatorname{cov}\left(a_{tn}^{(rj)}(x), a_{pn}^{(qk)}(x)\right) = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{l=1}^{n}\operatorname{cov}\left(\xi_{ti}^{(rj)}(x), \xi_{pl}^{(qk)}(x)\right)$$

$$= \frac{1}{n}\sum_{i=1}^{n}\operatorname{cov}\left(\xi_{ti}^{(rj)}(x), \xi_{pi}^{(qk)}(x)\right) + \frac{2}{n^2}\sum_{\tau=1}^{n-1}\sum_{i=1}^{n-\tau}\operatorname{cov}\left(\xi_{ti}^{(rj)}(x), \xi_{p(\tau+i)}^{(qk)}(x)\right)$$

$$= A_n(x) + R_n(x). \tag{A.1}$$

By Lemma 4 from [13], we have for the $A_n(x)$ summand and for $n \to \infty$:

$$\left| A_n(x) - \frac{S_{-(m+r+q)}}{nh_n^{m+r+q}} a_{t+p}(x) \int K^{(r)}(u) K^{(q)}(u) du \left(\int K^2(u) du \right)^{m-1} \right| = o\left(\frac{1}{nh_n^{m+r+q}} \right).$$

Denoting $U = \xi_{ti}^{(rj)}(x)$ and $V = \xi_{p(\tau+i)}^{(qk)}(x)$, we bound the summand $R_n(x)$. By Proposition 1 from [14], for $r = \dfrac{2+\delta}{\delta}$, $p = q = 2+\delta$, where $\delta > 0$ is an arbitrary number, we get

$$|\operatorname{cov}(U, V)| \leqslant 12[\alpha(\tau)]^{\frac{\delta}{2+\delta}} \left[\mathsf{E}|U|^{2+\delta} \mathsf{E}|V|^{2+\delta} \right]^{\frac{1}{2+\delta}}. \tag{A.2}$$

Since

$$\mathsf{E}|U|^{2+\delta} = \frac{1}{h_i^{(m+r)(2+\delta)}} \int\limits_{\mathsf{R}^{m+1}} \left| z^t \mathbf{K}^{(rj)} \left(\frac{x-t}{h_i} \right) \right|^{2+\delta} f(t, z) dt dz,$$

then, as in Lemma 1 from [13], we get for $2 + \delta = \dfrac{2}{1-\lambda}$:

$$\mathsf{E}|U|^{2+\delta} = \frac{1}{h_i^{(m+r)(2+\delta)-m}} a^{+}_{t(2+\delta)}(x) \int\limits_{\mathsf{R}^m} \left| \mathbf{K}^{(r)}(z) \right|^{2+\delta} dz + o\left(\frac{1}{h_i^{(m+r)(2+\delta)-m}} \right),$$

$$\mathsf{E}|V|^{2+\delta} = \frac{1}{h_{\tau+i}^{(m+q)(2+\delta)-m}} a^{+}_{p(2+\delta)}(x) \int\limits_{\mathsf{R}^m} \left| \mathbf{K}^{(q)}(z) \right|^{2+\delta} dz + o\left(\frac{1}{h_{\tau+i}^{(m+q)(2+\delta)-m}} \right).$$

Taking into account that $\alpha(\tau) \downarrow 0$ and $\lambda = \dfrac{\delta}{2+\delta}$, $0 < \lambda < 1$, we have

$$\sum_{\tau=1}^{\infty} [\alpha(\tau)]^{\frac{\delta}{2+\delta}} \leqslant \int\limits_0^1 [\alpha(\tau)]^{\frac{\delta}{2+\delta}} d\tau + \int\limits_1^2 [\alpha(\tau)]^{\frac{\delta}{2+\delta}} d\tau + \ldots = \int\limits_0^{\infty} [\alpha(\tau)]^{\lambda} d\tau < \infty. \tag{A.3}$$

By (A.2) and (A.3), expressing δ via λ, we get

$$|R_n(x)| = \frac{2}{n^2} \left| \sum_{\tau=1}^{n-1} \sum_{i=1}^{n-\tau} \operatorname{cov}\left(\xi_{ti}^{(rj)}(x), \xi_{p(\tau+i)}^{(qk)}(x) \right) \right|$$

$$\leqslant \frac{24}{n^2} [a^{+}_{t(2+\delta)}(x) a^{+}_{p(2+\delta)}(x)]^{\frac{1}{2+\delta}} \left[\int\limits_{\mathsf{R}^m} \left| \mathbf{K}^{(r)}(z) \right|^{2+\delta} dz \int\limits_{\mathsf{R}^m} \left| \mathbf{K}^{(q)}(z) \right|^{2+\delta} dz \right]^{\frac{1}{2+\delta}}$$

$$\times \sum_{\tau=1}^{n-1} \sum_{i=1}^{n-\tau} \frac{[\alpha(\tau)]^{\frac{\delta}{2+\delta}}}{h_i^{[(m+r)(2+\delta)-m]/(2+\delta)}} \frac{1}{h_{\tau+i}^{[(m+q)(2+\delta)-m]/(2+\delta)}} + o\left(\frac{1}{nh_n^{[(2m+r+q)(2+\delta)-2m]/(2+\delta)}} \right)$$

$$\leqslant \frac{24}{n^2} \left[a^{+}_{\frac{2t}{1-\lambda}}(x) a^{+}_{\frac{2p}{1-\lambda}}(x) \right]^{\frac{1-\lambda}{2}} \left[\int\limits_{\mathsf{R}^m} \left| \mathbf{K}^{(r)}(u) \right|^{\frac{2}{1-\lambda}} du \int\limits_{\mathsf{R}^m} \left| \mathbf{K}^{(q)}(u) \right|^{\frac{2}{1-\lambda}} du \right]^{\frac{1-\lambda}{2}}$$

$$\times \sum_{\tau=1}^{n-1} [\alpha(\tau)]^{\lambda} \sum_{i=\tau}^{n} \frac{1}{h_i^{m(1+\lambda)+r+q}} + o\left(\frac{1}{nh_n^{m(1+\lambda)+r+q}} \right) \leqslant \frac{24}{nh_n^{m(1+\lambda)+r+q}} \left[a^{+}_{\frac{2t}{1-\lambda}}(x) a^{+}_{\frac{2p}{1-\lambda}}(x) \right]^{\frac{1-\lambda}{2}}$$

$$\times \left[\int\limits_{\mathsf{R}^m} \left| \mathbf{K}^{(r)}(u) \right|^{\frac{2}{1-\lambda}} du \int\limits_{\mathsf{R}^m} \left| \mathbf{K}^{(q)}(u) \right|^{\frac{2}{1-\lambda}} du \right]^{\frac{1-\lambda}{2}} \int\limits_0^{\infty} [\alpha(\tau)]^{\lambda} d\tau + o\left(\frac{1}{nh_n^{m(1+\lambda)+r+q}} \right). \tag{A.4}$$

This proves the inequality (11).

Let us now prove (12). (A.1) implies that

$$|R_n(x)| \leqslant \frac{2}{n^2}\sum_{\tau=1}^{c(n)-1}\sum_{i=1}^{n-\tau}\left|\operatorname{cov}\left(\xi_{ti}^{(rj)}(x), \xi_{p\,(\tau+i)}^{(q\,k)}(x)\right)\right|$$
$$+\frac{2}{n^2}\sum_{\tau=c(n)}^{n}\sum_{i=1}^{n-\tau}\left|\operatorname{cov}\left(\xi_{ti}^{(rj)}(x), \xi_{p(\tau+1)}^{(q\,k)}(x)\right)\right| = J_1 + J_2.$$

Suppose that $c(n)$ are positive integers such that $c(n)h_n^m \to 0$, $c(n)h_n^{2m\lambda} \to \infty$ for $n \to \infty$ (for instance, one can take $c(n) \sim h_n^{m(\varepsilon-1)}$, $0 < \varepsilon < 1-2\lambda$, $0 < \lambda < 1/2$). Then for $n \to \infty$

$$J_1 \leqslant \frac{2}{n^2}\sum_{\tau=1}^{c(n)-1}\sum_{i=1}^{n-\tau}\frac{1}{h_i^{m+r}h_{\tau+i}^{m+q}}\int\limits_{\mathsf{R}^{2(m+1)}}\left|z^t\mathbf{K}^{(rj)}\left(\frac{x-u}{h_i}\right)\right.$$
$$\times\left. z^p\mathbf{K}^{(qk)}\left(\frac{x-v}{h_{\tau+i}}\right)\right|\left|f_{i(i+\tau)}(u,z,v,y) - f(u,z)f(v,y)\right| du\,dz\,dv\,dy$$
$$\leqslant \frac{2}{n^2}\sum_{\tau=1}^{c(n)-1}\sum_{i=1}^{n-\tau}\frac{1}{h_i^{m+r}h_{\tau+i}^{m+q}}\left[\sup_x a^+_{i(i+\tau),tp}(x,x) + \sup_x a_t^+(x)\sup_x a_p^+(x)\right]$$
$$\times\int\limits_{\mathsf{R}^{2m}}\left|\mathbf{K}^{(rj)}\left(\frac{x-u}{h_i}\right)\mathbf{K}^{(qk)}\left(\frac{x-v}{h_{\tau+i}}\right)\right| du\,dv$$
$$\leqslant \frac{C}{n^2}\sum_{\tau=1}^{c(n)-1}\sum_{i=1}^{n-\tau}\frac{1}{h_i^r h_{\tau+i}^q}\int\limits_{\mathsf{R}^m}\left|\mathbf{K}^{(r)}(u)\right| du\int\limits_{\mathsf{R}^m}\left|\mathbf{K}^{(q)}(u)\right| du \leqslant \frac{Cc(n)}{n}\frac{1}{h_n^{r+q}}$$
$$= O\left(\frac{h_n^m c(n)}{nh_n^{m+r+q}}\right) = o\left(\frac{1}{nh_n^{m+r+q}}\right).$$

In what follows, C will denote constants which are not necessarily equal even throughout the same proof.

Let us take $\delta = \dfrac{4\lambda}{1-2\lambda}$, $0 < \lambda < 1/2$. Then, similar to (A.4), we get

$$J_2 \leqslant \frac{24}{n^2}\left[a^+_{\frac{2t}{1-\lambda}}(x)a^+_{\frac{2p}{1-\lambda}}(x)\right]^{\frac{1-\lambda}{2}}\left[\int\limits_{\mathsf{R}^m}\left|\mathbf{K}^{(r)}(u)\right|^{\frac{2}{1-\lambda}} du\int\limits_{\mathsf{R}^m}\left|\mathbf{K}^{(q)}(u)\right|^{\frac{2}{1-\lambda}} du\right]^{\frac{1-\lambda}{2}}\sum_{\tau=c(n)}^{n-1}[\alpha(\tau)]^\lambda$$
$$\times\sum_{i=\tau}^{n}\frac{1}{h_i^{m(1+\lambda)+r+q}} + o\left(\frac{1}{nh_n^{m(1+\lambda)+r+q}}\right) \leqslant \frac{24}{nh_n^{m(1+2\lambda)+r+q}}\left[a^+_{\frac{2t}{1-\lambda}}(x)a^+_{\frac{2p}{1-\lambda}}(x)\right]^{\frac{1-2\lambda}{2}}$$
$$\times\left[\int\limits_{\mathsf{R}^m}\left|\mathbf{K}^{(r)}(z)\right|^{\frac{2}{1-2\lambda}} dz\int\limits_{\mathsf{R}^m}\left|\mathbf{K}^{(q)}(z)\right|^{\frac{2}{1-2\lambda}} dz\right]^{\frac{1-2\lambda}{2}}\sum_{\tau=c(n)}^{\infty}[\alpha(\tau)]^{2\lambda} + o\left(\frac{1}{nh_n^{m(1+2\lambda)+r+q}}\right).$$

Further,

$$J_2 \leqslant \frac{C}{c(n)nh_n^{m(1+2\lambda)+r+q}}\sum_{\tau=c(n)}^{\infty}c(n)[\alpha(\tau)]^{2\lambda}$$
$$\leqslant \frac{C}{c(n)nh_n^{m(1+2\lambda)+r+q}}\sum_{\tau=c(n)}^{\infty}\tau[\alpha(\tau)]^{2\lambda} \leqslant \frac{C}{c(n)nh_n^{m(1+2\lambda)+r+q}}\sum_{\tau=1}^{\infty}\tau[\alpha(\tau)]^{2\lambda}.$$

Since $\alpha(\tau)$ is nonincreasing, i.e., $1 \geqslant \alpha(1) \geqslant \alpha(2) \geqslant \dots$, then

$$\sum_{\tau=1}^{\infty}[\alpha(\tau)]^{\lambda} < \infty, \quad 0 < \lambda < \frac{1}{2},$$

implies that

$$\sum_{\tau=1}^{\infty}\tau[\alpha(\tau)]^{2\lambda} = \sum_{\gamma=1}^{\infty}\sum_{t=\gamma}^{\infty}[\alpha(t)]^{2\lambda} \leqslant \sum_{\gamma=1}^{\infty}[\alpha(\gamma)]^{\lambda}\sum_{t=\gamma}^{\infty}[\alpha(t)]^{\lambda} \leqslant \left[\sum_{\tau=1}^{\infty}[\alpha(\tau)]^{\lambda}\right]^2 < \infty.$$

Thus,

$$J_2 = O\left(\frac{1}{c(n)nh_n^{m(1+2\lambda)+r+q}}\right) = O\left(\frac{1}{nh_n^{m+r+q}c(n)h_n^{2m\lambda}}\right) = o\left(\frac{1}{nh_n^{m+r+q}}\right). \quad \triangledown$$

Proof of Lemma 2. We apply the techniques used to prove Lemma 4 in [15, pp. 239] and Lemma 1 in [15, pp. 270]. We denote

$$\eta_{ir}^{(rj)} = \frac{1}{h_i^{m+r}}\left[X_i^t\mathbf{K}^{(rj)}\left(\frac{x-U_{i-1}}{h_i}\right) - \mathsf{E}\left[X_i^t\mathbf{K}^{(rj)}\left(\frac{x-U_{i-1}}{h_i}\right)\right]\right].$$

The sequence $(Z_j)_{j\geqslant 1}$ is stationary, so

$$\mathsf{E}\left(S_{tn}^{(rj)}\right)^4 = \frac{1}{n^4}\,\mathsf{E}\left(\sum_{i=1}\eta_{ir}\right)^4 \leqslant \frac{4!}{n^4}\sum_{s,i,j,k}|\mathsf{E}\eta_{sr}\eta_{(i+s)r}\eta_{(i+j+s)r}\eta_{(i+j+k+s)r}|, \tag{A.5}$$

where the sum is taken over $s, i, j, k \geqslant 1$, $s+i+j+k \leqslant n$. Applying Proposition 1 from [14] again for $r = \dfrac{2+\delta}{\delta}$, $p = q = 2+\delta$, and taking into consideration condition (2) of this lemma and the fact that $\mathsf{E}\eta_{sr} = 0$, after changing variables in the integrals we get

$$\begin{aligned}&\left|\mathsf{E}\{\eta_{sr}(\eta_{(i+s)r}\eta_{(i+j+s)r}\eta_{(i+j+k+s)r})\}\right| \\ &\leqslant 12[\alpha(i)]^{\frac{\delta}{2+\delta}}\left[\mathsf{E}\left|\eta_{sr}\right|^{2+\delta}\mathsf{E}\left|\eta_{(i+s)r}\eta_{(i+j+s)r}\eta_{(i+j+k+s)r}\right|^{2+\delta}\right]^{\frac{1}{2+\delta}} \\ &\leqslant C[\alpha(i)]^{\frac{\delta}{2+\delta}}h_n^{\frac{4m}{2+\delta}-4(m+r)}\left(\int\limits_{\mathsf{R}^m}\left|\mathbf{K}^{(r)}(u)\right|^{2+\delta}du\right)^{\frac{4}{2+\delta}} \\ &\times\sup_x\left[a^+_{t(2+\delta)}(x)a^+_{s(s+j)(s+j+k),t(2+\delta)}(x,x,x)\right]^{\frac{1}{2+\delta}} \leqslant Ch_n^{\frac{4m}{2+\delta}-4(m+r)}[\alpha(i)]^{\frac{\delta}{2+\delta}}\end{aligned}$$

and

$$\left|\mathsf{E}\{(\eta_{sr}\eta_{(i+s)r}\eta_{(i+j+s)r})\eta_{(i+j+k+s)r}\}\right| \leqslant Ch_n^{\frac{4m}{2+\delta}-4(m+r)}[\alpha(k)]^{\frac{\delta}{2+\delta}}.$$

Similarly, we find

$$\begin{aligned}&\left|\mathsf{E}\{(\eta_{sr}\eta_{(i+s)r})(\eta_{(i+j+s)r}\eta_{(i+j+k+s)r})\}\right| = |\mathsf{E}\{(\eta_{sr}\eta_{(i+s)r})(\eta_{(i+j+s)r}\eta_{(i+j+k+s)r})\} \\ &- \mathsf{E}(\eta_{sr}\eta_{(i+s)r})\mathsf{E}(\eta_{(i+j+s)r}\eta_{(i+j+k+s)r}) + \mathsf{E}(\eta_{sr}\eta_{(s+i)r})\mathsf{E}(\eta_{(i+j+s)r}\eta_{(i+j+k+s)r})| \\ &\leqslant Ch_n^{\frac{4m}{2+\delta}-4(m+r)}[\alpha(j)]^{\frac{\delta}{2+\delta}} + \left|\mathsf{E}\,(\eta_{sr}\eta_{(s+i)r})\mathsf{E}(\eta_{(i+j+s)r}\eta_{(i+j+k+s)r})\right|,\end{aligned} \tag{A.6}$$

$$\left|\mathsf{E}\,(\eta_{sr}\eta_{(i+s)r})\right| \leqslant Ch_n^{\frac{2m}{2+\delta}-2(m+r)}[\alpha(i)]^{\frac{\delta}{2+\delta}}, \tag{A.7}$$

$$|\mathsf{E}\,(\eta_{(i+j+s)r}\eta_{(i+j+k+s)r})| \leqslant Ch_n^{\frac{2m}{2+\delta}-2(m+r)}[\alpha(k)]^{\frac{\delta}{2+\delta}}. \tag{A.8}$$

Substituting (A.7) and (A.8) into (A.6), we get:

$$
\begin{aligned}
&|\mathsf{E}\{(\eta_{sr}\eta_{(i+s)r})(\eta_{(i+j+s)r}\eta_{(i+j+k+s)r})\}| \\
&\leqslant Ch_n^{\frac{4m}{2+\delta}-4(m+r)}\left([\alpha(i)]^{\frac{\delta}{2+\delta}}[\alpha(k)]^{\frac{\delta}{2+\delta}}+[\alpha(j)]^{\frac{\delta}{2+\delta}}\right).
\end{aligned}
$$

So,

$$
\begin{aligned}
&C^{-1}h_n^{\frac{-4m}{2+\delta}+4(m+r)}\sum_{s,i,j,k}|\mathsf{E}\eta_{sr}\eta_{(i+s)r}\eta_{(i+j+s)r}\eta_{(i+j+k+s)r}| \\
&\leqslant 2n\sum_{i=1}^{n}\sum_{j,k=1}^{i}[\alpha(i)]^{\frac{\delta}{2+\delta}}+n\sum_{j=1}^{n}\sum_{i,k=1}^{\infty}[\alpha(i)]^{\frac{\delta}{2+\delta}}[\alpha(k)]^{\frac{\delta}{2+\delta}} \\
&\leqslant n\sum_{i=1}^{n}i^2[\alpha(i)]^{\frac{\delta}{2+\delta}}+n^2\sum_{i=1}^{\infty}[\alpha(i)]^{\frac{\delta}{2+\delta}}\sum_{k=1}^{\infty}[\alpha(k)]^{\frac{\delta}{2+\delta}} \\
&\leqslant n\sum_{i=1}^{\infty}i^2[\alpha(i)]^{\frac{\delta}{2+\delta}}+n^2\left(\sum_{i=1}^{\infty}[\alpha(i)]^{\frac{\delta}{2+\delta}}\right)^2 \\
&\leqslant n\int_1^{\infty}\tau^2[\alpha(\tau)]^{\frac{\delta}{2+\delta}}d\tau+n^2\left(\int_1^{\infty}[\alpha(\tau)]^{\frac{\delta}{2+\delta}}d\tau\right)^2,
\end{aligned}
\tag{A.9}
$$

which also uses the fact that $1=\alpha(0)\geqslant\alpha(1)\geqslant\alpha(2)\geqslant\ldots.$

Now (A.5) and (A.9) imply that

$$
\begin{aligned}
\mathsf{E}\left(S_{tn}^{(rj)}\right)^4 &\leqslant \frac{4!nCh_n^{\frac{4m}{2+\delta}}}{n^4h_n^{4(m+r)}}\left[n\left(\sum_{i=1}^{\infty}[\alpha(i)]^{\frac{\delta}{2+\delta}}\right)^2+3\sum_{k=1}^{\infty}k^2[\alpha(k)]^{\frac{\delta}{2+\delta}}\right] \\
&\leqslant \frac{Ch_n^{-\frac{2m\delta}{2+\delta}}}{n^2h_n^{2(m+2r)}}\left(\sum_{i=1}^{\infty}[\alpha(i)]^{\frac{\delta}{2+\delta}}\right)^2+\frac{Ch_n^{\frac{m(2-\delta)}{2+\delta}}}{n^2h_n^{2(m+2r)}nh_n^m}\sum_{k=1}^{\infty}k^2[\alpha(k)]^{\frac{\delta}{2+\delta}}.
\end{aligned}
$$

Consider the first summand in the right-hand side of (A.9). Take a sequence of positive numbers $c(n)$ such that $c(n)=o(n)$, $c(n)=O(h_n^{-m})$. In this case, since $c^2(n)h_n^{\frac{2m\delta}{2+\delta}}\to\infty$ for $n\to\infty$, we get

$$
\begin{aligned}
\frac{h_n^{-\frac{2m\delta}{2+\delta}}}{n^2h_n^{2(m+2r)}}\left(\sum_{i=c(n)}^{\infty}[\alpha(i)]^{\frac{\delta}{2+\delta}}\right)^2 &= \frac{h_n^{-\frac{2m\delta}{2+\delta}}}{n^2h_n^{2(m+2r)}c^2(n)}\left(\sum_{i=c(n)}^{\infty}c(n)[\alpha(i)]^{\frac{\delta}{2+\delta}}\right)^2 \\
&\leqslant \frac{C}{n^2h_n^{2(m+2r)}c^2(n)h_n^{\frac{2m\delta}{2+\delta}}}\left(\sum_{\tau=1}^{\infty}\tau[\alpha(\tau)]^{\frac{\delta}{2+\delta}}\right)^2=o\left(\frac{1}{n^2h_n^{2(m+2r)}}\right).
\end{aligned}
$$

We now show that

$$
\frac{Ch_n^{-\frac{2m\delta}{2+\delta}}}{n^2h_n^{2(m+2r)}}\left(\sum_{i=1}^{c(n)}[\alpha(i)]^{\frac{\delta}{2+\delta}}\right)^2\leqslant O\left(\frac{1}{n^2h_n^{2(m+2r)}}\right).
$$

By conditions (2) and (5) of Lemma 2, we conclude that

$$\left|\mathsf{E}\{\eta_{sr}(\eta_{(i+s)r}\eta_{(i+j+s)r}\eta_{(i+j+k+s)r})\}\right|$$
$$\leqslant Ch_n^{-4r}\left(\int\left|\mathbf{K}^{(r)}(u)\right|du\right)^4 \sup_x a^+_{s(i+s)(i+j+s)(i+j+k+s),t}(x,x,x,x) \leqslant Ch_n^{-4r}\overline{a}_{si(\cdot\cdot),t},$$
$$\left|\mathsf{E}\{(\eta_{sr}\eta_{(i+s)r}\eta_{(i+j+s)r})\eta_{(i+j+k+s)r}\}\right| \leqslant Ch_n^{-4r}\overline{a}_{s(\cdot\cdot)k,t},$$
$$\left|\mathsf{E}\{(\eta_{sr}\eta_{(i+s)r})(\eta_{(i+j+s)r}\eta_{(i+j+k+s)r})\}\right| \leqslant Ch_n^{-4r}\overline{a}_{s(\cdot)j(\cdot),t},$$

where

$$\overline{a}_{si(\cdot\cdot),t} = \max_{j,k}\sup_x a^+_{s(i+s)(i+j+s)(i+j+k+s),\,t}(x,x,x,x),$$
$$\overline{a}_{s(\cdot\cdot)k,t} = \max_{i,\,j}\sup_x a^+_{s(i+s)(i+j+s)(i+j+k+s),\,t}(x,x,x,x),$$
$$\overline{a}_{s(\cdot)j(\cdot),t} = \max_{i,\,k}\sup_x a^+_{s(i+s)(i+j+s)(i+j+k+s),\,t}(x,x,x,x).$$

Similar to inequalities (A.9),

$$\sum_{s,i,j,k}\left|\mathsf{E}\eta_{sr}\eta_{(i+s)r}\eta_{(i+j+s)r}\eta_{(i+j+k+s)r}\right| \leqslant 3Ch_n^{-4r}\sum_{s=1}^{n}\sum_{i,j,k=1}^{s}\overline{a}_{si,t}$$
$$\leqslant 3Ch_n^{-4r}\sum_{s=1}^{n}s^3\overline{a}_{s,t} \leqslant 3Ch_n^{-4r}n\sum_{k=1}^{\infty}k^2\overline{a}_{k,t}$$
$$= 3Ch_n^{-4r}n\left(\sum_{k=1}^{c(n)-1}k^2\overline{a}_{k,t} + \sum_{k=c(n)}^{\infty}k^2\overline{a}_{k,t}\right), \qquad \text{(A.10)}$$

where $\overline{a}_{si,t} = \max\left(\overline{a}_{si(\cdot\cdot),t}, \overline{a}_{s(\cdot\cdot)i,t}, \overline{a}_{s(\cdot)i(\cdot),t}\right)$, $\overline{a}_{s,t} = \max\limits_{i=1,s}\overline{a}_{si,t}$.

The second summand in the right-hand side of (A.10) has order $o\left(\dfrac{1}{n^2h_n^{2(m+2r)}}\right)$, and the first summand satisfies $\sum\limits_{k=1}^{c(n)-1}k^2\overline{a}_{k,t} \leqslant Cnc^2(n)$. Since $\lim\limits_{n\to\infty}c(n)h_n = o(1)$, we get

$$\mathsf{E}\left(S_{nr}^{(rj)}\right)^4 \leqslant \frac{n^2Cc^2(n)h_n^4}{n^4h_n^{4(m+r)}} + o\left(\frac{1}{n^2h_n^{2(m+2r)}}\right) = O\left(\frac{1}{n^2h_n^{2(m+2r)}}\right). \quad \triangledown$$

Proof of Theorem 2 follows from (9), Lemmas 1 and 3, Theorem 1 from [11] for $k = 2$ and $m = 4$, in which we set $t_n = A_n$, $d_n = O\left(nh_n^{m+2\max(r)} + h_n^{-2\nu}\right)$. Let us make two notes, though. A condition of Lemma 2:

$$\int_0^\infty \tau^2[\alpha(\tau)]^{\frac{\delta}{2+\delta}}d\tau < \infty, \quad 0 < \delta < 2,$$

implies the following condition of Lemma 1:

$$\int_0^\infty [\alpha(\tau)]^\lambda d\tau < \infty, \quad 0 < \lambda < 1/2.$$

Besides, $\sup\limits_x a^+_{t\beta}(x) < \infty$ and $\beta = 0, 4$ imply that $\sup\limits_x a^+_{t(2+\delta)}(x) < \infty$. $\triangledown$

REFERENCES

1. Masry, E. and Tjøstheim, D., Nonparametric Estimation and Identification of Nonlinear ARCH Time Series, *Econom. Theory*, 1995, vol. 11, pp. 258–289.
2. Härdle, W., Tsybakov, A., and Yang, L., Nonparametric Vector Autoregression, *J. Statist. Planning Inference*, 1998, vol. 68, pp. 221–245.
3. Jang, L., Härdle, W., and Nielsen, J.P., Nonparametric Autoregression with Multiplicative Volatility and Additive Mean, *J. Time Ser. Anal.*, 1999, vol. 20, no. 5, pp. 579–602.
4. Raibman, N.S., *Chto takoe identifikatsiya* (What is Identification), Moscow, Nauka, 1970.
5. Eykhof, P., *System Identification and State Estimation*, London: Wiley, 1974. Translated ubder the title *Osnovy identifikatsii sistem upravleniya. Otsenivanie parametrov i sostoyaniya*, Moscow, Mir, 1975.
6. Nadaraya, E.A., On Regression Estimates, *Theor. Prob. App.*, 1964, vol. 19, no. 1, pp. 147–149.
7. Watson, G.S., Smooth Regression Analysis, *Sankhya. Indian J. Statist.*, 1964, vol. A26, pp. 359–372.
8. Tjøstheim, D. and Auestad, B.H., Nonparametric Identification of Nonlinear Time Series: Projections, *J. Am. Statist. Associat.*, 1994, vol. 89, no. 428. pp. 1398–1409.
9. Walk, H., Strong Universal Pointwise Consistency of Recursive Kernel Regression Estimates, *Ann. Inst. Statist. Math.*, 2001, vol. 53, no. 4, pp. 691–707.
10. Pashchenko, F.F., Sensitivity Function and Its Applications to Choosing the Optimal Model, in *Sistemy upravleniya* (Control Systems), Moscow: Nauka, 1973, pp. 72–78.
11. Koshkin, G.M., Deviation Moments of the Substitution Error and Its Piecewise Smooth Approximations, *Sib. Mat. Zh.*, 1999, vol. 40, no. 3, pp. 605–618.
12. Bosq, D. and Cheze-Payaud, N., Optimal Asymptotic Quadratic Error of Nonparametric Regression Function Estimates for a Continuous-Time Process from Sampled-Data, *Statistics*, 1999, vol. 32, pp. 229–247.
13. Kitaeva, A.V. and Koshkin, G.M., Recurrent Nonparametric Estimation of Functions from Functionals of Multidimensional Density and Their Derivatives, *Autom. Remote Control*, 2009, no. 3, pp. 389–407.
14. Kitaeva, A.V. and Koshkin, G.M., *Neparametricheskoe otsenivanie funktsii ot uslovnykh momentov po mnogomernym nablyudeniyam s sil'nym peremeshivaniem* (Nonparametric Estimation of Functions of Conditional Moments by Multidimensional Observations with Strong Mixing), *Proc. 8th Int. Conf. "System Identification and Control Problems,"* January 26–30, 2009, Moscow: Inst. Probl. Upravlen., 2009, pp. 1001–1018.
15. Billingsley, P., *Convergence of Probability Measures*, New York: Wiley, 1968. Translated under the title *Skhodimost' veroyatnostnykh mer*, Moscow: Nauka, 1977.

This paper was recommended for publication by A.I. Kibzun, a member of the Editorial Board