Ann. Inst. Statist. Math.
35 (1983), Part A, 215-228

AN ORTHOGONAL SERIES ESTIMATE OF TIME-VARYING REGRESSION

WŁODZIMIERZ GREBLICKI, DANUTA RUTKOWSKA AND LESZEK RUTKOWSKI

(Received July 7, 1980; revised Mar. 19, 1982)

Summary

Let $(X_1, Y_1), (X_2, Y_2), \cdots$ be independent pairs of random variables according to the model $Y_n = t_n(X_n)R(X_n) + Z_n$, $n=1, 2, \cdots$, where t_n and R are unknown functions. Z_n's are i.i.d. random variables with zero mean and finite variance. The marginal density of X_n is independent of n. In the paper nonparametric estimates of a nonstationary regression function $\mathrm{E}\{Y_n | X_n = x\} = t_n(x)R(x)$ are proposed and their asymptotic properties are investigated.

1. Introduction

Let $(X_1, Y_1), (X_2, Y_2), \cdots$ be a sequence of independent pairs of random variables according to the model

$$Y_n = R_n(X_n) + Z_n , \qquad n=1, 2, \cdots , \tag{1}$$

where R_n's are Borel-measurable functions and Z_n's are i.i.d. random variables. Z_n is independent of X_n and

$$\mathrm{E}\, Z_n = 0 , \qquad \mathrm{E}\, Z_n^2 < \infty . \tag{2}$$

Y_n takes values in R, while X_n in $\mathcal{X}$, where $\mathcal{X}$ is a Borel subset of R^p. The marginal Lebesgue density f of X_n is independent of n. Our aim is to estimate the nonstationary regression function i.e. to track $R_n(x) = \mathrm{E}\{Y_n | X_n = x\}$.

In stationary case several nonparametric methods have been proposed. We mention works of Nadaraya [13], Rosenblatt [15], Noda [14], Collomb [4], Greblicki and Krzyżak [10] as well as Devroye and Wagner [6] based on the Rosenblatt-Parzen density estimate. The nearest neighbor estimate is represented by Stone [18] and Devroye [5]. The orthog-

Key words and phrases: Nonparametric estimation, regression function, nonstationary regression function, orthogonal series, stochastic approximation.

onal series approach has been discussed by Mirzahmedov and Hašimov [12] and Greblicki [9].

The next section consists of assumptions and preliminaries. The main results of the paper i.e. Theorems 1 and 2 are given in Sections 3 and 4. Concluding Theorems 3 and 4 are in Section 5. In the closing section an example is considered in which restrictions made in this paper are satisfied even if the regression function converges to infinity as n tends to infinity.

2. Preliminaries and assumptions

Throughout this paper we assume that

$$R_n(x)=t_n(x)R(x)\ . \tag{3}$$

A sequence of functions $\{t_n\}$ is unknown. It, however, becomes similar to some sequence of numbers; more precisely, there exists a sequence $\{c_n\}$ such that

$$\sup_x |t_n(x)-c_n| \xrightarrow{n} 0\ . \tag{4}$$

Further assumptions imposed on $\{t_n\}$ will be given in the sequel. The functional form of R is completely unknown.

All integrals and supremums are taken over $\mathcal{X}$. Besides, $K_1, K_2, \cdots$ denote positive constants numbered in order of appearance.

Furthermore we assume that

$$\int R^2(x)f(x)dx<\infty\ , \tag{5}$$

which implies

$$\int |R(x)|f(x)dx<\infty\ . \tag{6}$$

Note that from (1)–(5) it follows that

$$\mathrm{E}\,Y_n^2 \leqq K_1+K_2c_n^2\ . \tag{7}$$

In the next parts of the paper we refer (7) rather than (1)–(5).

We also introduce the following notations:

$$h_n(x)=R_n(x)f(x)\ , \qquad h(x)=R(x)f(x)\ .$$

We shall use a complete orthonormal system $\{g_j\}$, $j=0, 1, \cdots$, defined on $\mathcal{X}$, such that

$$|g_j(x)| \leqq G_j \tag{8}$$

for all $x \in \mathcal{X}$, where $\{G_j\}$ is a sequence of numbers.

From (4), (6) and (8) it follows that functions h_n can be expanded in the orthogonal series

$$h_n(x) \sim \sum_{j=0}^{\infty} a_{jn} g_j(x) . \tag{9}$$

It means that

$$a_{jn} = \int R_n(x) g_j(x) f(x) dx = \mathrm{E}\,\{Y_n g_j(X_n)\} . \tag{10}$$

In Section 5 unknown coefficients a_{jn}'s are estimated by the Robbins-Monro stochastic approximation method, see e.g. Wasan [21], i.e.

$$\hat{a}_{j,n+1} = \hat{a}_{jn} - \gamma_n(\hat{a}_{jn} - Y_{n+1} g_j(X_{n+1})) , \tag{11}$$

where $\hat{a}_{j0}=0$ for all j, and $\{\gamma_n\}$ is a sequence of positive numbers.

Let us expand f in the orthogonal series

$$f(x) \sim \sum_{j=0}^{\infty} b_j g_j(x) , \tag{12}$$

where

$$b_j = \int g_j(x) f(x) dx = \mathrm{E}\, g_j(X_1) . \tag{13}$$

Clearly

$$\hat{b}_{jn} = n^{-1} \sum_{i=1}^{n} g_j(X_i) \tag{14}$$

is an unbiased estimator of b_j.

As an estimator of R_n we take the statistics

$$\hat{R}_n(x) = \hat{h}_n(x) / \hat{f}_n(x) , \tag{15}$$

where

$$\hat{h}_n(x) = \sum_{j=0}^{N(n)} \hat{a}_{jn} g_j(x) , \tag{16}$$

$$\hat{f}_n(x) = \sum_{j=0}^{M(n)} \hat{b}_{jn} g_j(x) , \tag{17}$$

and where $\{N(n)\}$ and $\{M(n)\}$ are sequences of integers.

It should be mentioned that estimator (17) of a density function was proposed by Čencov [2] and studied by Schwartz [17], Kronmal and Tarter [11] and Bosq [1] among others. For $\gamma_n = 1/(n+1)$, $\hat{a}_{jn}$ is equal to $n^{-1} \sum_{i=1}^{n} Y_i g_j(X_i)$ and estimate (15) becomes that of studied by

Greblicki [9] for the stationary case.

In the paper we investigate asymptotic properties of (15), i.e. we show that, under suitable conditions,

$$|\hat{R}_n(x) - R_n(x)| \xrightarrow{n} 0$$

in probability and with probability one.

In order to prove convergence theorems we expand h in the orthogonal series

$$h(x) \sim \sum_{j=0}^{\infty} a_j g_j(x) , \tag{18}$$

where

$$a_j = \int R(x) g_j(x) f(x) dx . \tag{19}$$

Finally, we define

$$d_n = \sup_{N_1, N_2} \left\{ \sum_{j=N_1}^{N_2} \mathrm{E}\,(\hat{a}_{jn} - a_{jn})^2 \Big/ \sum_{j=N_1}^{N_2} G_j^2 \right\} , \tag{20}$$

where N_1 and N_2 run over the set of all integers. In Section 5 it will be shown that $\{d_n\}$ is bounded by a power sequence convergent to zero.

Herein we use the following two lemmas:

LEMMA A (Chung [3]). *Let $p_1, p_2, \cdots$ be real numbers such that for $n \geqq n_0$*

$$p_{n+1} \leqq (1 - c/n^{\omega}) p_n + c'/n^t ,$$

where $0 < \omega < 1$, $c > 0$, $c' > 0$, t real. Then

$$\limsup_{n \to \infty} n^{t-\omega} p_n \leqq c'/c .$$

LEMMA B (Van Ryzin [20]). *Let $\{A_n\}$ and $\{B_n\}$ be two sequences of random variables on a probability space (Ω, F, P). Let $\{F_n\}$ be a sequence of Borel fields such that $F_n \subset F_{n+1} \subset F$, and let A_n and B_n be measurable with respect to F_n. If $A_n \geqq 0$ a.e., $\mathrm{E}\, A_1$ is finite, and*

$$\mathrm{E}\,\{A_{n+1} | F_n\} \leqq A_n + B_n \qquad a.e.,$$

$$\sum_{n=1}^{\infty} \mathrm{E}\,|B_n| < \infty ,$$

then $\{A_n\}$ converges almost surely to a finite limit as n tends to infinity.

3. Convergence in probability

First we state and prove two lemmas.

LEMMA 1. *If* (6) *is satisfied and*

$$d_n^{1/2} \sum_{j=0}^{N(n)} G_j^2 \xrightarrow{n} 0 , \tag{21}$$

$$\sup_x |t_n(x) - c_n| \sum_{j=0}^{N(n)} G_j^2 \xrightarrow{n} 0 , \tag{22}$$

then

$$\mathrm{E}\,(\hat{h}_n(x) - h_n(x))^2 \xrightarrow{n} 0 \tag{23}$$

at every point $x \in \mathcal{X}$ *at which*

$$c_n\left[\sum_{j=0}^{N(n)} a_j g_j(x) - h(x)\right] \xrightarrow{n} 0 . \tag{24}$$

PROOF. Observe

$$\begin{aligned} \hat{h}_n(x) - h_n(x) = &\sum_{j=0}^{N(n)} (\hat{a}_{jn} - a_{jn}) g_j(x) + \sum_{j=0}^{N(n)} (a_{jn} - c_n a_j) g_j(x) \\ &+ c_n\left[\sum_{j=0}^{N(n)} a_j g_j(x) - h(x)\right] + (c_n h(x) - h_n(x)) . \end{aligned} \tag{25}$$

By Cauchy's inequality, the expectation of the squared first term on the right-hand side in (25) is not greater than

$$\sum_{j=0}^{N(n)} \mathrm{E}\,(\hat{a}_{jn} - a_{jn})^2 \sum_{j=0}^{N(n)} G_j^2 \leqq d_n \left[\sum_{j=0}^{N(n)} G_j^2\right]^2 . \tag{26}$$

Since

$$\begin{aligned} |a_{jn} - c_n a_j| &= \left| \int (t_n(x) - c_n) R(x) g_j(x) f(x) dx \right| \\ &\leqq G_j \sup_x |t_n(x) - c_n| \int |R(x)| f(x) dx , \end{aligned}$$

the absolute value of the second term in (25) does not exceed

$$(\sup_x |t_n(x) - c_n|) \sum_{j=0}^{N(n)} G_j^2 \int |R(x)| f(x) dx . \tag{27}$$

Moreover, the absolute value of the fourth term in (25) is majorized by

$$|R(x)| f(x) \sup_x |t_n(x) - c_n| . \tag{28}$$

In view of (25), (26), (27) and (28) the proof is complete.

LEMMA 2. *If*

$$c_n n^{-1/2} \sum_{j=0}^{M(n)} G_j^2 \xrightarrow{n} 0 ,$$

then

$$c_n^2 \mathrm{E}\,(\hat{f}_n(x)-f(x))^2 \xrightarrow{n} 0$$

at every point $x \in \mathcal{X}$, *at which*

$$c_n\left[\sum_{j=0}^{M(n)} b_j g_j(x)-f(x)\right] \xrightarrow{n} 0\ .$$

PROOF. Obviously

$$\hat{f}_n(x)-f(x)=\sum_{j=0}^{M(n)} (\hat{b}_{jn}-b_j)g_j(x)+\left[\sum_{j=0}^{M(n)} b_j g_j(x)-f(x)\right]\ . \tag{29}$$

Since $\mathrm{E}\,(\hat{b}_{jn}-b_j)^2 \leqq G_j^2/n$, the expectation of the squared first term in (29) is not greater than

$$\sum_{j=0}^{M(n)} \mathrm{E}\,(\hat{b}_{jn}-b_j)^2 \sum_{j=0}^{M(n)} G_j^2 \leqq n^{-1}\left[\sum_{j=0}^{M(n)} G_j^2\right]^2 ,$$

which completes the proof.

Combining Lemmas 1 and 2, we get the main result of this section.

THEOREM 1. *Let* (6), (21) *and* (22) *be satisfied. Let, moreover,*

$$(|c_n|+1)n^{-1/2} \sum_{j=0}^{M(n)} G_j^2 \xrightarrow{n} 0\ . \tag{30}$$

Then

$$|\hat{R}_n(x)-R_n(x)| \xrightarrow{n} 0$$

in probability at every point $x \in \mathcal{X}$ *at which* $f(x)>0$, (24) *holds and*

$$(|c_n|+1)\left[\sum_{j=0}^{M(n)} b_j g_j(x)-f(x)\right] \xrightarrow{n} 0\ . \tag{31}$$

PROOF. The result follows from the equality

$$\begin{aligned}\hat{R}_n(x)-R_n(x)&=(\hat{h}_n(x)-h_n(x))/\hat{f}_n(x)\\&\quad+h_n(x)(f(x)-\hat{f}_n(x))/f(x)\hat{f}_n(x)\ .\end{aligned} \tag{32}$$

4. Almost sure convergence

As in the previous section we start with two lemmas.

LEMMA 3. *Assume that conditions of Lemma* 1 *are satisfied. Furthermore, let*

$$\sum_{n=1}^{\infty} \gamma_n^2(1+c_n^2)\left[\sum_{j=0}^{N(n)} G_j^2\right]^2 < \infty\ , \tag{33}$$

and let

(34) $$\sum_{n=1}^{\infty} \gamma_n d_n \left[\sum_{j=0}^{N(n)} G_j^2\right]^2 < \infty ,$$

(35) $$\sum_{n=1}^{\infty} d_n \sum_{j=0}^{N(n)} G_j^2 \sum_{k=N(n)}^{N(n+1)} G_k^2 < \infty .$$

Then

(36) $$|\hat{h}_n(x) - h_n(x)| \xrightarrow{n} 0$$

with probability one, at every point $x \in \mathcal{X}$ *at which* (24) *holds.*

PROOF. By (25), (27) and (28) it suffices to show that

(37) $$\sum_{j=0}^{N(n)} (\hat{a}_{jn} - a_{jn}) g_j(x) = \sum_{j=0}^{N(n)} (\hat{a}_{jn} - \mathrm{E}\,\hat{a}_{jn}) g_j(x) + \sum_{j=0}^{N(n)} (\mathrm{E}\,\hat{a}_{jn} - a_{jn}) g_j(x)$$

converges to zero with probability one as n tends to infinity. Now we are concerned with the second term in (37). The absolute value of the term does not exceed

(38) $$\left[\sum_{j=0}^{N(n)} (\mathrm{E}\,\hat{a}_{jn} - a_{jn})^2 \sum_{k=0}^{N(n)} G_k^2\right]^{1/2} \leqq \left[\sum_{j=0}^{N(n)} \mathrm{E}\,(\hat{a}_{jn} - a_{jn})^2 \sum_{k=0}^{N(n)} G_k^2\right]^{1/2} \leqq d_n^{1/2} \sum_{j=0}^{N(n)} G_j^2 .$$

By making use of Lemma B we prove the convergence of the first term in (37). Denote

$$V_n(x) = \sum_{j=0}^{N(n)} (\hat{a}_{jn} - \mathrm{E}\,\hat{a}_{jn}) g_j(x) .$$

Observe

$$V_{n+1}(x) = V_n(x) + u_n(x) + w_n(x) ,$$

where

$$u_n(x) = \gamma_n \sum_{j=0}^{N(n+1)} [Y_{n+1} g_j(X_{n+1}) - \mathrm{E}\,(Y_{n+1} g_j(X_{n+1}))] g_j(x) ,$$

$$w_n(x) = (1-\gamma_n) \sum_{j=N(n)+1}^{N(n+1)} (\hat{a}_{jn} - \mathrm{E}\,\hat{a}_{jn}) g_j(x) - \gamma_n \sum_{j=0}^{N(n)} (\hat{a}_{jn} - \mathrm{E}\,\hat{a}_{jn}) g_j(x) .$$

Thus,

$$\mathrm{E}\,(V_{n+1}^2(x) | X_1, Y_1, X_2, Y_2, \cdots, X_n, Y_n) = V_n^2(x) + B_n(x) ,$$

where

$$B_n(x) = \mathrm{E}\,u_n^2(x) + w_n^2(x) + 2V_n(x) w_n(x) .$$

Now it will be verified that $\sum_{n=1}^{\infty} \mathrm{E}\,|B_n(x)| < \infty$ for every $x \in \mathcal{X}$. By Cauchy's inequality and (7) we obtain

$$\text{(39)}\qquad \mathrm{E}\, u_n^2(x) \leqq \gamma_n^2 \sum_{j=0}^{N(n+1)} \operatorname{var}\,[Y_{n+1} g_j(X_{n+1})] \sum_{j=0}^{N(n+1)} G_j^2$$

$$\leqq \gamma_n^2\, \mathrm{E}\, Y_{n+1}^2 \left[\sum_{j=0}^{N(n+1)} G_j^2\right]^2 \leqq \gamma_n^2 (K_1 + K_2 c_{n+1}^2) \left[\sum_{j=0}^{N(n+1)} G_j^2\right]^2 .$$

Using Cauchy's inequality again we get

$$w_n^2(x) \leqq 2(1-\gamma_n)^2 \sum_{j=N(n)}^{N(n+1)} (\hat{a}_{jn} - \mathrm{E}\,\hat{a}_{jn})^2 \sum_{j=N(n)}^{N(n+1)} G_j^2$$

$$+ 2\gamma_n^2 \sum_{j=0}^{N(n+1)} (\hat{a}_{jn} - \mathrm{E}\,\hat{a}_{jn})^2 \sum_{j=0}^{N(n+1)} G_j^2 .$$

Therefore

$$\text{(40)}\qquad \mathrm{E}\, w_n^2(x) \leqq 2(1-\gamma_n)^2 d_n \left[\sum_{j=N(n)}^{N(n+1)} G_j^2\right]^2 + 2\gamma_n^2 d_n \left[\sum_{j=0}^{N(n+1)} G_j^2\right]^2 .$$

In turn

$$|V_n(x) w_n(x)| \leqq \gamma_n \left[\sum_{j=0}^{N(n)} (\hat{a}_{jn} - \mathrm{E}\,\hat{a}_{jn}) g_j(x)\right]^2$$

$$+ \left|(1-\gamma_n) \sum_{j=0}^{N(n)} (\hat{a}_{jn} - \mathrm{E}\,\hat{a}_{jn}) g_j(x) \sum_{j=N(n)}^{N(n+1)} (\hat{a}_{jn} - \mathrm{E}\,\hat{a}_{jn}) g_j(x)\right| .$$

Applying Schwartz's and Cauchy's inequalities one gets

$$\text{(41)}\qquad \mathrm{E}\,|V_n(x) w_n(x)| \leqq \gamma_n \sum_{j=0}^{N(n)} \mathrm{E}\,(\hat{a}_{jn} - a_{jn})^2 \sum_{j=0}^{N(n)} G_j^2$$

$$+ |1-\gamma_n| \left[\sum_{j=0}^{N(n)} \mathrm{E}\,(\hat{a}_{jn} - a_{jn})^2 \sum_{j=0}^{N(n)} G_j^2\right.$$

$$\left. \times \sum_{j=N(n)}^{N(n+1)} \mathrm{E}\,(\hat{a}_{jn} - a_{jn})^2 \sum_{j=N(n)}^{N(n+1)} G_j^2\right]^{1/2}$$

$$\leqq \gamma_n d_n \left[\sum_{j=0}^{N(n)} G_j^2\right]^2 + |1-\gamma_n| d_n \sum_{j=0}^{N(n)} G_j^2 \sum_{j=N(n)}^{N(n+1)} G_j^2 .$$

In view of assumptions (33), (34) and (35), inequalities (39), (40) and (41) imply that $\sum_{n=1}^{\infty} \mathrm{E}\,|B_n(x)|$ is finite for every $x \in \mathcal{X}$. Consequently $V_n^2(x)$ converges to a finite limit almost surely as n tends to infinity. By Lemma 1 the limit is zero. The proof has been completed.

LEMMA 4. *If*

$$\text{(42)}\qquad \sum_{n=1}^{\infty} n^{-2} c_n^2 \left[\sum_{j=0}^{M(n)} G_j^2\right]^2 < \infty ,$$

then

$$\text{(43)}\qquad c_n |\hat{f}_n(x) - f(x)| \xrightarrow{n} 0$$

with probability one at every point at which (31) *holds.*

PROOF. By virtue of Lemma 2 it suffices to show that

$$c_n(\hat{f}_n(x)-\mathrm{E}\,\hat{f}_n(x))=c_n n^{-1}\sum_{i=1}^{n}\sum_{j=0}^{M(n)}(g_j(X_i)-\mathrm{E}\,g_j(X_i))g_j(x)$$

converges to zero with probability one as n tends to infinity. Let

$$\xi_n(x)=c_n\sum_{j=0}^{M(n)}(g_j(X_n)-\mathrm{E}\,g_j(X_n))g_j(x)\ .$$

Obviously

$$\mathrm{E}\,\xi_n^2(x)\leqq c_n^2\sum_{j=0}^{M(n)}\mathrm{var}\,g_j(X_n)\sum_{j=0}^{M(n)}g_j^2(x)\leqq c_n^2\left[\sum_{j=0}^{M(n)}G_j^2\right]^2 .$$

By virtue of the Kolmogorov strong law of large numbers (see Doob [7], p. 127) and (42), the proof is complete.

Finally we are able to establish the strong consistency of (15).

THEOREM 2. *Let* (6), (33), (34) *and* (35) *be satisfied. Let, moreover,*

$$\sum_{n=1}^{\infty}(c_n^2+1)n^{-2}\left[\sum_{j=0}^{M(n)}G_j^2\right]^2<\infty\ . \tag{44}$$

Then

$$|\hat{R}_n(x)-R_n(x)|\overset{n}{\rightarrow}0$$

w.p. 1 *at every point* $x\in\mathcal{X}$ *at which* $f(x)>0$, (24) *and* (31) *hold.*

5. The rate of the convergence of $\{d_n\}$

Here we show that the sequence $\{d_n\}$ defined by (20) converges to zero and is bounded by a power sequence. While proving Theorem 3 we use arguments similar to those used in Dupač [8].

THEOREM 3. *Let conditions* (6) *and* (7) *be satisfied and let*

$$\gamma_n=\delta n^{-r}\ ,\qquad \delta>0,\ 0<r<1\ , \tag{45}$$

$$\sup_x|t_{n+1}(x)-t_n(x)|=O(n^{-p})\ ,\qquad r<p\ , \tag{46}$$

$$c_n=O(n^q)\ ,\qquad 2q^+<r\ , \tag{47}$$

where $q^+=\max(0,q)$. *Then*

$$d_n=O(n^{-s})\ , \tag{48}$$

where

$$s=\begin{cases} 2(p-r) & \text{for } r\geqq 2(p+q^+)/3 \\ r-2q^+ & \text{otherwise}. \end{cases}$$

PROOF. To begin with, let us observe that

$$\mathrm{E}\,\{Y_{n+1}g_j(X_{n+1})\,|\,\hat{a}_{j1},\cdots,\hat{a}_{jn}\}=a_{j,n+1}\,. \tag{49}$$

By (7), (8) and (47),

$$\mathrm{var}\,[Y_{n+1}g_j(X_{n+1})\,|\,\hat{a}_{j1},\cdots,\hat{a}_{jn}]\leqq G_j^2(K_1+K_3n^{2q})\,. \tag{50}$$

From (3), (6), (8), (10) and (46) it follows that

$$|a_{j,n+1}-a_{jn}|\leqq G_j\sup_x|t_{n+1}(x)-t_n(x)|\int|R(x)|f(x)dx\leqq K_4G_jn^{-p}\,. \tag{51}$$

Subtracting $a_{j,n+1}$ on both sides of (11) we get

$$\begin{aligned}\hat{a}_{j,n+1}-a_{j,n+1}=(1-\gamma_n)(\hat{a}_{jn}-a_{jn})-\gamma_n(a_{j,n+1}-Y_{n+1}g_j(X_{n+1}))\\ -(1-\gamma_n)(a_{j,n+1}-a_{jn})\,.\end{aligned}$$

Now after squaring and taking conditional expectations, using (49), (50) and (51) we obtain

$$\begin{aligned}&\mathrm{E}\,\{(\hat{a}_{j,n+1}-a_{j,n+1})^2\,|\,\hat{a}_{j1},\cdots,\hat{a}_{jn}\}\\ &\quad\leqq(1-\gamma_n)(\hat{a}_{jn}-a_{jn})^2+K_4G_jn^{-p}|\hat{a}_{jn}-a_{jn}|\\ &\quad\quad+G_j^2(K_1\gamma_n^2+K_3\gamma_n^2n^{2q}+K_5n^{-2p})\,,\end{aligned} \tag{52}$$

for sufficiently large n satisfying $|1-\gamma_n|\leqq 1$. It is clear that for every $\varepsilon>0$ and every random variable Z with finite variance

$$2\,\mathrm{E}\,|Z|\leqq\varepsilon^{-1}+\varepsilon\,\mathrm{E}\,Z^2\,.$$

Thus, choosing $\varepsilon=K_6G_j^{-1}\gamma_nn^p$ (for some small K_6) one gets

$$2G_jn^{-p}\,\mathrm{E}\,|\hat{a}_{jn}-a_{jn}|\leqq K_6^{-1}G_j^2\gamma_n^{-1}n^{-2p}+K_6\gamma_n\,\mathrm{E}\,(\hat{a}_{jn}-a_{jn})^2\,. \tag{53}$$

Now taking unconditional expectation on both sides of (52) and using (53) one obtains

$$\begin{aligned}\mathrm{E}\,(\hat{a}_{j,n+1}-a_{j,n+1})^2\leqq(1-K_7n^{-r})\,\mathrm{E}\,(\hat{a}_{jn}-a_{jn})^2\\ +G_j^2(K_8n^{-2r}+K_9n^{2(q-r)}+K_{10}n^{r-2p})\end{aligned}$$

for sufficiently large n. Thus,

$$\mathrm{E}\,(\hat{a}_{j,n+1}-a_{j,n+1})^2\leqq(1-K_7n^{-r})\,\mathrm{E}\,(\hat{a}_{jn}-a_{jn})^2+K_{11}G_j^2n^{-(r+s)}\,. \tag{54}$$

Hence

$$d_{n+1}\leqq(1-K_7n^{-r})d_n+K_{11}n^{-(r+s)}\,.$$

Since $\hat{a}_{j0}=0$, d_1 is finite. A straightforward application of Lemma A completes the proof.

From Theorems 1, 2 and 3 one easily gets the next two concluding ones.

THEOREM 4. *Let* (6), (7) *and* (22) *hold. Let, moreover,* (45), (46) *and* (47) *be satisfied. If*

$$n^{(q^+-1/2)}\sum_{j=0}^{M(n)} G_j^2 \xrightarrow{n} 0 , \tag{55}$$

$$n^{-s/2}\sum_{j=0}^{N(n)} G_j^2 \xrightarrow{n} 0 , \tag{56}$$

then

$$|\hat{R}_n(x)-R_n(x)|\xrightarrow{n} 0$$

in probability at every $x \in \mathcal{X}$ *at which* $f(x)>0$,

$$\sum_{j=0}^{N(n)} a_j g_j(x)-h(x)=o(n^{-q}) \tag{57}$$

and

$$\sum_{j=0}^{M(n)} b_j g_j(x)-f(x)=o(n^{-q^+}) . \tag{58}$$

THEOREM 5. *Let* (6), (7) *and* (22) *be satisfied. Let, moreover,* (45), (46), (47) *and* (56) *be fulfilled. If*

$$\sum_{n=1}^{\infty} n^{-2(1-q^+)}\left[\sum_{j=0}^{M(n)} G_j^2\right]^2<\infty \tag{59}$$

and

$$\sum_{n=1}^{\infty} n^{-(s+r)}\left[\sum_{j=0}^{N(n)} G_j^2\right]^2<\infty , \tag{60}$$

then

$$|\hat{R}_n(x)-R_n(x)|\xrightarrow{n} 0$$

almost surely at every $x \in \mathcal{X}$ *at which* $f(x)>0$ *and both* (57) *and* (58) *hold.*

PROOF. Verifying that (33) is implied by (60), (35) is implied by (56) and (60), one can easily complete the proof.

6. Examples

The following examples illustrate the fact that conclusions of Theorems 4 and 5 are valid even if R_n tends to infinity as $n \to \infty$.

Let

$$t_n(x) = (1 + \rho(x)/n)n^q ,$$

where

$$\sup_x |\rho(x)| < \infty ,$$

and let q be unknown despite the fact that $0 < q \leqq Q$, where Q is a known number. Now $c_n = n^q$. One can select $\gamma_n = \delta n^{-2/3}$ and sequences $\{N(n)\}$ and $\{M(n)\}$ of types $\{n^\alpha\}$ and $\{n^\beta\}$, respectively, where α and β are positive numbers. This choice is decided by examples given below. In this case (46) and (48) hold with $p = 1 - q$ and $s = (2 - 6q)/3$.

We shall consider two examples of applicable orthogonal systems.

Hermite orthogonal system

If $\mathcal{X}$ is a real line, we can use a system

$$g_j(x) = (2^j j! \pi^{1/2})^{-1/2} e^{-x^2/2} H_j(x) ,$$

where

$$H_0(x) = 1 , \qquad H_j(x) = (-1)^j e^{x^2}(d^j e^{-x^2}/dx^j) , \quad j = 1, 2, \cdots$$

are Hermite polynomials. It can be found in Szegö ([19], p. 242) that $G_j = K_{12} j^{-1/12}$.

Suppose that series (12) and (18) converge at a point x to $f(x)$ and $h(x)$, respectively. Various conditions for the pointwise convergence of orthogonal expansions with the Hermite system can be found in Sansone [16]. Nevertheless, we mention here that the series under consideration converge to $f(x)$ and $h(x)$ at every differentiability point of f and h, respectively.

One can verify that conditions (22), (55) and (56) of Theorem 4 imposed on sequences $\{N(n)\}$ and $\{M(n)\}$ are satisfied for $\alpha < (2 - 6Q)/5$ and $\beta < (3 - 6Q)/5$. In turn, restrictions (22), (59) and (60) of Theorem 5 are fulfilled for $\alpha < (1 - 6Q)/5$ and $\beta < (3 - 6Q)/5$.

We are now interested in assumption (57). Let us assume that the function

$$\zeta(x) = e^{x^2/2} d^m (e^{-x^2/2} h(x))/dx^m$$

exists and is square integrable. By Schwartz's [17] result

$$|a_j| \leqq K_{13}(2j)^{-m/2} ,$$

where K_{13} is the L_2 norm of ζ. Hence, at every point x at which the series in (18) converges to $h(x)$,

$$\left|h(x)-\sum_{j=0}^{N(n)} a_j g_j(x)\right| = \left|\sum_{j=N(n)+1}^{\infty} a_j g_j(x)\right|$$
$$\leqq K_{14} \sum_{j=N(n)+1}^{\infty} j^{-(m+1/6)/2} \leqq K_{15} n^{-(m-11/6)\alpha/2} ,$$

which leads to $\alpha > 12Q/(6m-11)$. Similar result can be given for (58).

Legendre orthogonal system

If $\mathscr{X}=[-1, 1]$ we can apply the Legendre system

$$g_j(x)=(j+1/2)^{1/2}P_j(x) ,$$

where

$$P_0(x)=1 , \qquad P_j(x)=(2^j j!)^{-1}[d^j(x^2-1)^j/dx^j] , \quad j=1, 2, \cdots$$

are Legendre polynomials. In this case $G_j=K_{16}j^{1/2}$ (see Szegö ([19], p. 164)).

Criterion for the pointwise convergence of series (12) and (18) are given in Sansone [16]. In particular, the series converge to $f(x)$ and $h(x)$ at every point at which f and h satisfy the Lipschitz condition of a positive order.

By Jackson's theorem, see Sansone ([16], p. 206), if h is of bounded variation,

$$\left|h(x)-\sum_{j=0}^{n} a_j g_j(x)\right| = O(n^{-1}) ,$$

at every x in the interior of $\mathscr{X}$. In this case, (57) is satisfied for $\alpha > Q$. Similar result is true for (58).

The order restrictions of Theorems 4 and 5 are satisfied for $\alpha < (1-3Q)/6$, $\beta < (1-2Q)/4$ and $\alpha < (1-6Q)/12$, $\beta < (1-2Q)/4$, respectively.

Acknowledgment

The authors thank the referee for careful reading of the original manuscript.

TECHNICAL UNIVERSITY OF WROCŁAW
TECHNICAL UNIVERSITY OF CZĘSTOCHOWA
TECHNICAL UNIVERSITY OF CZĘSTOCHOWA

REFERENCES

[1] Bosq, D. (1969). Sur l'estimation d'une densite multivariee par une serie de fonctions orthogonales, *C. R. Acad. Sci. Paris*, A-B, **268**, 555-557.

[2] Čencov, N. N. (1962). Evaluation of an unknown distribution density from observations, *Soviet Math.*, **3**, 1559-1562.

[3] Chung, K. L. (1954). On a stochastic approximation method, *Ann. Math. Statist.*, **25**, 463-483.

[4] Collomb, G. (1977). Quelques proprietes de la methode du noyau pour l'estimation non parametrique de la regression en un point fixe, *C. R. Acad. Sci. Paris*, A, **285**, 289-292.

[5] Devroye, L. P. (1978). The uniform convergence of nearest neighbor regression function estimators and their application in optimization, *IEEE Trans. Inf. Theory*, **24**, 142-151.

[6] Devroye, L. P. and Wagner, T. J. (1980). On the L_1 convergence of kernel regression function estimators with applications in discrimination, *Zeit. Wahrscheinlichkeitsth.*, **51**, 15-25.

[7] Doob, J. L. (1953). *Stochastic Processes*, John Wiley, New York.

[8] Dupač, V. (1965). A dynamic stochastic approximation method, *Ann. Math. Statist.*, **36**, 1695-1702.

[9] Greblicki, W. (1974). Asymptotically optimal probabilistic algorithms for pattern recognition and identification, Scientific Papers of the Institute of Technical Cybernetics of Wrocław Technical University No. 18, Series: Monographs No. 3, Wrocław (in Polish).

[10] Greblicki, W. and Krzyżak, A. (1980). Asymptotic properties of kernel estimates of a regression function, *J. Statist. Plann. Inf.*, **4**, 81-90.

[11] Kronmal, R. and Tarter, M. (1968). The estimation of probability densities and cumulatives by Fourier series methods, *J. Amer. Statist. Ass.*, **63**, 925-952.

[12] Mirzahmedov, M. A. and Hašimov, S. A. (1972). On some properties of density estimation, *Colloq. Math. Soc. Janos Bolyai*, European meeting of statisticians, Budapest, 535-546.

[13] Nadaraya, E. A. (1964). On estimating regression, *Theory Prob. Appl.*, **9**, 141-142.

[14] Noda, K. (1976). Estimation of a regression function by the Parzen kernel-type density estimators, *Ann. Inst. Statist. Math.*, **28**, A, 221-234.

[15] Rosenblatt, M. (1969). Conditional probability density and regression estimators, *Multivariate Analysis* II, Academic Press, New York, 25-31.

[16] Sansone, G. (1959). *Orthogonal Functions*, Interscience Publishers Inc., New York.

[17] Schwartz, S. C. (1967). Estimation of probability density by an orthogonal series, *Ann. Math. Statist.*, **38**, 1261-1265.

[18] Stone, C. J. (1977). Consistent nonparametric regression, *Ann. Statist.*, **5**, 595-645.

[19] Szegö, G. (1959). *Orthogonal Polynomials*, Amer. Math. Soc. Coll. Publ., 23.

[20] Van Ryzin, J. (1969). On strong consistency of density estimates, *Ann. Math. Statist.*, **40**, 1765-1772.

[21] Wasan, M. T. (1969). *Stochastic Approximation*, Cambridge University Press.