APPLICATIONES MATHEMATICAE
26,2 (1999), pp. 121–131

W. POPIŃSKI (Warszawa)

LEAST-SQUARES TRIGONOMETRIC REGRESSION ESTIMATION

Abstract. The problem of nonparametric function fitting using the complete orthogonal system of trigonometric functions e_k, $k = 0, 1, 2, \ldots$, for the observation model $y_i = f(x_{in}) + \eta_i$, $i = 1, \ldots, n$, is considered, where η_i are uncorrelated random variables with zero mean value and finite variance, and the observation points $x_{in} \in [0, 2\pi]$, $i = 1, \ldots, n$, are equidistant. Conditions for convergence of the mean-square prediction error $(1/n)\sum_{i=1}^{n} E(f(x_{in}) - \widehat{f}_{N(n)}(x_{in}))^2$, the integrated mean-square error $E\|f - \widehat{f}_{N(n)}\|^2$ and the pointwise mean-square error $E(f(x) - \widehat{f}_{N(n)}(x))^2$ of the estimator $\widehat{f}_{N(n)}(x) = \sum_{k=0}^{N(n)} \widehat{c}_k e_k(x)$ for $f \in C[0, 2\pi]$ and $\widehat{c}_0, \widehat{c}_1, \ldots, \widehat{c}_{N(n)}$ obtained by the least squares method are studied.

1. Introduction. Let y_i, $i = 1, \ldots, n$, be observations at equidistant points $x_{in} = 2\pi(i-1)/n$, $i = 1, \ldots, n$, which follow the model $y_i = f(x_{in}) + \eta_i$, where $f : [0, 2\pi] \to \mathbb{R}$ is an unknown function satisfying appropriate conditions characterized in the sequel and η_i, $i = 1, \ldots, n$, are random variables satisfying the conditions $E\eta_i = 0$ and $E\eta_i\eta_j = \sigma_\eta^2 \delta_{ij}$, where $\sigma_\eta^2 > 0$ and δ_{ij} denotes the Kronecker delta.

The functions

$$(1)\qquad e_0(x) = 1,\ e_{2l-1}(x) = \sqrt{2}\sin(lx),\ e_{2l}(x) = \sqrt{2}\cos(lx),\ l = 1, 2, \ldots,$$

constitute a complete orthogonal system in the space $L^2[0, 2\pi]$, normalized so that

$$\frac{1}{2\pi}\int_0^{2\pi} e_k^2(s)\, ds = 1, \qquad k = 0, 1, 2, \ldots$$

1991 *Mathematics Subject Classification*: 62G07, 62F12.

Key words and phrases: Fourier coefficients, trigonometric polynomial, least squares method, regression function, consistent estimator.

In consequence, any function $f \in L^2[0,2\pi]$ has the representation

$$f = \sum_{k=0}^{\infty} c_k e_k, \quad \text{where} \quad c_k = \frac{1}{2\pi}\int_0^{2\pi} f(s) e_k(s)\, ds, \quad k = 0,1,2,\ldots$$

We consider estimators of the Fourier coefficients c_k, $k = 0,1,\ldots,N$, having the form

$$\widehat{c}_{kn} = \frac{1}{n}\sum_{i=1}^{n} y_i e_k(x_{in}), \qquad k = 0,1,\ldots,N. \tag{2}$$

It is well known that in the case of equidistant observation points $x_{in} = 2\pi(i-1)/n$, $i = 1,\ldots,n$, the above defined estimators are for $N = 2m$, $2m+1 \le n$, least squares estimators of the Fourier coefficients c_k, $k = 0,1,\ldots,N$, which is a consequence of the relations (see [1])

$$\frac{1}{n}\sum_{i=1}^{n} e_k(x_{in}) e_l(x_{in}) = \delta_{kl} \tag{3}$$

for $k,l = 0,1,\ldots,N$, $N = 2m$, $2m+1 \le n$.

Observe that if the regression function f is continuous the estimators $\widehat{c}_{kn}$ of the Fourier coefficients c_k, $k = 0,1,\ldots$, are asymptotically unbiased and consistent in the mean-square sense. Indeed, for fixed k, $0 \le k \le N$, $N = 2m$, $2m+1 \le n$,

$$E(\widehat{c}_{kn} - c_k)^2 = E(\widehat{c}_{kn} - E\widehat{c}_{kn})^2 + (E\widehat{c}_{kn} - c_k)^2$$

and taking into account (2) we immediately obtain

$$E(\widehat{c}_{kn} - E\widehat{c}_{kn})(\widehat{c}_{ln} - E\widehat{c}_{ln}) = \frac{\sigma_\eta^2}{n^2}\sum_{i=1}^{n} e_k(x_{in}) e_l(x_{in}),$$

$$E\widehat{c}_{kn} - c_k = \frac{1}{n}\sum_{i=1}^{n} f(x_{in}) e_k(x_{in}) - c_k,$$

which in view of (3) yields

$$\begin{aligned} E(\widehat{c}_{kn} - c_k)^2 &= \frac{\sigma_\eta^2}{n} + (E\widehat{c}_{kn} - c_k)^2, \\ E\widehat{c}_{kn} - c_k &= \frac{1}{2\pi}\frac{2\pi}{n}\sum_{i=1}^{n} f(x_{in}) e_k(x_{in}) - \frac{1}{2\pi}\int_0^{2\pi} f(s) e_k(s)\, ds. \end{aligned} \tag{4}$$

The above equalities and continuity of f and e_k imply that

$$\lim_{n\to\infty} E\widehat{c}_{kn} - c_k = 0 \quad \text{and} \quad \lim_{n\to\infty} E(\widehat{c}_{kn} - c_k)^2 = 0.$$

In the sequel we shall examine the asymptotic properties of the projection estimator of the regression function

$$\widehat{f}_N(x) = \sum_{k=0}^{N} \widehat{c}_{kn} e_k(x).$$

According to the Jackson theorem [6] for any 2π-periodic continuous function (i.e. for $f \in C[0, 2\pi]$ satisfying $f(0) = f(2\pi)$) there exists a trigonometric polynomial of degree l

$$T_l(x) = a_0 + \sum_{k=1}^{l} (a_k \cos(kx) + b_k \sin(kx)),$$

where $a_l^2 + b_l^2 \neq 0$, such that

$$\sup_{0 \le s \le 2\pi} |f(s) - T_l(s)| \le 12\omega(1/l, f),$$

where $\omega(\delta, f)$ (for $\delta > 0$) denotes the modulus of continuity of f.

2. Asymptotic mean-square prediction error. Consider first the mean-square prediction error of the estimator $\widehat{f}_N$, defined by

$$D_{nN} = \frac{1}{n} \sum_{i=1}^{n} E(f(x_{in}) - \widehat{f}_N(x_{in}))^2.$$

In view of the orthogonality relations (3) the standard squared bias plus variance decomposition yields

$$D_{nN} = \frac{1}{n} \sum_{i=1}^{n} (f(x_{in}) - E\widehat{f}_N(x_{in}))^2 + \sigma_\eta^2 \frac{N+1}{n}. \tag{5}$$

It can be easily seen that for $N = 2m$, $2m + 1 \le n$, the inequality

$$\frac{1}{n} \sum_{i=1}^{n} (f(x_{in}) - E\widehat{f}_N(x_{in}))^2 \le \frac{1}{n} \sum_{i=1}^{n} (f(x_{in}) - T_l(x_{in}))^2$$

holds for any trigonometric polynomial T_l of degree $l \le m$. Consequently, using (5) and applying the Jackson theorem we immediately see that for a 2π-periodic function $f \in C[0, 2\pi]$ we have $\lim_{n\to\infty} D_{nN(n)} = 0$ on condition that $\lim_{n\to\infty} N(n) = \infty$ and $\lim_{n\to\infty} N(n)/n = 0$.

From the equality (5) we see that for any regression function f the condition $\lim_{n\to\infty} N(n)/n = 0$ is also necessary for $\lim_{n\to\infty} D_{nN(n)} = 0$. For a continuous regression function f which is not a trigonometric polynomial of any finite order $\lim_{n\to\infty} D_{nN(n)} = 0$ also implies that $\lim_{n\to\infty} N(n) = \infty$. Indeed, if we assume that there exists a subsequence m_k, $k = 1, 2, \ldots$, such that the sequence $N(m_k)$ is bounded, then there also exists a subsequence n_l

such that $N(n_l) = M, l = 1, 2, \ldots$ In consequence, putting $f_M = \sum_{k=0}^{M} c_k e_k$ we would have

$$\frac{1}{n_l}\sum_{i=1}^{n_l}(f(x_{in_l}) - E\widehat{f}_{N(n_l)}(x_{in_l}))^2$$
$$= \frac{1}{n_l}\sum_{i=1}^{n_l}(f(x_{in_l}) - f_M(x_{in_l}))^2 + \frac{1}{n_l}\sum_{i=1}^{n_l}(f_M(x_{in_l}) - E\widehat{f}_{N(n_l)}(x_{in_l}))^2$$
$$+ \frac{2}{n_l}\sum_{i=1}^{n_l}(f_M(x_{in_l}) - E\widehat{f}_{N(n_l)}(x_{in_l}))(f(x_{in_l}) - f_M(x_{in_l}))$$

and since the functions f and f_M are continuous the second and third terms on the right-hand side would converge to zero because by the Schwarz inequality and (3),

$$\frac{1}{n_l}\sum_{i=1}^{n_l}(f_M(x_{in_l}) - E\widehat{f}_{N(n_l)}(x_{in_l}))^2 \le \frac{1}{n_l}\sum_{i=1}^{n_l}\sum_{k=0}^{M}(c_k - E\widehat{c}_{kn_l})^2\sum_{k=0}^{M}e_k^2(x_{in_l})$$
$$\le (M+1)\sum_{k=0}^{M}(c_k - E\widehat{c}_{kn_l})^2.$$

Consequently, we would have

$$\lim_{l\to\infty}\frac{1}{n_l}\sum_{i=1}^{n_l}(f(x_{in_l}) - E\widehat{f}_{N(n_l)}(x_{in_l}))^2 = \frac{1}{2\pi}\int_0^{2\pi}(f(s) - f_M(s))^2\,ds > 0,$$

so the sequence $D_{nN(n)}$ would not converge to zero.

The above conclusions allow us to formulate the following theorem.

Theorem 2.1. *If the regression function f is continuous, 2π-periodic and not a trigonometric polynomial of any finite order, then the projection estimator $\widehat{f}_{N(n)}$ is consistent in the sense of the mean-square prediction error, i.e.*

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}E(f(x_{in}) - \widehat{f}_{N(n)}(x_{in}))^2 = 0,$$

if and only if the sequence of natural numbers $N(n)$, $n = 1, 2, \ldots$, satisfies

$$\lim_{n\to\infty}N(n) = \infty, \qquad \lim_{n\to\infty}N(n)/n = 0.$$

Furthermore, since for a function f satisfying the Lipschitz condition with exponent $0 < \alpha \le 1$ we have $\omega(\delta, f) \le L\delta^\alpha$, where $L > 0$, it is easy to see that the following corollary holds.

Corollary 2.1. *Assume that the regression function f is 2π-periodic and satisfies the Lipschitz condition with exponent $0 < \alpha \le 1$. If the* *sequence of even natural numbers* $N(n)$, $n = 1, 2, \ldots,$ *satisfies* $N(n) \sim n^{1/(1+2\alpha)}$ (*i.e.* $r_1 \le n^{-1/(1+2\alpha)} N(n) \le r_2$ *for* $r_1, r_2 > 0$), *then*

$$\frac{1}{n}\sum_{i=1}^{n} E(f(x_{in}) - \widehat{f}_{N(n)}(x_{in}))^2 = O(n^{-2\alpha/(1+2\alpha)}).$$

The above results complement the ones presented in [2] which were proved for a more general fixed point design under more restrictive assumptions on the smoothness of the regression function.

3. Convergence of the integrated mean-square error. The formula for the bias of the estimator $\widehat{c}_{kn}$ (see (4)) can be rewritten in the form

$$E\widehat{c}_{kn} - c_k = \int_0^{2\pi} f e_k \, d(F_n - F),$$

where F denotes the uniform distribution function on $[0, 2\pi]$ and F_n the empirical distribution function of the "sample" x_{in}, $i = 1, \ldots, n$. If f is absolutely continuous, then integrating by parts gives

$$E\widehat{c}_{kn} - c_k = \int_0^{2\pi} (f' e_k + f e_k')(F_n - F),$$

and since $\sup |F - F_n| \le 1/n$ this yields

$$|E\widehat{c}_{kn} - c_k| \le \frac{1}{n}\Big(\int_0^{2\pi} |f' e_k| + \int_0^{2\pi} |f e_k'|\Big)$$

and finally in view of definition (1) we obtain

$$(6)\qquad \begin{aligned} (E\widehat{c}_{0n} - c_0)^2 &\le \frac{1}{n^2}\|f'\|_1^2, \\ (E\widehat{c}_{kn} - c_k)^2 &\le \frac{4}{n^2}(\|f'\|_1^2 + l^2\|f\|_1^2), \end{aligned}$$

for $k = 2l-1, 2l$, $l = 1, \ldots, m$, $2m+1 \le n$, where $\|*\|_p$ denotes the $L^p[0, 2\pi]$ norm.

Now consider the integrated mean-square error of the estimator $\widehat{f}_N$,

$$R_{nN} = \frac{1}{2\pi} E \int_0^{2\pi} (f - \widehat{f}_N)^2 = p_N + \sum_{k=0}^{N} E(\widehat{c}_{kn} - c_k)^2,$$

where $p_N = \sum_{k=N+1}^{\infty} c_k^2$. According to (4) we can write

$$(7)\qquad R_{nN} = p_N + \sum_{k=0}^{N} (E\widehat{c}_{kn} - c_k)^2 + \sigma_\eta^2 \frac{N+1}{n}.$$

For $N = 2m$, $2m+1 \le n$, taking into account (6) we obtain

$$(8)\quad \sum_{k=0}^{N} (E\widehat{c}_{kn} - c_k)^2 \le \frac{1}{n^2}\Big[(8m+1)\|f'\|_1^2 + 8\|f\|_1^2 \sum_{l=1}^{m} l^2\Big] \le \frac{1}{n^2}\Big[(4N+1)\|f'\|_1^2 + \frac{N(N+1)(N+2)}{3}\|f\|_1^2\Big]$$

since $\sum_{l=1}^{m} l^2 = m(m+1)(2m+1)/6$. The above estimate together with (7) allows us to formulate the following theorem.

THEOREM 3.1. *If the sequence of even natural numbers* $N(n)$, $n = 1, 2, \ldots,$ *satisfies*

$$\lim_{n\to\infty} N(n) = \infty, \qquad \lim_{n\to\infty} N(n)^{3/2}/n = 0,$$

then the projection estimator $\widehat{f}_{N(n)}$ *of the absolutely continuous regression function* f *is consistent in the sense of the integrated mean-square error*, *i.e.*

$$\lim_{n\to\infty} E\|f - \widehat{f}_{N(n)}\|_2^2 = 0.$$

Rafajłowicz [10] proved that a theorem similar to 3.1 holds in the case of regression functions for which the error of uniform approximation by trigonometric polynomials tends to zero as the polynomial degree increases, i.e. it holds for continuous and 2π-periodic regression functions [6]. It should be noted that our theorem extends the result from [10] since it is true for nonperiodic regression functions.

In order to obtain a result concerning the convergence rate of the integrated mean-square error we need the following lemma.

LEMMA 3.1. *If the function* f *is absolutely continuous*, *then for* $N = 2m$, $m = 1, 2, \ldots,$

$$p_N \le \frac{5\|f'\|_1^2}{\pi^2 N}.$$

Proof. Integrating by parts gives for $l = 1, 2, \ldots,$

$$c_{2l-1} = \frac{1}{\sqrt{2}\, l\pi}\Big[f(0) - f(2\pi) + \int_0^{2\pi} f'(s)\cos(ls)\, ds\Big],$$

$$c_{2l} = -\frac{1}{\sqrt{2}\, l\pi} \int_0^{2\pi} f'(s)\sin(ls)\, ds,$$

and in consequence

$$|c_{2l-1}| \le \frac{2}{\sqrt{2}\, l\pi}\|f'\|_1, \qquad |c_{2l}| \le \frac{1}{\sqrt{2}\, l\pi}\|f'\|_1.$$

Hence, for $N = 2m$,

$$p_N = \sum_{k=N+1}^{\infty} c_k^2 = \sum_{l=m+1}^{\infty} (c_{2l-1}^2 + c_{2l}^2) \le \frac{5\|f'\|_1^2}{2\pi^2} \sum_{l=m+1}^{\infty} \frac{1}{l^2}$$
$$\le \frac{5\|f'\|_1^2}{2\pi^2} \sum_{l=m+1}^{\infty} \frac{1}{l(l-1)} = \frac{5\|f'\|_1^2}{2\pi^2} \cdot \frac{1}{m}. \ \blacksquare$$

According to (7), (8) and by Lemma 3.1 we have for $N = 2m$,

$$R_{nN} \le \frac{5\|f'\|_1^2}{\pi^2 N} + \frac{1}{n^2}\left[(4N+1)\|f'\|_1^2 + \frac{N(N+1)(N+2)}{3}\|f\|_1^2\right] + \sigma_\eta^2 \frac{N+1}{n},$$

which can be rewritten in the form

$$R_{nN} \le \frac{A}{N} + \frac{1}{n^2}(BN + CN^3) + \frac{DN}{n},$$

where $A, B, C, D > 0$ are suitably chosen constants. From the last inequality it is easy to see that the following corollary holds.

COROLLARY 3.1. *If the regression function f is absolutely continuous and the sequence of even natural numbers $N(n)$, $n = 1, 2, \ldots,$ satisfies $N(n) \sim n^{1/2}$ (i.e. $r_1 \le n^{-1/2}N(n) \le r_2$ for $r_1, r_2 > 0$), then*

$$E\|f - \widehat{f}_{N(n)}\|_2^2 = O(n^{-1/2}).$$

In papers on wavelet methods of nonparametric function estimation (e.g. [5]) one can find results giving the IMSE decay rate of a wavelet projection estimator for an equidistant point design.

4. Pointwise mean-square consistency of the estimator. In this section we derive sufficient conditions for pointwise mean-square consistency of the projection estimator $\widehat{f}_N$ considered.

If the Fourier series of f converges to $f(x)$ at some $x \in [0, 2\pi]$, then

$$E(f(x) - \widehat{f}_N(x))^2 = E\Big(\sum_{k=0}^{N}(c_k - \widehat{c}_{kn})e_k(x)\Big)^2 + r_N^2(x)$$
$$+ 2r_N(x)\sum_{k=0}^{N}(c_k - E\widehat{c}_{kn})e_k(x),$$

where $r_N(x) = \sum_{k=N+1}^{\infty} c_k e_k(x)$. From the Cauchy–Schwarz inequality it further follows that

$$E(f(x)-\widehat{f}_N(x))^2 \le \sum_{k=0}^{N} E(c_k-\widehat{c}_{kn})^2 \sum_{k=0}^{N} e_k^2(x) + r_N^2(x)$$
$$+ 2|r_N(x)| \Big(\sum_{k=0}^{N}(c_k - E\widehat{c}_{kn})^2\Big)^{1/2} \Big(\sum_{k=0}^{N} e_k^2(x)\Big)^{1/2},$$

and according to (4) since $\sum_{k=0}^{N} e_k^2(x) = N+1$ for $N=2m$, $m \ge 0$, $x \in [0,2\pi]$, we finally have

$$(9) \qquad E(f(x)-\widehat{f}_N(x))^2 \le (N+1)\sum_{k=0}^{N}(c_k - E\widehat{c}_{kn})^2 + \sigma_\eta^2 \frac{(N+1)^2}{n}$$
$$+ 2|r_N(x)|(N+1)^{1/2}\Big(\sum_{k=0}^{N}(c_k - E\widehat{c}_{kn})^2\Big)^{1/2} + r_N^2(x).$$

If we assume that the regression function f is absolutely continuous, then since such a function is both continuous and of bounded variation in $[0,2\pi]$, its Fourier series converges uniformly to f in $(\delta, 2\pi-\delta)$ for $\delta > 0$ (see Corollary 2.62 [11]), so that $\lim_{n\to\infty} r_{N(n)}(x) = 0$ uniformly for $x \in (\delta, 2\pi - \delta)$ if $\lim_{n\to\infty} N(n) = \infty$. Hence, the estimates in (8) and (9) imply that the following theorem holds.

Theorem 4.1. *If the sequence $N(n)$, $n = 1,2,\ldots,$ of even natural numbers satisfies*

$$\lim_{n\to\infty} N(n) = \infty, \qquad \lim_{n\to\infty} N(n)^2/n = 0,$$

then for any $\delta > 0$ the projection estimator $\widehat{f}_{N(n)}$ of the absolutely continuous regression function f is uniformly consistent in the sense of the pointwise mean-square error in the interval $(\delta, 2\pi-\delta)$, i.e.

$$\lim_{n\to\infty} E(f(x)-\widehat{f}_{N(n)}(x))^2 = 0 \quad \textit{uniformly for } x \in (\delta, 2\pi-\delta).$$

It should be noted that for an absolutely continuous and 2π-periodic regression function the pointwise mean-square error converges uniformly in the closed interval $[0,2\pi]$, which follows from the fact that then the Fourier series of the regression function converges uniformly in this interval [11]. Let us also remark that Rafajłowicz [10] obtained sufficient conditions for uniform pointwise mean-square consistency of $\widehat{f}_N$ in $[0,2\pi]$ only for 2π-periodic regression functions.

5. Selecting the regression order. In this section we study a method of selecting a good value of N from the data, namely, the one based on

Mallows's C_p criterion [7]

$$C(N) = \frac{1}{n}\sum_{i=1}^{n}(y_i - \widehat{f}_N(x_{in}))^2 + \frac{2N\widehat{\sigma}^2}{n}, \tag{10}$$

where $\widehat{\sigma}^2$ is any consistent estimator of σ_η^2 (see [3], [4] for examples of such estimators). One selects a value $\widehat{N}_n$ by minimizing $C(N)$ over the integers $0 \le N \le n-1$.

We obtain results on asymptotic properties of this method for selecting the trigonometric regression order N. First, we prove the following theorem.

THEOREM 5.1. *Assume that*

(a) *the function f is absolutely continuous and is not a trigonometric polynomial of any finite order*,

(b) *there exists a sequence of nonnegative real numbers ε_k, $k = 0, 1, \ldots$, such that the sequence $(k+1)\varepsilon_k$ is nonincreasing and*

$$|c_k| \le \varepsilon_k, \qquad \sum_{k=0}^{\infty}\varepsilon_k < \infty,$$

(c) $\mu_4 = \sup E\eta_i^4 < \infty$,

(d) $\widehat{\sigma}^2 \xrightarrow{p} \sigma_\eta^2$ *as* $n \to \infty$.

If $\widehat{N}_n$ is the minimizer of (10), *then*

$$\int_0^{2\pi}(f - \widehat{f}_{\widehat{N}_n})^2 = O_p(n^{-1/2}).$$

Proof. Under the above assumptions Theorem 2 of [9] holds, which asserts that for the loss function $r_n(N) = \int_0^{2\pi}(f - \widehat{f}_N)^2$ we have

$$\frac{r_n(\widehat{N}_n)}{\min_{0\le N\le n-1} r_n(N)} \xrightarrow{p} 1, \qquad n \to \infty, \tag{11}$$

even if the absolute continuity assumption is not satisfied. If this assumption is satisfied we have, by Corollary 3.1,

$$\min_{0\le N\le n-1} E\int_0^{2\pi}(f - \widehat{f}_N)^2 = O(n^{-1/2}),$$

and consequently

$$\min_{0\le N\le n-1} \int_0^{2\pi}(f - \widehat{f}_N)^2 = O_p(n^{-1/2}).$$

The last equality together with (11) implies that $r_n(\widehat{N}_n) = O_p(n^{-1/2})$, which completes the proof. ■

We can also consider the loss function

$$d_n(N) = \frac{1}{n}\sum_{i=1}^{n}(f(x_{in}) - \widehat{f}_N(x_{in}))^2.$$

Under the assumptions of Theorem 5.1 (without the absolute continuity assumption) Theorem 2 of [9] also assures that

$$\frac{d_n(\widehat{N}_n)}{\min_{0\le N\le n-1} d_n(N)} \xrightarrow{p} 1, \qquad n\to\infty.$$

Thus, if the regression function is 2π-periodic we can prove analogously using Theorem 2.1 that $\lim_{n\to\infty} d_n(\widehat{N}_n) \stackrel{p}{=} 0$. If the 2π-periodic regression function f also satisfies the Lipschitz condition with exponent $0<\alpha\le 1$, then using Corollary 2.1 we can prove that $d_n(\widehat{N}_n) = O_p(n^{-2\alpha/(1+2\alpha)})$.

A sequence of real numbers ε_k, $k=0,1,\ldots$, which satisfies the conditions of assumption (b) in the above theorem exists for example when (see [9])

$$\sum_{k=0}^{\infty} |c_k| \ln\ldots\ln(k+1) < \infty,$$

where $\ln\ldots\ln(k+1)$ denotes a multiple logarithm of $k+1$.

Results concerning other asymptotic properties of the estimator considered, e.g. the limit distribution of its integrated squared error, are obtained in [8].

References

[1] B. Droge, *On finite-sample properties of adaptive least-squares regression estimates*, Statistics 24 (1993), 181–203.

[2] R. L. Eubank and P. Speckman, *Convergence rates for trigonometric and polynomial-trigonometric regression estimators*, Statist. Probab. Lett. 11 (1991), 119–124.

[3] T. Gasser, L. Sroka and C. Jennen-Steinmetz, *Residual variance and residual pattern in nonlinear regression*, Biometrika 73 (1986), 625–633.

[4] P. Hall, J. W. Kay and D. M. Titterington, *Asymptotically optimal difference-based estimation of variance in nonparametric regression*, ibid. 77 (1990), 521–528.

[5] P. Hall and P. Patil, *On wavelet methods for estimating smooth functions*, J. Bernoulli Soc. 1 (1995), 41–58.

[6] G. G. Lorentz, *Approximation of Functions*, Holt, Reinehart & Winston, New York, 1966.

[7] C. L. Mallows, *Some comments on C_p*, Technometrics 15 (1973), 661–675.

[8] E. Nadaraya, *Limit distribution of the integrated squared error of trigonometric series regression estimator*, Proc. Georgian Acad. Sci. Math. 1 (1993), 221–237.

[9] B. T. Polyak and A. B. Tsybakov, *Asymptotic optimality of the C_p criterion in projection type estimation of regression functions*, Teor. Veroyatnost. Primenen. 35 (1990), 305–317 (in Russian).

[10] E. Rafajłowicz, *Nonparametric least-squares estimation of a regression function*, Statistics 19 (1988), 349–358.
[11] A. Zygmund, *Trigonometrical Series*, Dover, 1955.

Waldemar Popiński
Department of Standards
Central Statistical Office
Al. Niepodległości 208
00-925 Warszawa, Poland
E-mail: w.popinski@stat.gov.pl

Received on 30.10.1997;
revised version on 18.1.1999