JOURNAL OF MULTIVARIATE ANALYSIS 25, 100–110 (1988)

Consistent Nonparametric Multiple Regression: The Fixed Design Case

ALEXANDER A. GEORGIEV*

Institute of Engineering Cybernetics, Technical University of Wroclaw, Wroclaw, Poland, and Lehrstuhl für Elektrische Steuerung und Regelung, Ruhr-Universität Bochum, Bochum, West Germany

Communicated by P. R. Krishnaiah

Consider the nonparametric regression model $Y_i^{(n)} = g(x_i^{(n)}) + \varepsilon_i^{(n)}$, $i = 1, ..., n$, where g is an unknown function, the design points $x_i^{(n)}$ are known and nonrandom, and $\varepsilon_i^{(n)}$'s are independent random variables. The regressor is assumed to take values in $A \subset R^p$, and the regressand to be real valued. This paper studies the behavior of the general nonparametric estimate

$$g_n(x) = \sum_{i=1}^{n} w_{ni}(x)\, Y_i^{(n)},$$

where the weight function w_{ni} is of the form $w_{ni}(x) = w_{ni}(x; x_1^{(n)}, ..., x_n^{(n)})$. Under suitable conditions, it is shown that the general linear smoother g_n for the unknown regression function g is asymptotically pointwise unbiased, weak, mean square and complete consistent, and asymptotically normal. The results of the limit theorems can be applied to extend or improve the conditions of the estimates with various particular weights w_{ni} including all those known in the literature.

1. INTRODUCTION

Let p be a natural number and A be a compact set in R^p. Consider observations

$$Y_i^{(n)} = g(x_i^{(n)}) + \varepsilon_i^{(n)}, \qquad i = 1, ..., n, \tag{1.1}$$

Received April 28, 1986; revised May 10, 1987.

AMS 1980 subject classifications: Primary 62G05, 62J02; Secondary 60F15.

Key words and phrases: nonparametric estimation, regression function, kernel estimate, nearest neighbor estimate, orthogonal series estimate, splines, multivariate, law of large numbers, central limit theorem, consistency.

* Research partially supported by CPBP Project 02.13 and partially supported by Alexander von Humboldt Foundation.

0047-259X/88 $3.00

where $x_1^{(n)}, ..., x_n^{(n)} \in A$ are design points, g is a bounded real valued function on A, and the random errors $\varepsilon_1^{(n)}, ..., \varepsilon_n^{(n)}$ are assumed independently, but not necessarily identically, distributed with either $E\varepsilon_i^{(n)}=0$, $i=1, ..., n$, or the law $(\varepsilon_i^{(n)})$ is symmetric about zero. The goal is to estimate g from observations. If g is assumed to belong to a space of functions which is parametrized by a finite number of parameters (e.g., the polynomials of degree m or less) then standard regression techniques may be applied. However, if we are only willing to assume that g possesses some smoothing properties, then the use of a nonparametric regression technique is in order. To estimate the unknown function g (only for $p=1$), many classes of estimates have been proposed, including the kernel method (Priestley and Chao [30], Clark [6], Gasser and Müller [13], Cheng and Lin [4, 5], Georgiev [15, 17], etc.), the nearest neighbor method (Greblicki [23], Georgiev [18, 19], Georgiev and Greblicki [22]), the spline method (Wahba [40], Rice and Rosenblatt [33], Silverman [37], etc.) and the orthogonal series method (Rutkowski [35], Rafajłowicz [31]). For the multivariate case ($p>1$) Ahmad and Lin [1], Georgiev [16], and Gałkowski and Rutkowski [11] discuss consistency of the kernel estimates, and Gałkowski and Rutkowski [12], and Rafajłowicz [32] establish consistency of the orthogonal series method. Generally, these estimates are linear in the $Y_i^{(n)}$'s.

The present paper investigates the following *general linear smoother* as an estimate of g, defined by formula

$$g_n(x)=\sum_{i=1}^{n} w_{ni}(x)\, Y_i^{(n)}, \qquad x \in A \subset R^p, \tag{1.2}$$

where weight functions $w_{ni}(x)$, $i=1, ..., n$, depend on the fixed design points $x_1^{(n)}, ..., x_n^{(n)}$ and on the number of observations n. The point x may be interpreted as a future value taken on by some x for which Y is not yet observed. The estimate (1.2) was proposed by Georgiev [19] and discussed by Georgiev and Greblicki [22] for dimension $p=1$. The main object of this paper is to prove pointwise laws of large numbers and the central limit theorem for $g_n(x)$ under some regularity conditions met by the unknown function g, the weights $w_{n1}(x), ..., w_{nn}(x)$,, and the random errors $\varepsilon_1^{(n)}, ..., \varepsilon_n^{(n)}$.

For the stochastic design model (or the correlation model) it is assumed that the $X_1, ..., X_n$ are random variables. Freedman [10] has emphasized the distinction between the regression model (1.1) and the correlation model under which $(X_1, Y_1), ..., (X_n, Y_n)$ is regarded as a random sample from the $(p+1)$-variate distribution of (X, Y), where X is a p-vector and Y is a scalar. For the correlation model Stone [39] has discussed a class of nonparametric regression estimates of type (1.2). His results imply that if the weight functions $w_{ni}(x)$, $i=1, ..., n$, satisfy certain conditions (in probability) and addition natural properties, then the nonparametric estimates are consistent in the mean. The results are valid for any distribution of X. Our model (1.1) is different from that of Stone and his results do not include results established in this paper. Moreover, Stone studied global behavior of the nonparametric regression estimates. We investigate local (pointwise) properties of the general linear smoother g_n. For more details on the study of the correlation model the reader should consult bibliographic reviews Collomb [7, 8] and Prakasa Rao [29], among others.

Recently Härdle and Luckhaus [25] and Härdle and Gasser [26] have introduced a robust nonparametric function fitting method. Their estimates are motivated by the theory of M-estimation and of kernel estimation of regression functions.

The paper is organized as follows. In Section 2 the main theorems are stated and discussed. The new results for the estimate of multiple function g allow the improvement of some recent results (for $p>1$) given by Ahmad and Lin [1]. The proofs of the theorems are given in Section 3. Our proofs are extremely simple.

We hope that this paper will achieve two main objects. First, the general results for the estimate g_n will give a useful tool for the analysis of a wide class of nonparametric regression estimates. For particular known weights $w_{n1}(x), ..., w_{nn}(x)$ the statisticians can obtain new sharper results (see the example in Section 2). Second, the established properties of the weights $w_{n1}(x), ..., w_{nn}(x)$ and of the random errors $\varepsilon_1^{(n)}, ..., \varepsilon_n^{(n)}$ should give the reader an insight into the way of constructing new practical, useful weight functions and new estimates.

2. Results

Throughout the next two sections the following assumptions are made. Let $A \subset R^p$ be a compact set. Let $w_{ni}(x)=0$ for $i>n$. To simplify our notation we shall write Y_i, x_i, ε_i for $Y_i^{(n)}$, $x_i^{(n)}$, $\varepsilon_i^{(n)}$.

Our first theorem establishes that $g_n(x)$ is asymptotically unbiased. The theorem holds an array of weight functions $w_{n1}(x), ..., w_{nn}(x)$, x fixed, with the properties

$$\sum_{i=1}^{n} w_{ni}(x) \to 1 \qquad \text{as} \quad n \to \infty, \tag{2.1}$$

and

$$\sum_{i=1}^{n} |w_{ni}(x)| \leqslant B \qquad \text{for all } n. \tag{2.2}$$

THEOREM 1. *Assume the function g is bounded on $A \subset R^p$. Under the conditions* (2.1) *and* (2.2), *and if*

$$\sum_{i=1}^{n} |w_{ni}(x)| \, I_{\{\|x_i - x\| > a\}} \to 0 \qquad as \quad n \to \infty \text{ for all } a > 0, \tag{2.3}$$

then

$$Eg_n(x) \to g(x) \qquad as \quad n \to \infty,$$

at every continuity point $x \in A$ of the function g.

Our next result describes a weak law of large numbers for the estimate g_n. Suppose that random variables ε_i are *uniformly bounded* by a random variable ε in the sense that

$$\sup_i P\{|\varepsilon_i| \geqslant t\} \leqslant P\{|\varepsilon| \geqslant t\} \qquad \text{for all} \quad t > 0. \tag{2.4}$$

Now we are in the position to give the following

THEOREM 2. *Assume the conditions of Theorem* 1, *and in addition that $\varepsilon_1, ..., \varepsilon_n$ are independent random variables with $E\varepsilon_i = 0$, uniformly bounded by a random variable ε with $E|\varepsilon| < \infty$. If*

$$\sup_i |w_{ni}(x)| \to 0 \qquad as \quad n \to \infty, \tag{2.5}$$

then

$$g_n(x) \to g(x) \qquad \text{in probability as } n \to \infty,$$

at every continuity point $x \in A$ of the function g.

EXAMPLE. The above theorem extends and improves one of the recent results given by Ahmad and Lin [1]. They introduced a multidimensional version of the Priestley and Chao estimate, i.e., the estimate (1.2) with weights

$$\tilde{w}_{ni}(x) = K\left(\frac{x - x_i}{a_n}\right) \frac{\Delta(A_i)}{a_n^p}, \tag{2.6}$$

where $A_1, ..., A_n$ is partition of $A = [0, 1]^p$ into n regions such that the volume $\Delta(A_i)$ is of order n^{-1}, $K(u)$ is a known p-dimensional bounded density, $\{a_n\}$ is a sequence of reals converging to zero as $n \to \infty$, and $x_i \in A_i$. Ahmad and Lin [1, p. 171] prove that if $K(u)$ is Lipschitz of order β, $\max_i \Delta(A_i) = O(n^{-1})$ and $na_n^{(1+1/\beta)p} \to \infty$ as $n \to \infty$, then $\tilde{g}_n(x) \to g(x)$ in probability as $n \to \infty$ for all $x \in A$, provided that $g(x)$ is continuous in $[0, 1]^p$. Following that theorem, the best result for the sequence $\{a_n\}$ is for $\beta = 1$, i.e.,

$$na_n^{2p} \to \infty \qquad \text{as} \quad n \to \infty.$$

It is easy to prove, by checking the conditions of Theorem 2, that if $K(u)$ is a continuous probability density function, nonincreasing as $\|u\| \to \infty$, $\max_i \Delta(A_i) = O(n^{-1})$ and $na_n^p \to \infty$ as $n \to \infty$, then $\tilde{g}_n(x) \to g(x)$ in probability as $n \to \infty$ at every continuity point $x \in A$. We remark that the above result improves that of Ahmad and Lin [1]. The sequence $\{a_n\}$ is such that

$$na_n^p \to \infty \qquad \text{as} \quad n \to \infty,$$

and the above conditions is the same for other multivariate weights

$$\hat{w}_{ni}(x) = \frac{1}{a_n^p} \int_{A_i} K\left(\frac{x-u}{a_n}\right) du, \tag{2.7}$$

investigated independently by Ahmad and Lin [1] and Georgiev [16].

Remark. In the above example, the weights (2.6) are constructed based on a positive kernel $K(u)$. There may be some drawback in using positive weights when some smoothness condition is assumed. For example, when the dimension $p = 1$ and x_0 is an endpoint of the support, then the use of nonnegative weights results in weighting values of g_n at x's which lie on one side of x_0 and there is no way of effectively using the smoothness of g to reduce the resulting bias. This problem does not occur in the correlation model setup as the smoothness assumption is not required in establishing Stone's consistency results, see Stone [39], Devroye and Wagner [9], Spiegelman and Sacks [38], and Greblicki *et al.* [24]. Notice that the weights w_{ni} used in this paper are not necessarilly positive.

THEOREM 3. *Assume the conditions of Theorem* 1, *and in addition that* $\varepsilon_1, ..., \varepsilon_n$ *are independent random variables with* $E\varepsilon_i = 0$, $\sup_i E\varepsilon_i^2 < \infty$. *If*

$$\sum_{i=1}^{n} w_{ni}^2(x) \to 0 \qquad \text{as} \quad n \to \infty, \tag{2.8}$$

then

$$E[g_n(x) - g(x)]^2 \to 0 \qquad \text{as} \quad n \to \infty,$$

at every continuity point $x \in A$ *of the function* g.

We can also establish sufficient conditions for strong pointwise consistency of $g_n(x)$.

THEOREM 4. *Assume the conditions of Theorem* 1, *and in addition that* $\varepsilon_1, ..., \varepsilon_n$ *are independent random variables with* $E\varepsilon_i = 0$, $\sup_i E|\varepsilon_i|^r < \infty$, *for some* $r > 2$. *If*

$$\sup_i w_{ni}^2(x) n \operatorname{loglog} n \to 0 \qquad as \quad n \to \infty, \tag{2.9}$$

then

$$g_n(x) \to g(x) \text{ with probability } 1 \text{ as } n \to \infty,$$

at every continuity point $x \in A$ *of the function* g.

We remark that the conclusion of Theorem 4 improves those of Benedetti [2] and Gasser and Müller [13] for dimension $p = 1$. It is noted that they assumed identically distributed random variables with $E\varepsilon_1^4 < \infty$ and other regularity conditions on the particular weights.

In the next two Theorems we give sufficient conditions for complete consistency of the estimate $g_n(x)$ in the sense of Hsu and Robbins [27]. The results for complete consistency imply strong consistency by Boole's inequality.

THEOREM 5. *Assume the conditions of Theorem* 1, *and in addition that* $\varepsilon_1, ..., \varepsilon_n$ *are independent random variables with* $E\varepsilon_i = 0$, *uniformly bounded by a random variable* ε *with* $E|\varepsilon|^{1+1/s} < \infty$, $s > 0$. *If*

$$\sup_i |w_{ni}(x)| = O(n^{-s}), \tag{2.10}$$

then

$$g_n(x) \to g(x) \qquad \text{almost completely as } n \to \infty,$$

at every continuity point $x \in A$ *of the function* g.

This result extends Theorem 1 of Georgiev [19] for a general linear smoother in one dimension ($p = 1$) and Theorem 3 of Ahmad and Lin [1] for particular weights of type (2.7) and random variables $\varepsilon_1, ..., \varepsilon_n$ which are independent but not necessarily identically distributed.

If we assume that the random errors in (1.1) are bounded random variables, we obtain the next result.

THEOREM 6. *Assume the conditions of Theorem* 1, *and in addition that*

$\varepsilon_1, ..., \varepsilon_n$ are independent random variables with $E\varepsilon_i = 0$, $\sup_i |\varepsilon_i| < \infty$ almost surely. If

$$\sup_i |w_{ni}(x)| \log n \to 0 \qquad as \quad n \to \infty, \tag{2.11}$$

then

$$g_n(x) \to g(x) \qquad \text{almost completely as} \quad n \to \infty$$

at every continuity point $x \in A$ of the function g.

Our last result provides a central limit theorem for the estimate $g_n(x)$. It generalizes known results for particular weights given by Benedetti [2] and Cheng and Lin [5] for the univariate case and by Ahmad and Lin [1] for the multivariate case.

THEOREM 7. *Assume that $\varepsilon_1, ..., \varepsilon_n$ are independent random variables with $E\varepsilon_i = 0$, $\sup_i E|\varepsilon_i|^{2+t} < \infty$, for some $t > 0$. If*

$$\frac{\sum_{i=1}^n |w_{ni}(x)|^{2+t}}{[\sum_{i=1}^n w_{ni}^2(x)]^{1+t/2}} \to 0 \qquad as \quad n \to \infty, \tag{2.12}$$

then

$$\frac{g_n(x) - Eg_n(x)}{[\text{Var } g_n(x)]^{1/2}} \to N(0, 1) \qquad \text{in distribution as } n \to \infty.$$

Denote by Φ the standard normal distribution. We obtain the following results concerning the rate of convergence in the central limit theorem expressed in terms of the weight functions $w_{n1}(x), ..., w_{nn}(x)$.

COROLLARY. *Assume the conditions of Theorem 7 with $t \leqslant 1$, then*

$$\sup_{\infty < z < \infty} \left| P\left\{ \frac{g_n(x) - Eg_n(x)}{[\text{Var } g_n(x)]^{1/2}} \leqslant z \right\} - \Phi(z) \right| \leqslant C_t \frac{\sum_{i=1}^n |w_{ni}(x)|^{2+t}}{[\sum_{i=1}^n w_{ni}^2(x)]^{1+t/2}},$$

where C_t is a universal constant.

In this section we have derived various pointwise properties of the general linear smoother $g_n(x)$. The results given above extend or improve the results known in the literature.

3. Proofs

The method of some proofs used here was apparently first used for the univariate case by Georgiev [19] and later modified by Georgiev and Greblicki [22]. In spite of the fact that some ideas of the proofs used here have been used in part before, the author considers the results obtained here to be sufficiently interesting to warrant comprehensible (and therefore necessarily complete) proofs.

Proof of Theorem 1. Choose an $a>0$ and let g be continuous at $x \in A$. Note that

$$
\begin{aligned}
&|Eg_n(x)-g(x)| \\
&\quad\leqslant \sum_{i=1}^{n} |w_{ni}(x)|\,|g(x_i)-g(x)|\,(I_{\{\|x_i-x\|\leqslant a\}}+I_{\{\|x_i-x\|>a\}}) \\
&\qquad+|g(x)|\left|\sum_{i=1}^{n} w_{ni}(x)-1\right| \\
&\quad\leqslant \delta(g,a)B+2C\sum_{i=1}^{n}|w_{ni}(x)|\,I_{\{\|x_i-x\|>a\}}+C\left|\sum_{i=1}^{n} w_{ni}(x)-1\right|, \qquad (3.1)
\end{aligned}
$$

where $C=\sup_A |g(x)|$ and $\delta(g,a)=\sup_{\|x-y\|\leqslant a}|g(x)-g(y)|$. Now we may deduce from (2.1), (2.2), and (2.3) that (3.1) tends to zero if $n\to\infty$ and a is sufficiently small. ∎

Proof of Theorem 2. Since

$$|g_n(x)-g(x)|\leqslant|g_n(x)-Eg_n(x)|+|Eg_n(x)-g(x)|, \qquad (3.2)$$

we show that the random part of r.h.s. in (3.2) tends to zero in probability as $n\to\infty$. Observe that

$$|g_n(x)-Eg_n(x)|=\left|\sum_{i=1}^{n} w_{ni}(x)\varepsilon_i\right|. \qquad (3.3)$$

Now the result follows via Theorem 1 given by Rohatgi [34, p. 305], and (2.5). ∎

Proof of Theorem 3. It suffices to prove that (3.3) tends to zero in quadratic mean as $n\to\infty$. Since

$$E(g_n(x)-Eg_n(x))^2\leqslant\sigma_{\max}^2\sum_{i=1}^{n} w_{ni}^2(x),$$

where $\sigma_{\max}^2=\sup_i E\varepsilon_i^2$, the proof follows from (2.8). ∎

Proof of Theorem 4. Observe that by virtue of the law of the iterated logarithm, we have for (3.3)

$$|g_n(x) - Eg_n(x)| = O((\sup_i w_{ni}^2(x)\, n \log\log n)^{1/2})$$

with probability 1 as $n \to \infty$. Hence the proof is complete by (2.9). ∎

Proof of Theorem 5. It is easily seen that the random part in (3.2) tends to zero almost completely as $n \to \infty$ as (2.10) is in force by Theorem 2 of Rohatgi [34, p. 306]. ∎

Proof of Theorem 6. As usual we shall start with (3.2). By a slight modification of the Bernstein inequality given in Bennett [3], Georgiev and Greblicki [22] have obtained that

$$P\{|g_n(x) - Eg_n(x)| \geqslant t\} < 2\exp\left(-\frac{M}{\sup_i |w_{ni}(x)|}\right), \tag{3.4}$$

where $t, M > 0$. (3.4) is summable with respect to n when (2.11) holds. ∎

Proof of Theorem 7. Write $\mu = \sup_i E|\varepsilon_i|^{2+t}$ for some $t > 0$ and $\sigma_{\min}^2 = \inf_i E\varepsilon_i^2$. Recall that

$$\frac{g_n(x) - Eg_n(x)}{[\operatorname{Var} g_n(x)]^{1/2}} = \frac{\sum_{i=1}^n w_{ni}(x)\varepsilon_i}{[\sum_{i=1}^n \operatorname{Var}(w_{ni}(x)\varepsilon_i)]^{1/2}}. \tag{3.5}$$

By applying Liapunov's central limit theorem we see that

$$\frac{\sum_{i=1}^n E|w_{ni}(x)\varepsilon_i|^{2+t}}{[\sum_{i=1}^n \operatorname{Var}(w_{ni}(x)\varepsilon_i)]^{1+t/2}} \leqslant \frac{\mu}{\sigma_{\min}^{2+t}} \cdot \frac{\sum_{i=1}^n |w_{ni}(x)|^{2+t}}{[\sum_{i=1}^n w_{ni}^2(x)]^{1+t/2}}.$$

(3.5) converges to $N(0, 1)$ in distribution as $n \to \infty$ when (2.12) is in force. ∎

The proof of the corollary follows immediatly from the Berry–Esseen theorem and some $t \in (0, 1]$.

Acknowledgment

Final parts of this work were done during the author's stay in Lehrstuhl für Elektrische Steuerung und Regelung, Ruhr-Universität Bochum, Bochum, West Germany, as a research fellow of Alexander von Humboldt Foundation. The author is grateful to Professor Dr. Heinz Unbehauen for his kindness and hospitality.

REFERENCES

[1] AHMAD, I. A., AND LIN, P. E. (1984). Fitting a multiple regression function. *J. Statist. Plann. Inference* **9** 163–176.

[2] BENEDETTI, J. (1977). On the nonparametric estimation of regression function. *J. Roy. Statist. Soc. Ser. B* **39** 248–253.

[3] BENNETT, G. (1962). Probability inequalities for the sum of independent random variables. *J. Amer. Statist. Assoc.* **57** 33–45.

[4] CHENG, K. F., AND LIN, P. E. (1981). Nonparametric estimation of regression function. *Z. Wahrsch. Verw. Gebiete* **57** 223–233.

[5] CHENG, K. F., AND LIN, P. E. (1981). Nonparametric estimation of a regression function: limiting distribution. *Austral. J. Statist.* **23** 186–195.

[6] CLARK, R. M. (1977). Non-parametric estimation of a smooth regression function. *J. Roy. Statist. Soc. Ser. B* **39** 107–113.

[7] COLLOMB, G. (1981). Estimation non paramétrique de la régression: Revue bibliographique. *Internat. Statist. Rev.* **49** 75–93.

[8] COLLOMB, G. (1985). Nonparametric regression: An up-to-date bibliography. *Statistics* **16** 309–324.

[9] DEVROYE, L., AND WAGNER, T. (1980). Distribution-free consistency results in nonparametric discrimination and regression function estimation. *Ann. Statist.* **8** 231–239.

[10] FREEDMAN, D. A. (1981). Bootstrapping regression models. *Ann. Statist.* **9** 1218–1228.

[11] GAŁKOWSKI, T., AND RUTKOWSKI, L. (1985). Nonparametric recovery of multivariate functions with applications to system identification. *Proc. IEEE* **73** 942–943.

[12] GAŁKOWSKI, T., AND RUTKOWSKI, L. (1986). Nonparametric fitting of multivariate functions. *IEEE Trans. Automat. Control* **AC-31** 785–787.

[13] GASSER, T., AND MÜLLER, H. G. (1979). Kernel estimation of regression functions. In *Smoothing Techniques for Curve Estimation* (T. Gasser and M. Rosenblatt, Eds.), Lecture Notes in Math. Vol. 757, pp. 23–68. Springer-Verlag, Berlin.

[14] GASSER, T., MÜLLER, H. G. KÖHLER, W., MOLINARI, L., AND PRADER, A. (1984). Nonparametric regression analysis of growth curves. *Ann. Statist.* **12** 210–229.

[15] GEORGIEV, A. A. (1984). Kernel estimates of functions and their derivatives with applications. *Statist. Probab. Lett.* **2** 45–50.

[16] GEORGIEV, A. A. (1984). On the estimation of multiple regression function and its mixed partial derivatives. Unpublished.

[17] GEORGIEV, A. A. (1984). Speed of convergence in nonparametric kernel estimation of a regression function and its derivatives. *Ann. Inst. Statist. Math.* **36** 455–462.

[18] GEORGIEV, A. A. (1984). On the recovery of functions and their derivatives from imperfect measurements. *IEEE Trans. Systems Man Cyberbet.* **SMC-14** 900–903.

[19] GEORGIEV, A. A. (1985).Local properties of function fitting estimates with application to system identification. In *Mathematical Statistics and Applications, Proceedings, 4th Pannonian Symp. Math. Statist., September 4–10, 1983, Bad Tatzmannsdorf, Austria* (W. Grossmann *et al.*, Eds.), pp. 141–151. Reidel, Dordrecht.

[20] GEORGIEV, A. A. (1985). Propriétés asymptotiques d'un estimateur fonctionnel non paramétrique. *C. R. Acad. Sci. Paris Sér. I Math.* **300** 407–410.

[21] GEORGIEV, A. A. (1985). *Consistent Nonparametric Multiple Regression: The Fixed Design Case.* Report 78/I-6/85. Technical University of Wrocław, Wrocław.

[22] GEORGIEV, A. A., AND GREBLICKI, W. (1986). Nonparametric function recovering from noisy observations. *J. Statist. Plann. Inference* **13** 1–14.

[23] GREBLICKI, W. (1982). *Nearest Neighbor Algorithm for Recovering Function from Noise.* Report 30/I-6/82. Technical University of Wrocław, Wrocław.

[24] Greblicki, W., Krzyżak, A., and Pawlak, M. (1984). Distribution-free pointwise consistency of kernel regression estimate. *Ann. Statist.* **12** 1570–1575.

[25] Härdle, W., and Luckhaus, S. (1984). Uniform consistency of a class of regression estimators. *Ann. Statist.* **12** 612–623.

[26] Härdle, W., and Gasser, T. (1984). Robust nonparametric function fitting. *J. Roy. Statist. Soc. Ser. B* **46** 42–51.

[27] Hsu, P. L., and Robbins, H. (1947). Complete convergence and the law of large numbers. *Proc. Nat. Acad. Sci. U.S.A.* **33** 25–31.

[28] Müller, H. G. (1984). Smooth optimum kernel estimates of densities, regression curves and modes. *Ann. Statist.* **12** 766–774.

[29] Prakasa Rao, B. L. S. (1983). *Nonparametric Functional Estimation.* Academic Press, New York.

[30] Priestley, M. B., and Chao, M. T. (1972). Non-parametric function fitting. *J. Roy. Statist. Soc. Ser. B* **34** 385–392.

[31] Rafajłowicz, E. (1987). Nonparametric orthogonal series estimators of regression: A class attaining the optimal convergence rate in L_2. *Statist. Probab. Lett.* **5** 219–224.

[32] Rafajłowicz, E. (1986). Nonparametric least squares estimation of a regression function. *Statistics*, in press.

[33] Rice, J., and Rosenblatt, M. (1983). Smoothing splines: Regression, derivatives and deconvolution. *Ann. Statist.* **11** 141–156.

[34] Rohatgi, V. K. (1971). Convergence of weighted sums of independent random variables. *Proc. Cambridge Philos. Soc.* **69** 305–307.

[35] Rutkowski, L. (1982). On system identification by nonparametric function fitting. *IEEE Trans. Automat. Control* **AC-27** 225–227.

[36] Schuster, E. F., and Yakowitz, S. J. (1979). Contribution to the theory of nonparametric regression with application to system identification. *Ann. Statist.* **7** 139–149.

[37] Silverman, B. W. (1984). Spline smoothing: The equivalent variable kernel method. *Ann. Statist.* **12** 898–916.

[38] Spiegelman, C., and Sacks, J. (1980). Consistent window estimation in nonparametric regression. *Ann. Statist.* **8** 240–246.

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

[40] Wahba, G. (1975). Smoothing noisy data with spline functions. *Numer. Math.* **24** 383–393.

[41] Yakowitz, S. J., and Szidarovszky, F. (1985). A comparison of kriging with nonparametric regression methods. *J. Multivariate Anal.* **16** 21–53.