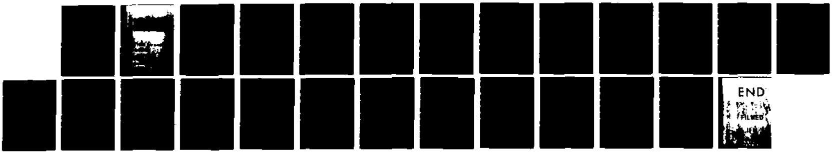
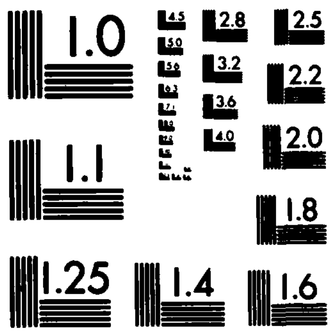
AD-A133 482 FITTING A MULTIPLE REGRESSION FUNCTION(U) FLORIDA STATE 1/1
UNIV TALLAHASSEE DEPT OF STATISTICS I A AHMAD ET AL.
DEC 82 FSU-STATISTICS-M650 N00014-80-C-0093

UNCLASSIFIED F/G 12/1 NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

FITTING A MULTIPLE REGRESSION FUNCTION

by

Ibrahim A. Ahmad[1] and Pi-Erh Lin[2]

AD A133482

DTIC FILE COPY

The Florida State University

Department

of

Statistics

Tallahassee, Florida

32306

FITTING A MULTIPLE REGRESSION FUNCTION

by

Ibrahim A. Ahmad[1] and Pi-Erh Lin[2]

FSU Statistics Report No. M650
ONR Technical Report No. 160

December, 1982

The Florida State University
Department of Statistics
Tallahassee, Florida 32306

[1]At University of Maryland Baltimore County

[2]This work was supported by the Army, Navy and Air Force under Office of Naval Research Contract No. N00014-80-C-0093. Reproduction in whole or in part is permitted for any purpose of the United States Government.

Key words: Function regression, consistency, asymptotic normality, optimal kernel, rates of convergence, kernel function.

AMS (1970) Subject classifications: Primary 62G05; Secondary 60F05, 60F15.

FITTING A MULTIPLE REGRESSION FUNCTION

by

Ibrahim A. Ahmad and Pi-Erh Lin
University of Maryland Baltimore County
and
Florida State University

ABSTRACT

Consider the p-dimensional unit cube $[0,1]^p$, $p \geq 1$. Partition $[0,1]^p$ into n regions, $R_{1,n},\ldots,R_{n,n}$ such that the volume $\Delta(R_{j,n})$ is of order n^{-1}, $j = 1,\ldots,n$. Select and fix a point in each of these regions so that we have $\underset{\sim}{x}_1^{(n)},\ldots,\underset{\sim}{x}_n^{(n)}$. Suppose that associated with the j-th predictor vector $\underset{\sim}{x}_j^{(n)}$ there is an observable variable $Y_j^{(n)}$, $j = 1,\ldots,n$ satisfying the multiple regression model $Y_j^{(n)} = g(\underset{\sim}{x}_j^{(n)}) + e_j^{(n)}$, where g is an unknown function defined on $[0,1]^p$ and $\{e_j^{(n)}\}$ are independent identically distributed random variables with $Ee_1^{(n)} = 0$ and $\mathrm{Var}\, e_1^{(n)} = \sigma^2 < \infty$. This paper proposes $g_n(\underset{\sim}{x}) = a_n^{-p}\sum_{j=1}^n Y_j^{(n)}\int_{R_{j,n}} k[(\underset{\sim}{x}-\underset{\sim}{u})/a_n]d\underset{\sim}{u}$ as an estimator of $g(\underset{\sim}{x})$, where $k(\underset{\sim}{u})$ is a known p-dimensional bounded density and $\{a_n\}$ is a sequence of reals converging to 0 as $n \to \infty$. Weak and strong consistency of $g_n(\underset{\sim}{x})$ and rates of convergence are obtained. Asymptotic normality of the estimator is established. Also proposed is $\sigma_n^2 = n^{-1}\sum_{j=1}^n (Y_j^{(n)} - g_n(\underset{\sim}{x}_j^{(n)}))^2$ as a consistent estimate of σ^2.

1. INTRODUCTION.

A statistical problem which finds a wide range of applications is the estimation of a regression function $g(\underset{\sim}{x}) = E(Y|\underset{\sim}{x})$, where Y is a dependent variable and $\underset{\sim}{x}$ is a $p\times 1$ vector of regressors ($p \geq 1$). If $g(\underset{\sim}{x})$ is specified except for a set of parameters, then a typical estimate for $g(\underset{\sim}{x})$ would be the least squares estimate which is, of course, the maximum likelihood estimate if the errors are normally distributed and $g(\underset{\sim}{x})$ is linear. But if $g(\underset{\sim}{x})$ is completely unknown, it would be desirable to estimate $g(\underset{\sim}{x})$ by a method that would provide good properties of the estimate. When x is univariate (i.e., $p=1$), such a method is proposed and studied by Priestley and Chao (1972). Their estimate has been further studied by Benedetti (1977), Cheng and Lin (1981a,b), and Schuster and Yakowitz (1979), among others. The Priestley-Chao estimate is nonparametric in the sense that the *conditional* distribution of Y given x is not specified. This estimate resembles the kernel estimate of a probability density function investigated by Rosenblatt (1956), Parzen (1962), and many others.

This document
In ~~the present investigation~~, an estimate of $g(\underset{\sim}{x})$ is proposed when there are at least two independent regressors. This is not a direct generalization of the Priestley-Chao estimate. Also presented is a consistent estimate for the error variance. With the aid of the variance estimate, an asymptotic confidence interval ~~for $g(\underset{\sim}{x})$~~ can be constructed.

The multiple regression function model we discuss here may be presented as follows: Let $[0,1]^p$ denote the p-dimensional unit cube $(p \geq 1)$. Divide the unit cube into n mutually disjoint and totally exhaustive regions $R_{1,n}, \ldots, R_{n,n}$ such that the volume of $R_{j,n}$ converges to 0 as $n \to \infty$. From each of these regions select and fix a point so that we have $\underset{\sim}{x}_1^{(n)}, \ldots, \underset{\sim}{x}_n^{(n)}$ where $\underset{\sim}{x}_j^{(n)} \in R_{j,n}$, $j = 1, \ldots n$. Suppose that $Y_1^{(n)}, \ldots, Y_n^{(n)}$ is a random sample obtained from the following model:

$$Y_j^{(n)} = g(\underset{\sim}{x}_j^{(n)}) + e_i^{(n)}, \; j = 1, \ldots, n, \tag{1.1}$$

where $e_1^{(n)}, \ldots, e_n^{(n)}$ are independent identically distributed (iid) random variables such that $Ee_1^{(n)} = 0$ and $\mathrm{Var} e_1^{(n)} = \sigma^2 < \infty$ and $g(\cdot)$ is an unknown p-dimensional function defined on $[0,1]^p$. The problem is to estimate $g(\underset{\sim}{x})$. Let $\Delta(R_{j,n})$ denote the volume of the j-th region $R_{j,n}$, $j = 1, \ldots, n$. A nonparametric estimate of $g(\underset{\sim}{x})$ is defined by:

$$g_n(\underset{\sim}{x}) = a_n^{-p} \sum_{j=1}^n Y_j^{(n)} \int_{R_{j,n}} k[(\underset{\sim}{x} - \underset{\sim}{u})/a_n] d\underset{\sim}{u}, \tag{1.2}$$

where $k(\underset{\sim}{u})$ is a known p-dimensional probability density function satisfying the following conditions:

(i) $\sup_{\underset{\sim}{u}} k(\underset{\sim}{u}) < \infty$, (ii) $\lim_{||\underset{\sim}{u}|| \to \infty} ||\underset{\sim}{u}|| k(\underset{\sim}{u}) = 0$,

where $|| \cdot ||$ is the Euclidean distance function.

If an approximate confidence interval for $g(\underset{\sim}{x})$ is desired, then one would need a consistent estimate of σ^2. One such estimate may be given by:

$$\sigma_n^2 = n^{-1} \sum_{j=1}^n (Y_j^{(n)} - g_n(\underset{\sim}{x}_j^{(n)}))^2. \tag{1.3}$$

The organization of this paper is as follows: In Section 2, weak and strong consistency and their rates, and asymptotic normality of $g_n(\underset{\sim}{x})$ are established. The conditions required to prove the consistency of $g_n(\underset{\sim}{x})$ are much weaker than those of Priestley and Chao (1972), Schuster and Yakowitz (1979), and Benedetti (1977); and the methods of proof are different. In Section 3, the (weak) consistency of σ_n^2 is established. Thus an approximate normal confidence interval for $g(\underset{\sim}{x})$ can be established. Finally in Section 4 we discuss the optimal choice of $k(\underset{\sim}{u})$.

In the rest of this paper we employ the following notations: $\xrightarrow{P}$, $\xrightarrow{wp1}$, and $\xrightarrow{D}$ to mean, respectively, convergence in probability, with probability one, and in distribution. Unless otherwise specified, hereafter, all integral signs will mean multiple integration. For easy of exposition, we shall write R_j for $R_{j,n}$ and suppress all superscripts for $Y_j^{(n)}$, $\underset{\sim}{x}_j^{(n)}$, and $e_j^{(n)}$, $j = 1, \ldots, n$ in the remainder of the paper.

2. PROPERTIES OF $g_n(\underset{\sim}{x})$.

In this section some basic properties of $g_n(\underset{\sim}{x})$ are established. Precisely we show that $g_n(\underset{\sim}{x})$ is asymptotically unbiased (Theorem 1), weakly consistent (Theorem 2), strongly consistent (Theorem 3), and asymptotically normal (Theorem 6). We also demonstrate that the rate of weak consistency is of the order $0(n^{-\rho})$ for some $\rho > 0$.

THEOREM 1. *If* $\max_{1\le j\le n} \Delta(R_j) = 0(n^{-1})$, *if* $na_n^p \to \infty$ *as* $n \to \infty$, *and if* $g(\underset{\sim}{x})$ *is continuous on* $[0,1]^p$, *then for each* $\underset{\sim}{x} \in [0,1]^p$,

$$Eg_n(\underset{\sim}{x}) \to g(\underset{\sim}{x}) \quad \text{as} \quad n \to \infty. \tag{2.1}$$

PROOF. Note that

$$Eg_n(\underset{\sim}{x}) = a_n^{-p}\sum_{j=1}^{n} g(\underset{\sim}{x}_j)\int_{R_j} k[(\underset{\sim}{x}-\underset{\sim}{u})/a_n]d\underset{\sim}{u},$$

where $d\underset{\sim}{u} = du \ldots du_p$. Thus,

$$|Eg_n(\underset{\sim}{x})-g(\underset{\sim}{x})| \le |\sum_{j=1}^{n}\int_{R_j}[g(\underset{\sim}{x}_j)-g(\underset{\sim}{u})]a_n^{-p}k[(\underset{\sim}{x}-\underset{\sim}{u})/a_n]d\underset{\sim}{u}|$$

$$+ |\sum_{j=1}^{n}\int_{R_j}[g(\underset{\sim}{u})-g(\underset{\sim}{x})]a_n^{-p}k[(\underset{\sim}{x}-\underset{\sim}{u})/a_n]d\underset{\sim}{u}|$$

$$= I_{1n} + I_{2n}, \text{ say.} \qquad (2.2)$$

But since $g(\underset{\sim}{x})$ is uniformly continuous on $[0,1]^p$ and $\max_{1\le j\le n} \Delta(R_j) = O(n^{-1})$ then for sufficiently large n, I_{1n} can be made arbitrary small. Note also that as $n \to \infty$

$$I_{2n} = |\int_{[0,1]^p}[g(\underset{\sim}{u})-g(\underset{\sim}{x})]a_n^{-p}k[(\underset{\sim}{x}-\underset{\sim}{u})/a_n]d\underset{\sim}{u}| \to 0, \qquad (2.3)$$

by Lemma 2.1 of Cacoullos (1966) provided $na_n^p \to \infty$, as $n \to \infty$. ||

THEOREM 2. If the conditions of Theorem 1 hold then for all $\underset{\sim}{x} \in [0,1]^p$,

$$g_n(\underset{\sim}{x}) \xrightarrow{P} g(\underset{\sim}{x}) \text{ as } n \to \infty. \qquad (2.4)$$

PROOF. In light of Theorem 1 we need only to show that for all $\underset{\sim}{x} \in [0,1]^p$, $g_n(\underset{\sim}{x}) - Eg_n(\underset{\sim}{x}) \xrightarrow{P} 0$ as $n \to \infty$. To this end we use the following result of Pruitt (1966, Theorem 1): Let $\{U_j\}$ be a sequence of iid random variables such that $EU_1 = 0$, and let $Z_n = \sum_{j=1}^{n} C_{nj}U_j$ where $\{C_{nj}\}$ is an array of constants such that $\lim_{n\to\infty} C_{nj} = 0$ for every integer j. Then $Z_n \xrightarrow{P} 0$ if and only if $\max_{1\le j\le n} C_{nj} \to 0$ as $n \to \infty$. Now set

$$U_j = Y_j - EY_j \quad \text{and} \quad C_{nj} = a_n^{-p} \int_{R_j} k[(\underset{\sim}{x}-\underset{\sim}{u})/a_n]d\underset{\sim}{u}. \tag{2.5}$$

Then there exists a positive constant C not depending on n such that

$$\max_{1\le j\le n} C_{nj} \le a_n^{-p} \max_{1\le j\le n} \Delta(R_j) \sup_{\underset{\sim}{v}} k(\underset{\sim}{v}) \le C(na_n^p)^{-1} \to 0, \quad \text{as} \quad n \to \infty.$$

Thus the result follows. ||

<u>THEOREM 3</u>. <u>Assume that the conditions of Theorem 1 are in force and that</u> $a_n \ge Cn^{-(1-\beta)/p}$ <u>for some</u> $C > 0$ <u>and</u> $\beta \in (0,1)$. <u>If</u> $E|e_1|^{1+1/\beta} < \infty$, <u>then for all</u> $\underset{\sim}{x} \in [0,1]^p$,

$$g_n(\underset{\sim}{x}) \xrightarrow{wp1} g(\underset{\sim}{x}), \quad \underline{\text{as}} \quad n \to \infty. \tag{2.6}$$

<u>PROOF</u>. We use another result of Pruit (1966, Theorem 2), in which it is stated that if $\max_{1\le j\le n} C_{nj} = O(n^{-\beta})$ for some $0 < \beta < 1$ and if $E|U_1|^{1+1/\beta} < \infty$, then $Z_n \xrightarrow{wp1} 0$, as $n \to \infty$, where C_{nj}, U_j, and Z_n are as defined in the proof of our Theorem 2. Thus $\max_{1\le j\le n} C_{nj} \le C/(na_n^p) = O(n^{-\beta})$ if we choose $a_n \ge Cn^{-(1-\beta)/p}$, and the desire conclusion follows. ||

It is possible to obtain the rates of convergence in weak consistency (Theorem 2). First we state and prove the following lemma.

<u>LEMMA 1</u>. <u>Suppose that there exist positive constants</u> C_1 <u>and</u> C_2 <u>such that, for all</u> $n \ge 1$,

$$a_n \geq \max(C_1 n^{-\alpha/p},\ C_2 n^{-(1-\beta)/p}) \tag{2.7}$$

for some $0 < \alpha < \beta < 1$. If $G(u) = P(|Y_1 - EY_1| \geq u)$, and if $u^t G(u) \leq M < \infty$ for some $t > 1 + \alpha/\beta$, then for every $\varepsilon > 0$,

$$P(|g_n(\underset{\sim}{x}) - Eg_n(\underset{\sim}{x})| \geq \varepsilon) = O(n^{-\rho}), \tag{2.8}$$

where $\rho = \beta(t-1) - \alpha > 0$.

PROOF. We make use of Theorem 1 of Franck and Hanson (1966). Recall the definitions of $\{C_{nj}\}$, $\{U_j\}$ and Z_n from the proof of Theorem 2. Franck and Hanson (1966) show that if for some constants $0 < \alpha < \beta$, $\sum_{j=1}^n C_{nj} \leq C_3 n^{\alpha}$, $\max\limits_{1\leq j\leq n} C_{nj} \leq C_4 n^{-\beta}$, and if for some $t > 1$, $\sum_{j=1}^n C_{nj}^t \leq C_5 n^{-\rho}$ for some $\rho > 0$ then it follows that $P[|Z_n| \geq \varepsilon] = O(n^{-\rho})$. Identifying $C_{nj} = a_n^{-p} \int_{R_j} k[(\underset{\sim}{x}-\underset{\sim}{u})/a_n]d\underset{\sim}{u}$ and $U_j = Y_j - EY_j$, $j = 1,\ldots,n$. we see that if $a_n \geq C_2 n^{-(1-\beta)/p}$ then $\max\limits_{1\leq j\leq n} C_{nj} \leq C_4 n^{-\beta}$ and if $a_n \geq C_1 n^{-\alpha/p}$, then

$$\sum_{j=1}^n C_{nj} = a_n^{-p} \int_{[0,1]^p} k[(\underset{\sim}{x}-\underset{\sim}{u})/a_n]d\underset{\sim}{u} \leq C^* a_n^{-p} \leq C_3 n^{\alpha} .$$

Finally, $\sum_{j=1}^n C_{nj}^t \leq \max\limits_{1\leq j\leq n} C_{nj}^{t-1} \sum_{j=1}^n C_{nj}$

$$\leq C_5 n^{-\beta(t-1)+\alpha} \leq C_5 n^{-\rho},$$

where $\rho = \beta(t-1) - \alpha.$ ||

THEOREM 4. Assume that the conditons of Lemma 1 are in force. Then for any $\varepsilon > 0$ and all $\underset{\sim}{x} \in [0,1]^p$,

$$P[|g_n(\underset{\sim}{x}) - g(\underset{\sim}{x})| \geq \varepsilon] = O(n^{-\rho}). \tag{2.9}$$

PROOF. Since $Eg_n(\underset{\sim}{x}) - g(\underset{\sim}{x}) \to 0$ as $n \to \infty$ for all $\underset{\sim}{x} \in [0,1]^p$ we have for n sufficiently large that $|Eg_n(\underset{\sim}{x}) - g(\underset{\sim}{x})| \le \varepsilon/2$ and hence

$$P[|g_n(\underset{\sim}{x}) - g(\underset{\sim}{x})| \ge \varepsilon] \le P[|g_n(\underset{\sim}{x}) - Eg_n(\underset{\sim}{x})| \ge \varepsilon/2] = 0(n^{-\rho}), \tag{2.10}$$

by Lemma 1.||

We can also establish rates of convergence in the mean square consistency, i.e., the rate of $E[g_n(\underset{\sim}{x}) - g(\underset{\sim}{x})]^2$. To this end we establish the following lemma

LEMMA 2. Suppose that $k(u)$ is such that $\int u_{i_1} \dots u_{i_j} k(\underset{\sim}{u})d\underset{\sim}{u} = 0$ for all $i_1,\dots,i_j = 1,2,\dots,p$ and $j = 1,\dots,M-1$ and $\int |u_{i_1}| \dots |u_{i_M}| k(\underset{\sim}{u})d\underset{\sim}{u} < \infty$ for all $i_1,\dots,i_M = 1,2,\dots,p$. Assume also that all partial derivatives of order M or less of $g(\underset{\sim}{x})$ exist and are bounded. Then for any $\underset{\sim}{x} \in [0,1]^p$,

$$[Eg_n(\underset{\sim}{x}) - g(\underset{\sim}{x})]^2 = 0(a_n^{2Mp}). \tag{2.11}$$

REMARK. Under the present setting Lemma 2 holds only for $M \le 2$. In order for the lemma to hold for $M > 2$, however, the kernel function can no longer be a probability density function. It must be allowed to take both positive and negative values. Then Conditions (i) and (ii) for k(.) given in Section 1 must be appropriately modified. This phenomenon is also noted by Cacoullos (1966) in th' 'rne estimate of a multivariate density function.

PROOF OF LEMMA 2. Recall that

$$|Eg_n(\underset{\sim}{x}) - g(\underset{\sim}{x})| \le |\sum_{j=1}^{n} \int_{R_j} [g(\underset{\sim}{x}_j) - g(\underset{\sim}{u})] a_n^{-p} k[(\underset{\sim}{x}-\underset{\sim}{u})/a_n] d\underset{\sim}{u}|$$

$$+ |\int [g(\underset{\sim}{u}) - g(\underset{\sim}{x})] a_n^{-p} k[(\underset{\sim}{x}-\underset{\sim}{u})/a_n] d\underset{\sim}{u}| = I_{1n} + I_{2n},$$

say $\qquad (2.12)$

We shall show that $I_{in} = 0(a_n^{Mp})$, $i = 1,2$, $\underset{\sim}{x} \in [0,1]^p$. We shall prove that $I_{2n} = 0(a_n^{Mp})$, and note that $I_{1n} = 0(a_n^{Mp})$ may be shown analogously. Now, writing $\underset{\sim}{x}' = (x_1,\ldots,x_p)$,

$$I_{2n} = |\int_{C*}[g(\underset{\sim}{x}-a_n\underset{\sim}{w}) - g(\underset{\sim}{x})]k(\underset{\sim}{w})d\underset{\sim}{w}|, \tag{2.13}$$

where $C^* = [-\frac{1-x_1}{a_n}, \frac{x_1}{a_n}] \times \ldots \times [-\frac{1-x_p}{a_n}, \frac{x_p}{a_n}]$. Now, using multidimensional Taylor expansion we see that

$$g(\underset{\sim}{x}-a_n\underset{\sim}{w}) - g(\underset{\sim}{x}) = \sum_{\ell=1}^{M-1} \frac{1}{\ell!} g^{(\ell)}(\underset{\sim}{x};-a_n\underset{\sim}{w}) + \frac{1}{M!} g^{(M)}(\underset{\sim}{x}-\theta a_n\underset{\sim}{w};-a_n\underset{\sim}{w}), \tag{2.14}$$

where $g^{(\ell)}(\underset{\sim}{x};\underset{\sim}{t}) = \sum_{i_1=1}^{p} \cdots \sum_{i_\ell=1}^{p} \frac{\partial^\ell g(\underset{\sim}{x})}{\partial x_{i_1}\ldots\partial x_{i_\ell}} t_{i_1} \ldots t_{i_\ell}$, $\ell = 1,\ldots,M$.

Hence

$$I_{2n} = |\frac{1}{M!} \int k(\underset{\sim}{w}) g^{(M)}(\underset{\sim}{x}-\theta a_n\underset{\sim}{w};-a_n\underset{\sim}{w})d\underset{\sim}{w}|$$

$$\leq \frac{a_n^{Mp}}{M!} \int k(\underset{\sim}{w}) |D(i_1,\ldots,i_M;g)(|\underset{\sim}{x}-\theta a_n\underset{\sim}{w}|)||w_{i_1}| \ldots |w_{i_M}|(-1)^M d\underset{\sim}{w}, \tag{2.15}$$

where $D(i_1,\ldots,i_M;g)(\underset{\sim}{x}) = \frac{\partial^M g(\underset{\sim}{x})}{\partial x_{i_1}\ldots\partial x_{i_M}}$ and

$|\underset{\sim}{x} - \theta a_n\underset{\sim}{w}| = (|x_1-\theta a_n w_1|,\ldots,|x_p-\theta a_n w_p|)$. But

$|D(i_1,\ldots,i_M;g)(\underset{\sim}{x})| \leq M$ for all $\underset{\sim}{x}$; it follows that

$$I_{2n} \leq C^* a_n^{Mp}. \tag{2.16}$$

Thus the lemma follows.||

<u>THEOREM 5</u>. <u>Assume that the conditions of Lemma 2 are satisfied. Then for all</u> $\underset{\sim}{x} \in [0,1]^p$ <u>and</u> $a_n = 0(n^{-1/(2M+1)p})$,

$$E[g_n(\underset{\sim}{x}) - g(\underset{\sim}{x})]^2 = 0(n^{-2M/(2M+1)}). \tag{2.17}$$

PROOF. Note that $E[g_n(\underset{\sim}{x}) - g(\underset{\sim}{x})]^2 \le \mathrm{Var} g_n(\underset{\sim}{x}) + [Eg_n(\underset{\sim}{x}) - g(\underset{\sim}{x})]^2$.

Now from Lemma 2, $[Eg_n(\underset{\sim}{x}) - g(\underset{\sim}{x})]^2 = 0(n^{-2M/2(M+1)})$. Thus we need to evaluate $\mathrm{Var} g_n(\underset{\sim}{x})$

$$\begin{aligned}\mathrm{Var} g_n(\underset{\sim}{x}) &= \sigma^2 a_n^{-2p} \textstyle\sum_{j=1}^n \{\int_{R_j} k[(\underset{\sim}{x}-\underset{\sim}{u})/a_n] d\underset{\sim}{u}\}^2 \\ &\le (\sigma^2 a_n^{-2p} \max_{1\le j\le n} \Delta(R_j) \int_{[0,1]^p} k[(\underset{\sim}{x}-\underset{\sim}{u})/a_n] d\underset{\sim}{u} \\ &\le C^* a_n^{-p} n^{-1} = 0((na_n^p)^{-1}) = 0(n^{-2M/(2M+1)}),\end{aligned} \qquad (2.18)$$

by choosing $a_n = 0(n^{-1/(2M+1)p})$. ||

As for the asymptotic normality of $g_n(\underset{\sim}{x})$ we proceed as follows: Choose the regions $R_1, \ldots, R_n$ so that $\Delta(R_j) = c_j/n$, $j = 1, \ldots, n$ where $c_1, \ldots, c_n$ are positive constants with $\sum_{j=1}^n c_j = n$ and that $\sup_{\underset{\sim}{u},\underset{\sim}{v} \in R_j} \|\underset{\sim}{u} - \underset{\sim}{v}\| = O(n^{-1/p})$. Write $c_{min} = \min\{c_j\}$ and $c_{max} = \max\{c_j\}$. Assume that $\nu_3 = E|e_1|^3 < \infty$. Then, for large n,

$$\begin{aligned}&(na_n^p)^{3/2} \textstyle\sum_{j=1}^n E|a_n^{-p}(Y_j - EY_j) \int_{R_j} k[(\underset{\sim}{x}-\underset{\sim}{u})/a_n] d\underset{\sim}{u}|^3 \\ &= (na_n^{-p})^{3/2} \textstyle\sum_{j=1}^n E|Y_j - EY_j|^3 \{\int_{R_j} k[(\underset{\sim}{x}-\underset{\sim}{u})/a_n] d\underset{\sim}{u}\}^3 \\ &= \nu_3 (na_n^{-p})^{3/2} \textstyle\sum_{j=1}^n \{\int_{R_j} k[\ \underset{\sim}{x}-\underset{\sim}{u})/a_n] d\underset{\sim}{u}\}^3 \\ &\le C\nu_3 (na_n^{-p})^{3/2} (c_{max}/n)^2 \int_{[0,1]^p} k[(\underset{\sim}{x}-\underset{\sim}{u})/a_n] d\underset{\sim}{u} \\ &\le C' \nu_3 (na_n^p)^{-1/2},\end{aligned}$$

provided that k(.) is bounded. Similarly, for n sufficiently large, and k(.) is of Lipschitz of order β,

$$na_n^p \operatorname{Var} g_n(\underset{\sim}{x}) = na_n^{-p}\sum_{j=1}^n E(Y_j-EY_j)^2\{\int_{R_j} k[(\underset{\sim}{x}-\underset{\sim}{u})/a_n]d\underset{\sim}{u}\}^2$$

$$= \sigma^2(na_n^{-p})\sum_{j=1}^n\{\int_{R_j} k[(\underset{\sim}{x}-\underset{\sim}{u})/a_n]d\underset{\sim}{u}\}^2$$

$$\approx \sigma^2(na_n^{-p})\sum_{j=1}^n \Delta(R_j)k[(\underset{\sim}{x}-\underset{\sim}{x}_j)/a_n]\int_{R_j} k[(\underset{\sim}{x}-\underset{\sim}{u})/a_n]d\underset{\sim}{u}$$

$$\approx \sigma^2\sum_{j=1}^n c_j a_n^{-p}\int_{R_j} k^2[(\underset{\sim}{x}-\underset{\sim}{u})/a_n]d\underset{\sim}{u}$$

$$\geq \sigma^2 c_{min}\int_{R^p} k^2(\underset{\sim}{u})d\underset{\sim}{u}. \qquad (2.19)$$

In fact, when $\Delta(R_j) = 1/n$ for all j, $(na_n^p)\operatorname{Var} g_n(\underset{\sim}{x}) \to \sigma^2\int k^2(\underset{\sim}{u})d\underset{\sim}{u}$. Hence applying Liapounov's central limit theorem (Loéve (1963), p.277), we see that

$$(na_n^p)^{1/2}[g_n(\underset{\sim}{x}) - Eg_n(\underset{\sim}{x})]/\{\sigma^2\sum_{j=1}^n c_j a_n^{-p}\int_{R_j} k^2[(\underset{\sim}{x}-\underset{\sim}{u})/a_n]d\underset{\sim}{u}\}^{1/2}$$

converges to N(0,1) in distribution as n tends to infinity and that from Lemma 2 with M = 2

$$(na_n^p)^{1/2}[Eg_n(\underset{\sim}{x}) - g(\underset{\sim}{x})] = O((na_n^{5p})^{1/2}). \qquad (2.20)$$

Hence we arrive at the following theorem.

<u>THEOREM 6</u>. <u>Assume</u> <u>that</u> $E|e_1|^3 < \infty$, $g(\underset{\sim}{x})$ <u>has</u> <u>bounded</u> <u>second</u> <u>partial</u> <u>derivatives</u>. <u>If</u> k <u>is</u> Lip(β), $\int u_i k(\underset{\sim}{u})d\underset{\sim}{u} = 0$, <u>and</u> $\int |u_i u_j| k(\underset{\sim}{u})d\underset{\sim}{u} < \infty$, <u>for</u> <u>all</u> i,j = 1,...,p, <u>if</u> $na_n^p \to \infty$ <u>and</u> <u>if</u> $na_n^{5p} \to 0$ <u>as</u> $n \to \infty$, <u>then</u>

$$(na_n^p)^{1/2}[g_n(\underset{\sim}{x}) - g(\underset{\sim}{x})]/[\sigma^2\sum_{j=1}^n c_j V_j^2]^{1/2} \xrightarrow{D} N(0,1), \text{ as } n \to \infty,$$

where

$$V_j^2 = a_n^{-p}\int_{R_j} k^2[(\underset{\sim}{x}-\underset{\sim}{u})/a_n]d\underset{\sim}{u}. \qquad (2.21)$$

Note that when $p = 1$, the estimate $g_n(\underset{\sim}{x})$ is a competitor to that of Priestley and Chao (1972). The properties of our estimator held under much weaker conditions than those of Priestley and Chao (1972) and Benedetti (1977). To make this remark more precise let us define a multidimensional extension of the Priestley and Chao estimate and discuss briefly its properties. Let

$$\tilde{g}_n(\underset{\sim}{x}) = a_n^{-p}\sum_{j=1}^{n} Y_j \Delta(R_j) k[(\underset{\sim}{x}-\underset{\sim}{x}_j)/a_n]. \tag{2.22}$$

Note that if $p = 1$ and we select $R_j = [x_{j-1}, x_j]$, $j=1,\ldots,n$ where $0 = x_0 < x_1 < \ldots < x_n = 1$, (2.22) reduces to the estimate proposed by Priestley and Chao (1972). They prove that if $g(x)$ and $k(u)$ are both Lipschitz of orders α and β respectively, if $\max\limits_{1\leq j\leq n}\Delta(R_j) = 0(n^{-1})$, and if $a_n = n^{-\gamma}$, $0 < \gamma < \min(\alpha, \frac{\beta}{1+\beta})$, then $g_n(x) \xrightarrow{P} g(x)$ for all $x \in [0,1]$, provided that $g(x)$ is continuous on $[0,1]$. A better result can be obtained for $\tilde{g}_n(\underset{\sim}{x})$, $p \geq 1$ as follows: Consider

$$\begin{aligned} E(g_n(\underset{\sim}{x}) - \tilde{g}_n(\underset{\sim}{x}))^2 &= E\{a_n^{-p}\sum_{j=1}^{n} Y_j \int_{R_j} [k[(\underset{\sim}{x}-\underset{\sim}{x}_j)/a_n] - k[(\underset{\sim}{x}-\underset{\sim}{u})/a_n]]d\underset{\sim}{u}\}^2 \\ &= \sigma^2\sum_{j=1}^{n}\{a_n^{-p}\int_{R_j}[k[(\underset{\sim}{x}-\underset{\sim}{x}_j)/a_n] - k[(\underset{\sim}{x}-\underset{\sim}{u})/a_n]]d\underset{\sim}{u}\}^2 \\ &+ \{a_n^{-p}\sum_{j=1}^{n} g(\underset{\sim}{x}_j)\int_{R_j}[k[(\underset{\sim}{x}-\underset{\sim}{x}_j)/a_n] - k(\underset{\sim}{x}-\underset{\sim}{u})/a_n]]d\underset{\sim}{u}\}^2 \end{aligned} \tag{2.23}$$

It is not difficult to see that if $k(\underset{\sim}{u})$ is Lipschitz of order β, then the first term is of order $0(n^{-(2\beta+1)}a_n^{-(2\beta+2)p})$ and the second term is of order $0(n^{-2\beta}a_n^{-(2\beta+2)p})$, provided that $\max\limits_{1\leq j\leq n}\Delta(R_j) = 0(n^{-1})$. Thus if $na_n^{(1+1/\beta)p} \to \infty$ as

$n \to \infty$, we conclude, in view of Theorem 2, that: If the conditions of Theorem 1 hold, if $k(\underset{\sim}{u})$ is Lipschitz of order β and if $na_n^{(1+1/\beta)p} \to \infty$ then $\tilde{g}_n(\underset{\sim}{x}) \xrightarrow{P} g(\underset{\sim}{x})$ as $n \to \infty$ for all $\underset{\sim}{x} \in [0,1]^p$.

Note that while the above result improves that of Priestley and Chao (1972) it is still weaker than that of Theorem 2, but the calculation of $\tilde{g}_n(\underset{\sim}{x})$ is easier than the calculation of $g_n(\underset{\sim}{x})$, since the latter requires evaluation of $\int_{R_j} k[(\underset{\sim}{x}-\underset{\sim}{u})/a_n]d\underset{\sim}{u}$ which may not be easy for some kernels such as a multivariate normal. We shall show in Section 4, however, that the optimum choice of $k(\underset{\sim}{u})$, is a p.d.f. such that the evaluation of $\int_{R_j} k[(\underset{\sim}{x}-\underset{\sim}{u})/a_n]d\underset{\sim}{u}$ is not difficult whenever R_j is properly devised.

On the other hand, note that it follows from the above discussion that $(na_n^p)E(g_n(\underset{\sim}{x})-\tilde{g}_n(\underset{\sim}{x}))^2 = O(n^{-2\beta+1}a_n^{-(2\beta+1)p})$. Hence, if $na_n^{(2\beta+1)p/(2\beta-1)} \to \infty$ as $n \to \infty$, if the conditions of Theorem 6 are satisfied, and if $k(\underset{\sim}{u})$ is Lipschitz of order β, then $(na_n^p)^{1/2}(\tilde{g}_n(\underset{\sim}{x})-g(\underset{\sim}{x}))/[\sigma^2\sum_{j=1}^n c_j V_j^2]^{1/2} \xrightarrow{D} N(0,1)$, as $n \to \infty$; compare this result with Theorem 2 of Benedetti (1977).

We close this section by noticing that the estimate $\tilde{g}_n(\underset{\sim}{x})$ gives rise to a different estimate of σ^2, namely we can define

$$\tilde{\sigma}_n^2 = \frac{1}{n}\sum_{j=1}^n (Y_j-\tilde{g}_n(\underset{\sim}{x}_i))^2. \qquad (2.24)$$

In Section 3, we shall demonstrate that $\tilde{\sigma}_n^2$ is also consistent but it requires stronger conditions than those needed for the consistency of σ_n^2.

3. CONSISTENCY OF σ_n^2.

In this section we show that σ_n^2 is weakly consistent. To this end note that we can write

$$n\sigma_n^2 = \sum_{j=1}^n (Y_j - g(\underset{\sim}{x}_j))^2 + \sum_{j=1}^n (g_n(\underset{\sim}{x}_j) - g(\underset{\sim}{x}_j))^2$$

$$+ 2\sum_{j=1}^n (Y_j - g(\underset{\sim}{x}_j))(g_n(\underset{\sim}{x}_j) - g(\underset{\sim}{x}_j)) = I_{1n} + I_{2n} + I_{3n}, \quad \text{say.} \tag{3.1}$$

Note that $n^{-1}I_{1n} \xrightarrow{P} \sigma^2$ as $n \to \infty$ by the weak law of large numbers. Thus we need to show $n^{-1}I_{in} \xrightarrow{P} 0$, $i = 2,3$, as $n \to \infty$. A stronger conclusion would be to show that $n^{-1}EI_{in} \to 0$, $i = 2,3$, as $n \to \infty$. But

$$n^{-1}EI_{2n} = n^{-1}\{\sum_{j=1}^n \mathrm{Var}(g_n(\underset{\sim}{x}_j)) + \sum_{j=1}^n [Eg_n(\underset{\sim}{x}_j) - g(\underset{\sim}{x}_j)]^2\}. \tag{3.2}$$

In view of (2.19), $na_n^p \mathrm{Var} g_n(\underset{\sim}{x}) \approx \sigma^2 \sum_{j=1}^n c_j V_j^2 = O(1)$, as $n \to \infty$. Hence the first term of (3.2) is readily seen to be of order $O((na_n^p)^{-1}) = o(1)$ if $na_n^p \to \infty$ as $n \to \infty$. Next, if the conditions of Lemma 2 hold with $M = 2$, then the second term of (3.2) is of order $O(a_n^{4p}) = o(1)$. As for $n^{-1}EI_{3n}$, we have

$$n^{-1}EI_{3n} \leq 2n^{-1}\{\sum_{j=1}^n E^{1/2}(Y_j - g(\underset{\sim}{x}_j))^2 E^{1/2}(g_n(\underset{\sim}{x}_j) - g(\underset{\sim}{x}_j))^2\}$$

$$\leq O(n^{-1/2}a_n^{3p/2}) = o(1), \tag{3.3}$$

under the same conditions used in the proof of (3.2). Hence we arrive at

THEOREM 7. *Assume that the conditions of Lemma 2 are satisfied with* $M = 2$. *Then*

$$\sigma_n^2 \xrightarrow{P} \sigma^2 \quad \text{as} \quad n \to \infty. \tag{3.4}$$

Note that we can derive an analogous result for $\tilde{\sigma}_n^2$ under a bit stronger conditions, viz., write

$$\tilde{\sigma}_n^2 = n^{-1}\sum_{j=1}^{n}(Y_j-\tilde{g}_n(\underset{\sim}{x}_j))^2 = \sigma_n^2 + n^{-1}\sum_{j=1}^{n}(g_n(\underset{\sim}{x}_j)-\tilde{g}_n(\underset{\sim}{x}_j))^2$$
$$+ 2n^{-1}\sum_{j=1}^{n}(Y_j-g_n(\underset{\sim}{x}_j))(g_n(\underset{\sim}{x}_j)-\tilde{g}_n(\underset{\sim}{x}_j)) .$$

But $\sigma_n^2 \xrightarrow{P} \sigma^2$ and $E(g_n(\underset{\sim}{x})-\tilde{g}_n(\underset{\sim}{x}))^2 \longrightarrow 0$ if $k(\underset{\sim}{u})$ is Lipschitz of order β and $na_n^{(1+1/\beta)p} \to \infty$. Thus we can easily see that the second and the third terms of the above expression of $\tilde{\sigma}_n^2$ have expected values converging to 0, as $n \to \infty$.

It is possible also to obtain the second mean convergence of σ_n^2 (and thus of $\tilde{\sigma}_n^2$) under the extra assumption that $na_n^{2p} \to \infty$ as $n \to \infty$ and $Ee_1^4 < \infty$. To see this, we need to show that $\mathrm{Var}\sigma_n^2 \to 0$ as $n \to \infty$. Now,

$$\mathrm{Var}(\sigma_n^2) = n^{-2}\{\sum_{j=1}^{n}\mathrm{Var}[(Y_j-g_n(\underset{\sim}{x}_j))^2] + \sum_{j\neq j^*}\mathrm{Cov}[(Y_j-g_n(\underset{\sim}{x}_j)),$$
$$(Y_{j^*}-g_n(\underset{\sim}{x}_{j^*}))]\} = n^{-2}\{J_{1n}+J_{2n}\}, \text{ say.} \tag{3.5}$$

But

$$\mathrm{Var}[(Y_j-g_n(\underset{\sim}{x}_j))^2] = \mathrm{Var}(Y_j^2) + \mathrm{Var}(g_n^2(\underset{\sim}{x}_j)) + 4\mathrm{Var}(Y_jg_n(\underset{\sim}{x}_j))$$
$$+ 2\mathrm{Cov}(Y_j^2, g_n^2(\underset{\sim}{x}_j))$$
$$- 4\mathrm{Cov}(Y_j^2, Y_jg_n(\underset{\sim}{x}_j)) - 4\mathrm{Cov}(Y_jg_n(\underset{\sim}{x}_j), g_n^2(\underset{\sim}{x}_j)). \tag{3.6}$$

Whenever $\max_{1\le j\le n} \Delta(R_j) = 0(n^{-1})$ and $na_n^{2p} \to \infty$ as $n \to \infty$, it is not difficult to obtain that

$\mathrm{Var}(g_n^2(\underset{\sim}{x}_j)) = 0(n^{-2}a_n^{-4p})$, $\mathrm{Var}(Y_j g_n(\underset{\sim}{x}_j)) = 0(n^{-1}a_n^{-2p})$,

$\mathrm{Cov}(Y_j^2, g_n^2(\underset{\sim}{x}_j)) = 0(n^{-1}a_n^{-2p})$, $\mathrm{Cov}(Y_j^2, Y_j g_n(\underset{\sim}{x}_j)) = 0(n^{-1}a_n^{-p})$, and $\mathrm{Cov}(g_n^2(\underset{\sim}{x}_j), Y_j g_n(\underset{\sim}{x}_j)) = 0(n^{3/2}a_n^{-3p})$.

Hence it follows that $n^{-2}\sum_{j=1}^{n}\mathrm{Var}[(Y_j - g_n(\underset{\sim}{x}_j))^2] \to 0$ as $n \to \infty$. Similar but perhaps more tedious algegra also reveals that $n^{-2}\sum_{j\neq j*}\mathrm{Cov}((Y_j - g_n(\underset{\sim}{x}_j)), (Y_{j*} - g_n(\underset{\sim}{x}_{j*}))) \to 0$ as $n \to \infty$, provided that $na_n^{2p} \to \infty$.

Finally note that one can combine Theorems 6 and 7 to conclude that an asymptotically normal confidence interval for $g(\underset{\sim}{x})$ can be constructed with the limits

$$g_n(\underset{\sim}{x}) \pm z_{\alpha/2}[\sigma_n^2 \sum_{j=1}^{n} c_j V_j^2]^{1/2}$$

where z_α denote the upper $100\alpha\%$ point of the standard normal distribution and V_j^2 is given in (2.21).

4. OPTIMAL CHOICE OF THE KERNEL.

We now proceed to find the kernel $k(\underset{\sim}{u})$ which minimizes the mean square error $E(g_n(\underset{\sim}{x}) - g(\underset{\sim}{x}))^2$. Note that since the present regression estimation problem resembles the density estimation problem, it is not surprising that the optimal choice of the kernel function for our problem turns out to be exactly the same as that derived for the density estimation problem. For the latter case when $k(\underset{\sim}{u}) = \Pi_{i=1}^{p} k(u_i)$ where $k(.)$ is a bounded univariate p.d.f. such that $|u|k(u) \to 0$ as $|u| \to \infty$, see Epanechnikov (1969). Assume for the remaining of the study that $\Delta(R_j) = 1/n$, $j = 1,\ldots,n$. Let $\tilde{k}(\underset{\sim}{t})$ denote the characteristic function of $k(\underset{\sim}{u})$, i.e.,

$$\tilde{k}(\underset{\sim}{t}) = \int e^{i\underset{\sim}{t}'\underset{\sim}{u}} k(\underset{\sim}{u})d\underset{\sim}{u}. \tag{4.1}$$

Then we can write

$$Eg_n(\underset{\sim}{x}) - g(\underset{\sim}{x}) = (2\pi)^{-p}\{\textstyle\sum_{j=1}^{n} g(\underset{\sim}{x}_j)\int_{R_j}[\int e^{i\underset{\sim}{w}'(\underset{\sim}{x}-\underset{\sim}{u})}\tilde{k}(a_n\underset{\sim}{w})d\underset{\sim}{w}]d\underset{\sim}{u}\}$$

$$- (2\pi)^{-p}\int e^{i\underset{\sim}{w}'\underset{\sim}{x}}\phi_g(\underset{\sim}{w})d\underset{\sim}{w}, \qquad (4.2)$$

where $\phi_g(\underset{\sim}{w}) = \int_{[0,1]^p} e^{i\underset{\sim}{w}'\underset{\sim}{x}} g(\underset{\sim}{x})d\underset{\sim}{x}$.

Thus

$$Eg_n(\underset{\sim}{x}) - g(\underset{\sim}{x}) = (2\pi)^{-p}\{\int e^{i\underset{\sim}{w}'\underset{\sim}{x}}[\tilde{k}(a_n\underset{\sim}{w})\textstyle\sum_{j=1}^{n} g(\underset{\sim}{x}_j)\int_{R_j} e^{-i\underset{\sim}{w}'\underset{\sim}{u}}d\underset{\sim}{u} - \phi_g(\underset{\sim}{w})]d\underset{\sim}{w}\}$$

$$= (2\pi)^{-p}\{\int e^{i\underset{\sim}{w}'\underset{\sim}{x}}\tilde{k}(a_n\underset{\sim}{w})[\textstyle\sum_{j=1}^{n} g(\underset{\sim}{x}_j)\int_{R_j} e^{-i\underset{\sim}{w}'\underset{\sim}{u}}d\underset{\sim}{u} - \phi_g(\underset{\sim}{w})]d\underset{\sim}{w}$$

$$+ \int e^{i\underset{\sim}{w}'\underset{\sim}{x}}\phi_g(\underset{\sim}{w})[\tilde{k}(a_n\underset{\sim}{w}) - 1]d\underset{\sim}{w}. \qquad (4.3)$$

Now, if g is bounded and continuous, then by the dominated convergence theorem,

$$\textstyle\sum_{j=1}^{n} g(\underset{\sim}{x}_j)\int_{R_j} e^{i\underset{\sim}{w}'\underset{\sim}{u}}d\underset{\sim}{u} \to \phi_g(\underset{\sim}{w}) \quad \text{as} \quad n \to \infty. \qquad (4.4)$$

Thus the first term in the right-hand-side (rhs) of (4.3) converges to 0 as $n\to\infty$. Next, if there exist positive $r_1,\ldots,r_p$ such that $[1-k(\underset{\sim}{u})]/\prod_{i=1}^{p}|u_i|^{r_i} \to k_{r_1,\ldots,r_p}$, a non zero constant, as $||\underset{\sim}{u}|| \to 0$, then $r_1,\ldots,r_p$ are called the characteristic exponents of k and $k_{r_1,\ldots,r_p}$ the characteristic coefficient.

Thus

$$a_n^{-\sum_{i=1}^{n} r_i}(2\pi)^{-p}\int e^{-i\underset{\sim}{w}'\underset{\sim}{x}}\phi_g(\underset{\sim}{w})[\tilde{k}(a_n\underset{\sim}{w}) - 1]d\underset{\sim}{w}$$

$$= (2\pi)^{-p}\int e^{i\underset{\sim}{w}'\underset{\sim}{x}}\phi_g(\underset{\sim}{w})\ \frac{\tilde{k}(a_n.\underset{\sim}{w})-1}{\prod_{i=1}^{p}|a_n w_i|^{r_i}}\left(\prod_{i=1}^{p}|w_i|^{r_i}\right)d\underset{\sim}{w}$$

$$\to k_{r_1,\ldots,r_p}\int e^{i.\underset{\sim}{w}'\underset{\sim}{x}}\prod_{i=1}^{p}|w_i|^{r_i}\phi_g(\underset{\sim}{w})d\underset{\sim}{w}$$

$$= k_{r_1,\ldots,r_p} h^{r_1,\ldots,r_p}(\underset{\sim}{x}), \quad \text{say.} \qquad (4.5)$$

Now, since $\mathrm{Var}\, g_n(\underset{\sim}{x}) \simeq \sigma^2 (na_n^p)^{-1} \int k^2(\underset{\sim}{u})d\underset{\sim}{u}$ for n sufficiently large we obtain that

$$E(g_n(\underset{\sim}{x}) - g(\underset{\sim}{x})))^2 \simeq \sigma^2 (na_n^p)^{-1} \int k^2(\underset{\sim}{u})d\underset{\sim}{u} + a_n^{2(\sum r_i)} |k_{r_1,\ldots,r_p} h^{r_1,\ldots,r_p}(\underset{\sim}{x})|^2, \qquad (4.6)$$

where $\sum r_i = \sum_{i=1}^p r_i$. Thus the rhs of (4.6) is minimized by choosing a_n as follows:

$$a_n = \{(p\sigma^2/n) \int k^2(\underset{\sim}{u})d\underset{\sim}{u}/2(\sum r_i)|k_{r_1,\ldots,r_p} h^{r_1,\ldots,r_p}(\underset{\sim}{x})|^2\}^{1/(p+2\sum r_i)}, \qquad (4.7)$$

and the minimum value is

$$(p + 2\sum_{i=1}^p r_i)\{\sigma^2 \int k^2(\underset{\sim}{u})d\underset{\sim}{u}\}^{2\sum r_i/(p+\sum r_i)} [(2\sum r_i)n]^{2\sum r_i/(p+\sum r_i)}$$

$$\times |k_{r_1,\ldots,r_p} h^{r_1,\ldots,r_p}(\underset{\sim}{x})|^{2/(p+\sum r_i)}, \qquad (4.8)$$

which tends to 0 at the rate $n^{-2\sum r_i/(p+\sum r_i)}$. Now suppose that among the special class of kernels $k(\underset{\sim}{u}) = \prod_{i=1}^p k(u_i)$, with $k(\cdot)$ a known bounded p.d.f., with $r_i = 2, i = 1,\ldots,p$ we want to find the one that minimizes (4.8). This problem is precisely that of Epanechnikov (1969) whose solution is found to be $k_0(u) = (3/4)(1-u^2)$, $|u| \le 1$, $= 0$, elsewhere.

REFERENCES

BENEDETTI, J.K. (1977). On nonparametric estimation of regression functions. J. Royal Statist. Soc. (B), 39, 248-253.

CACOULLOS, T. (1966). Estimation of a multivariate density. Ann. Inst. Statist. Math., 18, 179-189.

CHENG, K.F. and LIN, P.E. (1981a). Nonparametric estimation of a regression function. z. Wahrs. verw. Gebiete 57, 223-233.

CHENG, K.F. and LIN, P.E. (1981b). Nonparametric estimation of a regression function: limiting distribution. Austral. J. Statist. 23, 186-195.

EPANECHNIKOV, V.A. (1969). Nonparametric estimates of a multivariate probability density. Theor. Probability Appl. 14, 153-158.

FRANCK, W.E. and HANSON, D.L. (1966). Some results giving rates of convergence in the law of large numbers for weighted sum of independent random variables. Trans. Amer. Math. Soc. 124, 347-359.

LOÉVE, M. (1963). Probability Theory. (Third Edition). Van Nostrand Co., Princeton, N.J.

PARZEN, E. (1962). On estimation of a probability density and mode. Ann. Math. Statist., 33, 1065-1076.

PRIESTLEY, M.E. and CHAO, M.T. (1972). Nonparametric function fitting. J. Royal Statist. Soc. (B), 34, 385-392.

PRUITT, W.E. (1966). Summability of independent random variables. J. Math. Mech., 15, 769-776.

ROSENBLATT, M. (1956). Remarks on some nanparametric estimates of a density function. Ann. Math. Statist., 27, 822-837.

SCHUSTER, E.F. and YAKOWITZ, S. (1979). Contributions to the theory of nonparametric regression with application to system identification. Ann. Statist., 7, 139-149.

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER FSU Statistics M650 ONR Technical 160	2. GOVT ACCESSION NO. A133 482	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Fitting a Multiple Regression Function		5. TYPE OF REPORT & PERIOD COVERED Technical
		6. PERFORMING ORG. REPORT NUMBER FSU Statistics M650
7. AUTHOR(s) Ibrahim A. Ahmad and Pi-Erh Lin		8. CONTRACT OR GRANT NUMBER(s) ONR Contract N00014-80-C-0093
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics Florida State University Tallahassee, FL 32306		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Statistics & Probability Program Arlington, VA 22217		12. REPORT DATE December, 1982
		13. NUMBER OF PAGES 20
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report)
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Distribution Unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Function regression, consistency, asymptotic normality, optimal kernel, rates of convergence, kernel function.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Consider the p-dimensional unit cube $[0,1]^p$, $p \geq 1$. Partition $[0,1]^p$ into n regions, $R_{1,n},\ldots,R_{n,n}$ such that the volume $\Delta(R_{j,n})$ is of order $n^{-1}, j = 1,\ldots,n$. Select and fix a point in each of these regions so that we have $\underset{\sim}{x}_1^{(n)},\ldots,\underset{\sim}{x}_n^{(n)}$ Suppose that associated with the j-th predictor vector

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

$\underset{\sim}{x}_j^{(n)}$ there is an observable variable $Y_j^{(n)}$, $j = 1,\ldots,n$ satisfying the multiple regression model $Y_j^{(n)} = g(\underset{\sim}{x}_j^{(n)}) + e_j^{(n)}$, where g is an unknown an unknown function defined on $[0,1]^p$ and $\{e_j^{(n)}\}$ are independent identically distributed random variables with $Ee_1^{(n)} = 0$ and Var $e_1^{(n)} = \sigma^2 < \infty$. This paper proposes $g_n(\underset{\sim}{x}) = a_n^{-p}\sum_{j=1}^n Y_j^{(n)}\int_{R_{j,n}} k[(\underset{\sim}{x}-\underset{\sim}{u})/a_n]d\underset{\sim}{u}$ as an estimator of $g(\underset{\sim}{x})$, where $k(\underset{\sim}{u})$ is a known p-dimensional bounded density and $\{a_n\}$ is a sequence of reals converging to 0 as $n \to \infty$. Weak and strong consistency of $g_n(\underset{\sim}{x})$ and rates of convergence are obtained. Asymptotic normality of the estimator is established. Also proposed is $\sigma_n^2 = n^{-1}\sum_{j=1}^n(Y_j^{(n)} - g_n(\underset{\sim}{x}_j^{(n)}))^2$ as a consistent estimate of σ^2.

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

END

FILMED

10-83

DTIC