# A CLASS OF NONPARAMETRIC RECURSIVE ESTIMATORS OF A MULTIPLE REGRESSION FUNCTION

Isogai, Eiichi
Department of Mathematics, Faculty of Science, Niigata University

https://doi.org/10.5109/13343

出版情報：Bulletin of informatics and cybernetics. 20 (3/4), pp.33-44, 1983-03. Research Association of Statistical Sciences
バージョン：

*Bulletin of Informatics and Cybernetics Vol. 20, No. 3~4, 1983*

# A CLASS OF NONPARAMETRIC RECURSIVE ESTIMATORS OF A MULTIPLE REGRESSION FUNCTION

By

**Eiichi ISOGAI***

**Abstract**

Let $Z=(X, Y)$ be a $R^p \times R$-valued random vector having a (unknown) probability density function $f^*(x, y)$ with respect to Lebesgue measure. We wish to estimate a regression function $m(x) = E[Y|X=x]$. In this paper we propose a class of recursive estimators $\{m_n(x)\}$ based on a random sample $Z_1=(X_1, Y_1)$, $Z_2=(X_2, Y_2), \cdots$ from $Z$, and show the strong pointwise consistency and the asymptotic normality of $m_n(x)$ at a point $x$. We also deal with the optimality in the sense of asymptotic minimum variance.

## 1. Introduction

Let $(X, Y)$ be a $R^p \times R$-valued random vector having a probability density function (p. d. f.) $f^*(x, y)$ with respect to Lebesgue measure. Based on a sequence $Z_1=(X_1, Y_1)$, $Z_2=(X_2, Y_2)$, $Z_3=(X_3, Y_3), \cdots$, of independent identically distributed random vectors defined on a probability space $(\Omega, \mathfrak{F}, P)$ with the common (unknown) p. d. f. $f^*(x, y)$, we wish to estimate a regression function $m(x)=E[Y|X=x]$ (of $Y$ on $X$) which is assumed to exist. Györfi [5] discusses about the estimators of a nonparametric regression function and investigates universal consistency of these estimators.

Ahmad and Lin [1] proposed the recursive estimator $\tilde{m}_n(x)$ of the form

$$\tilde{m}_0(x)=\tilde{f}_0(x)\equiv 0 \tag{1.1}$$
$$\tilde{f}_n(x)=(h_n/h_{n-1})^p \tilde{f}_{n-1}(x)+K((x-X_n)/h_n)$$
$$\tilde{m}_n(x)=\tilde{m}_{n-1}(x)+\tilde{f}_n^{-1}(x)(Y_n-\tilde{m}_{n-1}(x))K((x-X_n)/h_n),$$

where $K(x)$ is a p. d. f. with certain properties and $\{h_n\}$ is a sequence of positive numbers. They give pointwise and uniform convergence results with weak and strong convergence. In addition they treat the joint asymptotic normality of $(nh_n^p)^{1/2}(\tilde{m}_n(x_1) -m(x_1), \cdots, \tilde{m}_n(x_k)-m(x_k))$ at distinct points $x_1, \cdots, x_k$. Devroye and Wagner [4] proposed the still simpler recursive estimator than that of $\{\tilde{m}_n(x)\}$. They discuss weak and strong consistency in $L^1$.

In this paper we propose a class of recursive estimators $\{m_n(x)\}$ of the form

$$m_0(x)\equiv 0, \quad f_0(x)\equiv c \quad \text{with } c \text{ being an arbitrary positive constant} \tag{1.2}$$
$$f_n(x)=f_{n-1}(x)+a_n\{K_n(x, X_n)-f_{n-1}(x)\}$$
$$m_n(x)=m_{n-1}(x)+a_n G(f_n(x))(Y_n-m_{n-1}(x))K_n(x, X_n)$$

* Department of Mathematics, Faculty of Science, Niigata University, Niigata, 950-21, Japan.

for each $n \geqq 1$, where

$$a_n = a/n \quad \text{with} \quad 0 < a \leqq 1, \tag{1.3}$$

$$G(y) = y^{-1} \quad \text{if} \quad y > 0$$

$$= 0 \quad \text{if} \quad y \leqq 0,$$

$$K_n(x, s) = h_n^{-p} K((x-s)/h_n) \quad \text{for} \quad x, s \in R^p \tag{1.4}$$

and $K(x)$ is a bounded p. d. f. on $R^p$ with respect to Lebesgue measure.

Remark 1.1. It is easy to see that $G(f_n(x)) = f_n^{-1}(x)$ for each $n \geqq 1$ and each $x \in R^p$ if either $a \neq 1$ in (1.3) or $K(x) > 0$ for all $x \in R^p$. Thus, in the case where $a = 1$ in (1.3) and $K(x) > 0$ for all $x \in R^p$, $m_n(x)$ in (1.2) coincides with $\tilde{m}_n(x)$ in (1.1) for each $n \geqq 1$ and each $x \in R^p$.

The purpose of this paper is to show the strong pointwise consistency and the asymptotic normality of $m_n(x)$ at a point $x$. We also deal with an optimal choice of the coefficient $a$ in (1.3) in the sense of asymptotic minimum variance. For a special type of the sequence $\{h_n\}$ our estimators are shown to be asymptotically more efficient than those of Ahmad and Lin [1], in the sense of Definition 5.1 given in Section 5, by choosing the coefficient a suitably.

In Section 2 we shall give auxiliary results needed later. In Section 3 the strong pointwise consistency of $m_n(x)$ will be shown. In Section 4 we shall give the asymptotic normality of $(nh_n^p)^{1/2}(m_n(x) - m(x))$ at a point $x$. Section 5 is devoted to the optimal choice of the coefficient a and comparison between the estimators of $\{\tilde{m}_n(x)\}$ and $\{m_n(x)\}$.

## 2. Preliminaries and Auxiliary Results

In this section we shall give some results which are needed for the sections that follow.

Let the bounded p. d. f. $K(x)$ in (1.4) satisfy

$$\|x\|^p |K(x)| \to 0 \quad \text{as} \quad \|x\| \to \infty, \tag{K1}$$

where $\|\cdot\|$ denotes the Euclidean norm on $R^p$. Let $\{h_n\}$ in (1.4) be a sequence of positive numbers converging to zero, on which some of the following conditions are imposed:

$$\sum_{n=1}^{\infty} (n^2 h_n^p)^{-1} < \infty \tag{H1}$$

$$n^{1+\eta} h_n^{p+4} \to 0 \quad \text{as} \quad n \to \infty \quad \text{for some} \quad \eta > 0 \tag{H2}$$

$$n h_n^p \to \infty \quad \text{as} \quad n \to \infty. \tag{H3}$$

We introduce some notations. Let

$$f(x) = \int_R f^*(x, y) dy, \qquad q(x) = \int_R y f^*(x, y) dy$$

$$g(x) = \int_R y^2 f^*(x, y) dy,$$

$$Var[Y \mid X=x]=(g(x)/f(x))-(q(x)/f(x))^2 \quad \text{if} \quad f(x)>0$$
$$=0 \quad \text{otherwise}$$

and

$$Q_n(x, z)=yK_n(x, u) \quad \text{for} \quad x\in R^p, \quad z=(u, y)\in R^p\times R \quad \text{and} \quad n\geqq 1,$$

where $K_n(x, u)$ is defined as (1.4). We assume that $f(x)$, $q(x)$ and $g(x)$ are finite for all $x\in R^p$ and that $m(x)=q(x)/f(x)$ if $f(x)>0$. Throughout this paper we assume that $E[|Y|]<\infty$, which guarantees the existence of $m(x)$. Define a sequence $\{q_n(x)\}$ as follows:

(2.1) $$q_0(x)\equiv 0$$
$$q_n(x)=\sum_{j=1}^{n} a_j\beta_{jn}Q_j(x, Z_j) \quad \text{for} \quad n=1, 2, \cdots,$$

where

$$\beta_{mn}=\prod_{j=m+1}^{n}(1-a_j) \quad \text{if} \quad n>m\geqq 0$$
$$=1 \quad \text{if} \quad n=m\geqq 0$$

and $a_n$ is given in (1.3).

Let $\gamma_1=1$ and $\gamma_n=\prod_{j=2}^{n}(1-a_j)$ for all $n\geqq 2$. It is clear that $\beta_{mn}=\gamma_n\gamma_m^{-1}$ for $n\geqq m\geqq 1$. Throughout this paper $C_1, C_2, \cdots$ denote suitable positive constants, not depending on all positive integers $n$. Isogai [6] gives the result

(2.2) $$C_1n^{-a}\leqq\gamma_n\leqq C_2n^{-a} \quad \text{for all} \quad n\geqq 1.$$

By the definition of $f_n(x)$ we get

(2.3) $$f_n(x)=\sum_{j=1}^{n} a_j\beta_{jn}K_j(x, X_j)+\beta_{0n}c \quad \text{for} \quad n=1, 2, \cdots.$$

Define $N$ as follows:

(2.4) $N$=smallest integer $n\geqq 1$ such that $f_n(x)>0$ if such an $n$ exists
$=+\infty$ otherwise.

LEMMA 2.1. *Suppose that for some* $\omega\in\Omega$ $N=N(\omega)$ *is finite. Then, for such* an $\omega$,

$$m_n(x)=q_n(x)/f_n(x) \quad \textit{for all} \quad n\geqq N$$
$$=0 \quad \textit{for} \quad N>n\geqq 1.$$

PROOF. We can easily get the following facts:

(2.5) $$f_n(x)>0 \quad \text{for} \quad n\geqq N \quad \text{and} \quad f_n(x)=0 \quad \text{for} \quad N>n\geqq 1$$

and

(2.6) $$K_n(x, X_n)=0 \quad \text{for} \quad N>n\geqq 1.$$

By (2.5) we have $G(f_n(x))=0$ for $N>n\geqq 1$, which yields $m_n(x)=0$ for $N>n\geqq 1$. By (1.2), (2.5) and the definition of $G(y)$ we get

(2.7) $$f_n(x)m_n(x)=\sum_{j=N}^{n} a_j\beta_{jn}Q_j(x, Z_j) \quad \text{for all} \quad n\geqq N.$$

Since by (2.6) $Q_n(x, Z_n)=0$ for $N>n\geqq 1$, using (2.1) and (2.7) we obtain $f_n(x)m_n(x)=$

$q_n(x)$ for all $n \geqq N$. Thus, by (2.5) we get $m_n(x)=q_n(x)/f_n(x)$ for all $n \geqq N$. This completes the proof.

LEMMA 2.2. *Let* $\{d_n\}$ *be a sequence of positive numbers converging to zero. If it holds that for some* $a>0$ *and some* $p>0$

$$\lim_{n\to\infty} n^{1-2a} d_n^p = 0$$

*and*

$$\lim_{n\to\infty} n^{1-2a} d_n^p \sum_{j=1}^{n} j^{2(a-1)} d_j^{-p} = \beta \quad \textit{with some constant } \beta>0,$$

*then for any positive integer* $m$

$$\sum_{j=m}^{n} a^2 j^{-2} \gamma_j^{-2} d_j^{-p} \sim a^2 \beta (n d_n^p \gamma_n^2)^{-1} \qquad \textit{as} \quad n\to\infty,$$

*where* " $\phi_n \sim \psi_n$ *as* $n\to\infty$ " *means that* $\phi_n/\psi_n \to 1$ *as* $n\to\infty$.

PROOF. It suffices to show that $\beta^{-1} n d_n^p \gamma_n^2 \sum_{j=m}^{n} j^{-2}\gamma_j^{-2} d_j^{-p} = \beta^{-1} n d_n^p \sum_{j=m}^{n} j^{-2} \beta_{jn}^2 d_i^{-p} \to 1$ as $n\to\infty$. Let any $\varepsilon$ $(0<\varepsilon<1)$ be fixed. Choose $\xi$ with $0<\xi<1$ such that $(1+2\varepsilon/3)(1+\xi)<1+\varepsilon$ and $(1-\varepsilon/3)(1-\xi)>1-\varepsilon$. Since $\beta_{jn} \sim j^a n^{-a}$ as $n \geqq j \to \infty$, there exists a positive integer $m_1$ $(>m)$ such that

$$\left(1-\frac{\varepsilon}{3}\right) j^{2a} n^{-2a} \leqq \beta_{jn}^2 \leqq \left(1+\frac{\varepsilon}{3}\right) j^{2a} n^{-2a} \qquad \text{for} \quad n \geqq j \geqq m_1. \tag{2.8}$$

It follows from (2.2) that

$$n^{2a}\gamma_n^2 \leqq C_3 \qquad \text{for all} \quad n \geqq 1. \tag{2.9}$$

By the assumptions of the lemma we have $\sum_{j=m_1}^{n} j^{2(a-1)} d_j^{-p} \to \infty$ as $n\to\infty$ and $\beta^{-1} n^{1-2a} d_n^p \cdot \sum_{j=m_1}^{n} j^{2(a-1)} d_j^{-p} \to 1$ as $n\to\infty$. Thus there exists a positive integer $m_2$ $(>m_1)$ such that for $n \geqq m_2$

$$1-\xi < \beta^{-1} n^{1-2a} d_n^p \sum_{j=m_1}^{n} j^{2(a-1)} d_j^{-p} < 1+\xi \tag{2.10}$$

and

$$0 \leqq C_3 \sum_{j=m}^{m_1-1} j^{-2}\gamma_j^{-2} d_j^{-p} \Big/ \sum_{j=m_1}^{n} j^{2(a-1)} d_j^{-p} < \frac{\varepsilon}{3}. \tag{2.11}$$

Hence by (2.8)$\sim$(2.11) we obtain that for $n \geqq m_2$

$$\begin{aligned} &\beta^{-1} n d_n^p \sum_{j=m}^{n} j^{-2}\beta_{jn}^2 d_j^{-p} \\ &\leqq \beta^{-1} n^{1-2a} d_n^p \sum_{j=m_1}^{n} j^{2(a-1)} d_j^{-p} \left\{ \left( C_3 \sum_{j=m}^{m_1-1} j^{-2}\gamma_j^{-2} d_j^{-p} \Big/ \sum_{j=m_1}^{n} j^{2(a-1)} d_j^{-p} \right) + \left(1+\frac{\varepsilon}{3}\right) \right\} \\ &< \left(1+\frac{2\varepsilon}{3}\right)(1+\xi) < 1+\varepsilon. \end{aligned} \tag{2.12}$$

On the other hand, by (2.8) and (2.10) we get

$$\beta^{-1} n d_n^p \sum_{j=m}^{n} j^{-2}\beta_{jn}^2 d_j^{-p} \geqq \left(1-\frac{\varepsilon}{3}\right) \beta^{-1} n^{1-2a} d_n^p \sum_{j=m_1}^{n} j^{2(a-1)} d_j^{-p} \tag{2.13}$$

$$>\left(1-\frac{\varepsilon}{3}\right)(1-\xi)>1-\varepsilon \qquad \text{for all} \quad n\geqq m_2.$$

Thus, by (2.12) and (2.13) we obtain the lemma. This completes the proof.

For any real-valued function $\theta$ on $R^p$, let $C(\theta)$ be a set of continuity of points of $\theta$ and let $\|\theta\|_\infty=\sup_{x\in R^p}|\theta(x)|$. By Lemmas 2.1 and 2.2 of Isogai [7] we have

LEMMA 2.3. *Let $\{d_n\}$ be a sequence of positive numbers converging to zero. Suppose that $\theta(x)$ is an integrable, real-valued Borel measurable function on $R^p$ and that $K(x)$ is a bounded, integrable, real-valued Borel measurable function on $R^p$ satisfying* (K1). *Then, for each $x\in C(\theta)$,*

$$\int_{R^p} d_n^{-p}K((x-u)/d_n)\theta(u)du\to\theta(x)\int_{R^p}K(u)du \qquad \text{as} \quad n\to\infty$$

*and*

$$\sup_{n\geqq 1}\int_{R^p} d_n^{-p}|K((x-u)/d_n)||\theta(u)|du\leqq M,$$

*where $M$ is a positive constant depending on $x$.*

## 3. Strong pointwise consistency

In this section the strong pointwise consistency of $m_n(x)$ will be shown. In order to prove the strong pointwise consistency of $m_n(x)$ we shall give two lemmas.

LEMMA 3.1. *Let* (H1) *be satisfied. Then, for each point $x\in C(f)$,*

$$\lim_{n\to\infty} f_n(x)=f(x) \quad \textit{with probability one } (w.p.1).$$

PROOF. By (2.3) we get

$$f_n(x)-E[f_n(x)]=\gamma_n\sum_{j=1}^n \gamma_j^{-1}a_jU_j(x) \qquad \text{for each} \quad n\geqq 1, \tag{3.1}$$

where

$$U_n(x)=K_n(x, X_n)-E[K_n(x, X_n)]. \tag{3.2}$$

It follows from Lemma 2.3 that

$$h_n^pE[U_n^2(x)]\leqq h_n^pE[K_n^2(x, X_n)]\leqq C_3 \qquad \text{for all} \quad n\geqq 1. \tag{3.3}$$

From (H1) and (3.3) we get

$$\sum_{n=1}^\infty a_n^2E[U_n^2(x)]<\infty. \tag{3.4}$$

Thus by making use of (3.1), (3.4), the Khintchine-Kolmogorov convergence theorem (see Chow and Teicher [3], page 110) and the Kronecker lemma (see Loève [8], page 238) we have

$$f_n(x)-E[f_n(x)]\to 0 \quad \text{w.p.1 as} \quad n\to\infty. \tag{3.5}$$

Since $E[f_n(x)]=\sum_{j=1}^n a_j\beta_{jn}E[K_j(x, X_j)]+\beta_{0n}c$ and $\lim_{n\to\infty}\beta_{0n}=0$, it follows from Lemma 2.3 and the Toeplitz lemma (see Loève [8], page 238) that

(3.6) $$E[f_n(x)]\to f(x) \quad \text{as} \quad n\to\infty.$$

Thus by (3.5) and (3.6) the lemma is proved.

LEMMA 3.2. *Assume that* $E[Y^2]<\infty$. *Let* (H1) *be satisfied. Then, for each point* $x\in C(q)\cap C(g)$,

$$\lim_{n\to\infty} q_n(x)=q(x) \quad \text{w.p.1},$$

*where* $A\cap B$ *denotes intersection of two sets* $A$ *and* $B$.

PROOF. Since by (2.1) $E[q_n(x)]=\sum_{j=1}^{n} a_j\beta_{jn}E[Q_j(x, Z_j)]$ for each $n\geqq 1$, it follows from Lemma 2.3 and the Toeplitz lemma that

(3.7) $$E[q_n(x)]\to q(x) \quad \text{as} \quad n\to\infty.$$

It is clear that $q_n(x)-E[q_n(x)]=\gamma_n \sum_{j=1}^{n} \gamma_j^{-1}a_jV_j(x)$ for each $n\geqq 1$, where

(3.8) $$V_n(x)=Q_n(x, Z_n)-E[Q_n(x, Z_n)].$$

Since by Lemma 2.3 $h_n^pE[Q_n^2(x, Z_n)]\leqq C_3$ for all $n\geqq 1$, in the same manner as (3.5) we have that $q_n(x)-E[q_n(x)]\to 0$ w.p.1 as $n\to\infty$, which, together with (3.7), yields the conclusion of the lemma.

The strong pointwise consistency of $m_n(x)$ is obtained in the following:

THEOREM 3.1. *Assume that* $E[Y^2]<\infty$. *Let* (H1) *be satisfied. Then, for each point* $x\in C(f)\cap C(q)\cap C(g)$ *with* $f(x)>0$, *we have*

$$\lim_{n\to\infty} m_n(x)=m(x) \quad w.p.1.$$

PROOF. By Lemmas 3.1 and 3.2 there exists an event $\tilde{\Omega}$ with $P\{\tilde{\Omega}\}=1$ such that on $\tilde{\Omega}$

(3.9) $$\lim_{n\to\infty} f_n(x)=f(x) \quad \text{and} \quad \lim_{n\to\infty} q_n(x)=q(x).$$

Let any $\omega\in\tilde{\Omega}$ be fixed. Then by (3.9) and $f(x)>0$ $N(\omega)$ defined by (2.4) is finite. Since $m(x)=q(x)/f(x)$ we obtain, by (3.9) and Lemma 2.1, that $\lim_{n\to\infty} m_n(x)=m(x)$, which concludes the theorem.

## 4. Asymptotic normality

In this section we shall show the asymptotic normality of $(nh_n^p)^{1/2}(m_n(x)-m(x))$ at a point $x$. It will turn out that we are able to express in a direct way the dependency of the asymptotic variance on the choice of the sequence $\{h_n\}$ and the coefficient $a$.

Let $U_n(x)$ and $V_n(x)$ be defined as (3.2) and (3.8), respectively. Also, let

$$t(x)=\int_R |y|^3f^*(x, y)dy, \quad W_n(x)=a_n\gamma_n^{-1}(U_n(x), V_n(x))'$$

and

(4.1) $$B_n(x)=(nh_n^p)^{1/2}\gamma_n \sum_{j=1}^{n} W_j(x),$$

where the prime denotes transpose and we assume that $t(x)$ is finite for all $x\in R^p$. In order to prove the asymptotic normality we shall prove the following lemma.

LEMMA 4.1. *Assume that* $E[|Y|^3]<\infty$. *Let* Condition $A$ *hold.*

Condition $A$: *For some* $0<a\leqq 1$

(A1) $$\lim_{n\to\infty} n^{1-2a}h_n^p=0,$$

(A2) $$\lim_{n\to\infty} n^{1-2a}h_n^p\sum_{j=1}^{n} j^{2(a-1)}h_j^{-p}=\beta \quad \text{with some constant } \beta>0$$

*and*

(A3) $$\lim_{n\to\infty} (nh_n^p)^{3/2}n^{-3a}\sum_{j=1}^{n} j^{3(a-1)}h_j^{-2p}=0.$$

*Consider a point* $x\in C(f)\cap C(q)\cap C(g)$ *with* $f(x)>0$ *and* $Var[Y|X=x]>0$. *If either* $x\in C(t)$ *or* $\|t\|_\infty<\infty$ *holds then*

$$B_n(x)\xrightarrow[L]{} N_2(0,\Gamma(x)) \quad \text{as} \quad n\to\infty,$$

*where*

(4.2) $$\Gamma(x)=a^2\beta\int_{R^p}K^2(u)du\begin{pmatrix} f(x) & q(x)\\ q(x) & g(x)\end{pmatrix},$$

$N_k(0,\Gamma)$ *denotes the* $k$*-variate normal with mean vector* 0 *and variance-covariance matrix* $\Gamma$, *and* "$\xrightarrow[L]{}$" *means convergence in law.*

PROOF. By the Cramér-Wold theorem (see Billingsley [2], page 49), it suffices to show that for any $D'=(d_1, d_2)\in R^2$

(4.3) $$D'B_n(x)\xrightarrow[L]{} N_1(0, D'\Gamma(x)D) \quad \text{as} \quad n\to\infty.$$

We may assume $D\neq 0$. Let $\sigma_n^2(x)$ be the variance of $D'B_n(x)$. It holds that

(4.4) $$D'B_n(x)/\sigma_n(x)\xrightarrow[L]{} N_1(0,1) \quad \text{as} \quad n\to\infty$$

if we verify Lyapounov's condition

(4.5) $$(nh_n^p)^{3/2}\gamma_n^3\sum_{j=1}^{n} E[|D'W_j(x)|^3]/\sigma_n^3(x)\to 0 \quad \text{as} \quad n\to\infty.$$

It is easy to see that

(4.6) $$\sigma_n^2(x)=nh_n^p\gamma_n^2\sum_{j=1}^{n} a_j^2\gamma_j^{-2}\{d_1^2E[U_j^2(x)]+d_2^2E[V_j^2(x)]+2d_1d_2E[U_j(x)V_j(x)]\}.$$

By Lemma 2.3 we get as $n\to\infty$

(4.7) $$h_n^pE[U_n^2(x)]\to f(x)\int K^2(u)du,$$

$$h_n^pE[V_n^2(x)]\to g(x)\int K^2(u)du$$

and

$$h_n^pE[U_n(x)V_n(x)]\to q(x)\int K^2(u)du,$$

where the domain of integral is $R^p$ unless otherwise specified. According to Lemma 2.2, (A1), (A2) and (2.2), we have that $\sum_{j=1}^{n} a_j^2\gamma_j^{-2}h_j^{-p}\nearrow\infty$ as $n\to\infty$, which, together with

(4.7) and the Toeplitz lemma, yields that

$$\left(\sum_{j=1}^{n} a_j^2\gamma_j^{-2}h_j^{-p}\right)^{-1}\sum_{j=1}^{n} a_j^2\gamma_j^{-2}E[U_j^2(x)]\to f(x)\int K^2(u)du \tag{4.8}$$

as $n\to\infty$. Thus by making use of Lemma 2.2 and (4.8) we obtain

$$nh_n^p\gamma_n^2\sum_{j=1}^{n} a_j^2\gamma_j^{-2}E[U_j^2(x)]\to a^2\beta f(x)\int K^2(u)du \quad \text{as} \quad n\to\infty. \tag{4.9}$$

By the same argument for (4.9) we have as $n\to\infty$

$$nh_n^p\gamma_n^2\sum_{j=1}^{n} a_j^2\gamma_j^{-2}E[V_j^2(x)]\to a^2\beta g(x)\int K^2(u)du \tag{4.10}$$

and

$$nh_n^p\gamma_n^2\sum_{j=1}^{n} a_j^2\gamma_j^{-2}E[U_j(x)V_j(x)]\to a^2\beta q(x)\int K^2(u)du\,. \tag{4.11}$$

Combining (4.6), (4.9), (4.10) and (4.11) we get

$$\sigma_n^2(x)\longrightarrow D'\Gamma(x)D \quad \text{as} \quad n\to\infty. \tag{4.12}$$

By the assumptions that $f(x)>0$ and $Var[Y\,|\,X=x]>0$ we get

$$D'\Gamma(x)D>0\,. \tag{4.13}$$

It can be easily shown that

$$E[|D'W_j(x)|^3]\leqq 4(|d_1|^3+|d_2|^3)a_j^3\gamma_j^{-3}\max\{E[|U_j(x)|^3],\ E[|V_j(x)|^3]\}. \tag{4.14}$$

It follows from Lemma 2.3 that

$$E[|U_j(x)|^3]\leqq C_3h_j^{-2p} \quad \text{for all} \quad j\geqq 1. \tag{4.15}$$

It is easy to see that

$$E[|V_j(x)|^3]\leqq 8\|K\|_\infty^2h_j^{-2p}\int h_j^{-p}K((x-u)/h_j)t(u)du \tag{4.16}$$

for all $j\geqq 1$. If $x\in C(t)$, then we get, by Lemma 2.3 and (4.16), $E[|V_j(x)|^3]\leqq C_4h_j^{-2p}$ for all $j\geqq 1$. If $\|t\|_\infty<\infty$, then by (4.16) we get $E[|V_j(x)|^3]\leqq 8\ \|K\|_\infty^2\|t\|_\infty h_j^{-2p}$ for all $j\geqq 1$. Thus we have, under either $x\in C(t)$ or $\|t\|_\infty<\infty$,

$$E[|V_j(x)|^3]\leqq C_5h_j^{-2p} \quad \text{for all} \quad j\geqq 1. \tag{4.17}$$

Combining (4.14), (4.15) and (4.17) we obtain

$$E[|D'W_j(x)|^3]\leqq C_6a_j^3\gamma_j^{-3}h_j^{-2p} \quad \text{for all} \quad j\geqq 1. \tag{4.18}$$

Since by (2.2) and (4.18)

$$(nh_n^p)^{3/2}\gamma_n^3\sum_{j=1}^{n} E[|D'W_j(x)|^3]\leqq C_7(nh_n^p)^{3/2}n^{-3a}\sum_{j=1}^{n} j^{3(a-1)}h_j^{-2p}\,,$$

it follows from (A3) that

$$(nh_n^p)^{3/2}\gamma_n^3\sum_{j=1}^{n} E[|D'W_j(x)|^3]\to 0 \quad \text{as} \quad n\to\infty,$$

which, together with (4.12) and (4.13), implies (4.5). In virtue of (4.4) and (4.12) we obtain (4.3). Thus the lemma is proved.

The following theorem is concerned with the asymptotic normality of $m_n(x)$ at a point $x$. In the remainder of this paper, let $K(x)$ satisfy the following additional conditions:

(K2) $$\int_{R^p} u_i K(u_1, \cdots, u_p) du_1 \cdots du_p = 0 \quad \text{for} \quad i=1, \cdots, p$$

and

(K3) $$\int_{R^p} \|u\|^2 K(u) du < \infty .$$

THEOREM 4.1. *Assume that* $E[|Y|^3]<\infty$. *Suppose that there exist bounded, continuous second partial derivatives* $\partial^2 f(x)/\partial x_i \partial x_j$ *and* $\partial^2 q(x)/\partial x_i \partial x_j$ *for* $i, j=1, \cdots, p$ *and that* $g(x)$ *is continuous on* $R^p$. *Let* Condition $A$ *in Lemma* 4.1, (H1), (H2) *and* (H3) *be fulfilled. Consider a point* $x$ *with* $f(x)>0$ *and* $Var[Y|X=x]>0$. *Then, under either* $x \in C(t)$ *or* $\|t\|_\infty<\infty$, *we have*

$$(nh_n^p)^{1/2}(m_n(x)-m(x)) \xrightarrow[L]{} N_1(0, \sigma^2(x)) \quad \text{as} \quad n\to\infty,$$

*where* $\sigma^2(x)=a^2\beta \, Var[Y|X=x]\int_{R^p} K^2(u)du/f(x)$.

PROOF. We note that by the assumptions Lemma 4.1 holds. Setting $B_n^*(x)=(nh_n^p)^{1/2}(f_n(x)-f(x), q_n(x)-q(x))'$, we get

(4.19) $$B_n^*(x)-B_n(x)=(nh_n^p)^{1/2}\gamma_n \sum_{j=1}^n a_j\gamma_j^{-1}D_j(x)+(nh_n^p)^{1/2}\beta_{0n}D_0(x),$$

where $B_n(x)$ is defined as (4.1) and let

$$d_{j1}(x)=E[K_j(x, X_j)]-f(x), \quad d_{j2}(x)=E[Q_j(x, Z_j)]-q(x)$$
$$D_j(x)=(d_{j1}(x), d_{j2}(x))' \quad \text{for} \quad j\geqq 1$$
$$D_0(x)=(c-f(x), -q(x))'.$$

It follows from (2.2) and (A1) that $(nh_n^p)^{1/2}\beta_{0n}\to 0$ as $n\to\infty$. Thus, if we show that for each $i=1, 2$

(4.20) $$(nh_n^p)^{1/2}\gamma_n \sum_{j=1}^n a_j\gamma_j^{-1}|d_{ji}(x)|\to 0 \quad \text{as} \quad n\to\infty,$$

then, it follows from (4.19) that $\lim_{n\to\infty}\|B_n^*(x)-B_n(x)\|=0$ with $\|\cdot\|$ denoting the Euclidean norm on $R^2$, which, together with Theorem 4.1 of Billingsley [2] and Lemma 4.1, yields that

(4.21) $$B_n^*(x) \xrightarrow[L]{} N_2(0, \Gamma(x)) \quad \text{as} \quad n\to\infty.$$

We shall now show (4.20). By the Taylor theorem, (K2), (K3) and the boundedness of $\partial^2 f(x)/\partial x_i \partial x_j$ and $\partial^2 q(x)/\partial x_i \partial x_j$, we get that for each $i=1, 2$ $|d_{ji}(x)|\leqq C_3 h_j^2$ for all $j\geqq 1$, which yields that for each $i=1, 2$

(4.22) $$(nh_n^p)^{1/2}\gamma_n \sum_{j=1}^n a_j\gamma_j^{-1}|d_{ji}(x)|\leqq C_4(nh_n^p)^{1/2}\gamma_n \sum_{j=1}^n j^{-1}\gamma_j^{-1}h_j^2 .$$

By making use of the Cauchy-Schwarz inequality we obtain

(4.23) $$nh_n^p\gamma_n^2\Big(\sum_{j=1}^n j^{-1}\gamma_j^{-1}h_j^2\Big)^2 \leqq \Big(\sum_{j=1}^\infty j^{-(1+\eta)}\Big)nh_n^p\gamma_n^2 \sum_{j=1}^n j^{-1+\eta}\gamma_j^{-2}h_j^4 ,$$

where $\eta$ is given in (H2). It follows from (H2) and the Toeplitz lemma that

$$\left(\sum_{j=1}^{n} j^{-2}\gamma_j^{-2}h_j^{-p}\right)^{-1}\sum_{j=1}^{n} j^{-1+\eta}\gamma_j^{-2}h_j^4\to 0 \quad \text{as} \quad n\to\infty,$$

which, together with Lemma 2.2, yields that

$$nh_n^p\gamma_n^2\sum_{j=1}^{n} j^{-1+\eta}\gamma_j^{-2}h_j^4\to 0 \quad \text{as} \quad n\to\infty. \tag{4.24}$$

Thus the relations (4.22), (4.23) and (4.24) imply (4.20). Let us define a function $T$ on $R^2$ as

$$\begin{aligned} T(u, v)&=v/u \quad \text{if} \quad u\neq 0 \\ &=0 \quad \text{if} \quad u=0. \end{aligned}$$

By the Taylor theorem and (2.5) we get

$$\begin{aligned} A_n(x)&\equiv(nh_n^p)^{1/2}\{T(f_n(x), q_n(x))-T(f(x), q(x))\} \\ &=L'B_n^*(x)+\varepsilon_n\|B_n^*(x)\| \quad \text{for all} \quad n\geqq N, \end{aligned} \tag{4.25}$$

where $N$ defined by (2.4) is assumed to be finite and $L=(-q(x)/f^2(x), f^{-1}(x))'$, and

$$\varepsilon_n\to 0 \quad \text{if} \quad \|(f_n(x)-f(x), q_n(x)-q(x))'\|\to 0. \tag{4.26}$$

It follows from Lemmas 3.1 and 3.2 that $\varepsilon_n\to 0$ w.p.1 as $n\to\infty$. Hence by (4.21) we get

$$\varepsilon_n\|B_n^*(x)\|\to 0 \text{ in probability as } n\to\infty. \tag{4.27}$$

Since by Lemma 3.1 $P\{N=\infty\}=0$, using (4.21), (4.25) and (4.27) we have

$$A_n(x)\xrightarrow[L]{} N_1(0, \sigma^2(x)) \quad \text{as} \quad n\to\infty. \tag{4.28}$$

On the other hand, by Lemma 2.1 and (2.5) we get

$$\begin{aligned} &P\{|A_n(x)-(nh_n^p)^{1/2}(m_n(x)-m(x))|\geqq\varepsilon\} \\ &\leqq P\{n<N\}\to 0 \quad \text{as} \quad n\to\infty \quad \text{for each} \quad \varepsilon>0, \end{aligned}$$

which, together with (4.28), yields that

$$(nh_n^p)^{1/2}(m_n(x)-m(x))\xrightarrow[L]{} N_1(0, \sigma^2(x)) \quad \text{as} \quad n\to\infty.$$

Thus the theorem is established.

## 5. Optimal Choice of $a$

In this section we shall give an optimal choice of the coefficient $a$, under a certain criterion, for a special sequence of $\{h_n\}$. A comparison between the estimators of (1.1) and (1.2) is also given. The following definition gives our criterion.

Definition 5.1. Let $\{b_n\}$ be a strictly increasing sequence of positive numbers with $\lim_{n\to\infty} b_n=\infty$. Suppose two sequences of estimators of some constant $\theta$, $\{U_n\}$ and $\{V_n\}$ satisfy

$$b_n(U_n-\theta)\xrightarrow[L]{} N_1(0, \sigma_1^2(\theta)) \quad \text{as} \quad n\to\infty$$

and

$$b_n(V_n-\theta)\xrightarrow[L]{} N_1(0, \sigma_2^2(\theta)) \quad \text{as} \quad n\to\infty,$$

where $\sigma_1^2(\theta)$ and $\sigma_2^2(\theta)$ are two positive constants depending on $\theta$. Then we define $ef(\{U_n\}, \{V_n\})$ as $ef(\{U_n\}, \{V_n\})=\sigma_1^2(\theta)/\sigma_2^2(\theta)$ and call $ef(\{U_n\}, \{V_n\})$ the relative asymptotic efficiency of $\{V_n\}$ to $\{U_n\}$. If $ef(\{U_n\}, \{V_n\})<1$ then the sequence of estimators $\{U_n\}$ is said to be asymptotically more efficient than that of the estimators $\{V_n\}$.

The following theorem gives the asymptotic variance of the estimators $\{m_n(x)\}$ and the optimal choice of the coefficient $a$, in the sence of the asymptotic minimum variance, for a special sequence of $\{h_n\}$.

THEOREM 5.1. *Let*

$$h_n=n^{-r/p} \quad with \quad p/(p+4)<r<1. \tag{5.1}$$

*Let the coefficient $a$ in* (1.3) *satisfy* $1\geqq a>(1-r)/2$. *Then, under all conditions of Theorem* 4.1 *except the conditions about the coefficient $a$ and $\{h_n\}$, we have*

$$n^{(1-r)/2}(m_n(x)-m(x)) \xrightarrow[L]{} N_1(0, \sigma^2(x)) \qquad as \quad n\to\infty,$$

*where* $\sigma^2(x)=a^2(2a+r-1)^{-1}Var[Y|X=x]\int_{R^p}K^2(u)du/f(x)$. *Furthermore,* $\sigma^2(x)$ *attains its minimum value* $(1-r)Var[Y|X=x]\int_{R^p}K^2(u)du/f(x)$ *at* $a=1-r$ *for fixed* $r$.

PROOF. We shall verify (H1), (H2), (H3) and Condition $A$. As $r<1$, (H1) and (H3) are satisfied. Since $p/(p+4)<r$, there exists a positive number $\eta$ such that $p(1+\eta)/(p+4)<r$, which implies (H2). After some calculations with $r>1-2a$, (A1), (A2) with $\beta=(2a+r-1)^{-1}$ and (A3) hold. Thus by Theorem 4.1 the first assertion is established. Since $a^2(2a+r-1)^{-1}$ attains its minimum at $a=1-r$ for fixed $r$, so does $\sigma^2(x)$, and its minimum value is $(1-r)Var[Y|X=x]\int_{R^p}K^2(u)du/f(x)$. This completes the proof.

We shall now compare two estimators of $\{\tilde{m}_n(x)\}$ and $\{m_n(x)\}$ under the criterion of the relative asymptotic efficiency. Let $\{h_n\}$ be given in (5.1). The following corollary is concerned with the comparison of $\{\tilde{m}_n(x)\}$ and $\{m_n(x)\}$, and show that the sequence of the estimators $\{m_n(x)\}$ is asymptotically more efficient than that of the estimators $\{\tilde{m}_n(x)\}$.

COROLLARY 5.1. *Let $K(x)$ be positive for all $x\in R^p$. Then, under all conditions of Theorem* 5.1 *with* $1>a>(1-r)/(1+r)$ *we obtain* $ef\equiv ef(\{m_n(x)\}, \{\tilde{m}_n(x)\})=a^2(1+r)/(2a+r-1)$ $(<1)$. *Furthermore, ef attains its minimum value* $1-r^2$ *at* $a=1-r$ *for fixed* $r$ *with* $p/(p+4)<r<1$.

PROOF. Note that $1-r>(1-r)/(1+r)>(1-r)/2$ for $0<r<1$. It follows from Remark 1.1 that $\tilde{m}_n(x)$ is obtained by putting $a=1$ in (1.3). Thus by Theorem 5.1 we get the first assertion. The second assertion is easily obtained. This completes the proof.

## References


[1] AHMAD, I. A. and LIN, P.: *Nonparametric sequential estimation of a multiple regression function*, Bull. Math. Statist., 17, No. 1~2 (1976), 63-75.

[2] BILLINGSLEY, P.: *Convergence of probability measures*, John Wiley and Sons, Inc., New York, (1968).

[3] CHOW, Y. S. and TEICHER, H.: *Probability theory: Independence, interchangeability, martingales*, Springer-Verlag, New York and Berlin, (1978).

[4] DEVROYE, L. P. and WAGNER, T. J.: *On the $L_1$ convergence of kernel estimators of regression functions with applications in discrimination,* Z. Wahrscheinlichkeitstheor. Verw. Geb. **51** (1980), 15-25.
[5] GYÖRFI, L.: *Recent results on nonparametric regression estimate and multiple classification,* Probl. Control. Inf. Theory, **10** (1981), 43-52.
[6] ISOGAI, E.: *Strong consistency and optimality of a sequential density estimator,* Bull. Math. Statist., **19**, No. 1~2 (1980), 55-69.
[7] ISOGAI, E.: *Strong uniform consistency of recursive kernel density estimators,* Sci. Rep. Niigata Univ., Ser. A, **18** (1982), 15-27.
[8] LOÈVE, M.: *Probability theory, 3rd ed.* Van Nostrand, New York, (1963).

*Communicated by N. Furukawa*
*Received May 10, 1982*