Chin. Ann. of Math.
8B (4) 1987

NECESSARY CONDITIONS OF L_1-CONVERGENCE OF KERNEL REGRESSION ESTIMATORS*

SUN DONGCHU (孙东初)**

Abstract

Let $(X_1,Y_1),\cdots,(X_n,Y_n)$ be iid. and $R^d\times R$-valued samples of (X,Y). The kernel estimator of the regression function $m(x)\triangleq E(Y|X=x)$ (if it exists), with kernel K, is denoted by

$$m_n(x)=\sum_{i=1}^n Y_iK\left(\frac{X_i-x}{h_n}\right)\Big/\sum_{j=1}^n K\left(\frac{X_j-x}{h_n}\right).$$

Many authors discussed the convergence of $m_n(x)$ in various senses, under the conditions $h_n\to 0$ and $nh_n^d\to\infty$ as $n\to\infty$. Are these conditions necessary? This paper gives an affirmativea nswer to this bprolemuithe case of L_1-conversence, when K satisfies (1.3) and $E(|Y|\log^+|Y|)<\infty$.

§ 1. Introduction and Results

Let (X_i, Y_i), $i=1,\cdots,n$ be a random sample from the $(d+1)$-dimensional distribution of (X, Y), where X is a d-vector and Y a scalar. Let $m(x)\triangleq E(Y|X=x)$ be the regression of Y on X (assuming it exists). Kernel estimators of $m(x)$ were introduced by Watson [1] and Nadaraya[2]. They proposed the estimator as

$$m_n(x)=\sum_{i=1}^n Y_iK\left(\frac{X_i-x}{h_n}\right)\Big/\sum_{j=1}^n K\left(\frac{X_j-x}{h_n}\right), \tag{1.1}$$

where kernel $K\geqslant 0$ is an integrable function on R^d and window size h_n is a sequence of positive numbers; we will treat 0/0 as 0. Criteria measuring the closeness of m_n to m include the distance in L_1,

$$J_n=\int|m_n(x)-m(x)|\,dF(x), \tag{1.2}$$

where F is the (unknown) marginal distribution of X. Recently, an increasing amount of attention is being given to the convergence of J_n. Devroye[3] proved the following theorem.

Theorem. *Suppose that $E|Y|<\infty$ and there exist positive numbers β, C_1, C_2, such that*

$$C_1I_{(\|x\|\leqslant\beta)}\leqslant K(x)\leqslant C_2I_{(\|x\|\leqslant\beta)}, \tag{1.3}$$

where I is the indicator function. If

Manuscript received January 23, 1985.

* Projects supported by he Science Fund of the Chinese Academy of Sciences.

** Department of Mathematical Statistics, East China Normal University, Shanghai, China.

$$h_n \to 0 \text{ as } n \to \infty \tag{1.4}$$

and

$$nh_n^d \to \infty \text{ as } n \to \infty, \tag{1.5}$$

then

$$E\,J_n = E\,|m_n(X) - m(X)| \to 0 \text{ as } n \to \infty. \tag{1.6}$$

It is easy to see (1.6) implies

$$J_n \to 0 \text{ in probability as } n \to \infty. \tag{1.7}$$

We remark that all norms are either all L_∞ or all L_2 throughout the paper.

In addition, many authors discussed the various convergence of $m_n(x)$ (see [4, 5] etc.). But all of them based on the same conditions (1.4) and (1.5). Intuitively, condition (1.4) needs the information of samples which is as near as possible to x, and condition (1.5) the information of samples as much as possible, i. e., the speed of h_n tends to 0 must not be too fast. In the present paper we discuss the inverse problem of the L_1-convergence of $m_n(x)$, i.e., whether or not conditions (1.4) and (1.5) are necessary when (1.7) holds.

Throughout this paper we impose the following conditions. Assume that the kernel function K satisfies (1.3), that the marginal distribution F of X has a density function f, and that $E|Y|\log^+|Y|<\infty$, where

$$\log^+ a = \begin{cases} \log a, & \text{as } a \geq 1, \\ 0, & \text{as } 0<a<1. \end{cases}$$

Our first theorem is used to prove the necessity of (1.5).

Theorem 1. *Under the above conditions, if* (1.4) *holds, then* (1.5) *is necessary when* (1.7) *holds.*

Our next result deals with the necessity of (1.4).

Theorem 2. *In addition to the above conditions, if f is bounded, K is continuous in the neighbor of zero, and for every $r>0$, $F\{X: g(n, r, X) = m(X)\}<1$, where*

$$g(n, r, x) = E\sum_{i=1}^{n} Y_i K\left(\frac{X_i - x}{r}\right) \Big/ \sum_{j=1}^{n} K\left(\frac{X_j - x}{r}\right), \quad x \in R^d,$$

then (1.4) *is necessary when* (1.7) *holds.*

To read smoothly, the following basic notations will be used whenever they are needed. Let $S_{x,r} = \{u: \|u-x\| \leq r\}$ be the closed sphere of radius r centered at x, $V(n, r, x) = \sum_{i=1}^{n} K((X_i - x)/r)$, and $W_{ni}(x) = K((X_i - x)/r)/V(n, r, x)$ for every $r>0$ and $x \in R^d$. For simplicity we sometimes write h for h_n.

§ 2. The Proof of Theorems

In order to prove these theorems we need some lemmas. For the proof of Lemma 1, see Theorem 3.1 in [3].

Lemma 1. *If $E|Y|<\infty$ and K satisfies* (1.3), *then there exists a positive* *constant* C_3 *independent of* F, $x\in R^d$, $r>0$ and positive *integers* n, *such that*

$$E\left\{\sum_{i=1}^{n}K((X_i-x)/r)\,|Y_i|/V(n,r,x)\right\}$$
$$\leqslant C_3\sup_{r>0}\left\{\int_{S_{x,r}}E(|Y|\,|X=u)F(du)/F(S_{x,r})\right\}$$
$$\triangleq C_3 m^*(x). \tag{2.1}$$

Furthermore, if $E|Y|\log^+|Y|<\infty$, *then* $Em^*(X)<\infty$.

Proof of Theorem 1 We assume first that $\lim_n nh_n^d=b\in(0,\infty)$ and prove by contradiction. By Fatou's Lemma and Fubini's Theorem, we have

$$EJ_n=E\left(\int|m_n(x)-m(x)|\,dF(x)\right)=\int E|m_n(x)-m(x)|\,dF(x)$$
$$\geqslant\int|Em_n(x)-m(x)|\,dF(x)=\int|g(n,h_n,x)-m(x)|\,dF(x).$$

Since (1.7) holds, by Lemma 1 and the Lebesgue dominated convergence theorem, we know (1.6) also holds. Hence

$$\int|g(n,h_n,x)-m(x)|\,dF(x)\to0 \text{ as } n\to\infty, \tag{2.2}$$

$$E\int|m_n(x)-g(n,h_n,x)|\,dF(x)\to0 \text{ as } n\to\infty. \tag{2.3}$$

Let $l\in(0,\infty)$ be a number. Define $A=\{X_i\,\overline{\in}\,S_{x,hl},\ i=1,\cdots,n\}$. Clearly

$$E\int|m_n(x)-g(n,h,x)|\,dF(x)=\int E|m_n(x)-g(n,h,x)|\,dF(x)$$
$$\geqslant\int|g(n,h,x)|P(A)\,dF(x)-\int|E\{I_Am_n(x)\}|\,dF(x)$$
$$\triangleq W_n-V_n. \tag{2.4}$$

By Lemma 1, we have

$$\liminf_n\int|g(n,h,x)|\,dF(x)\leqslant C_3Em^*(X)<\infty.$$

From (2.2), we get

$$g(n,h,X)\to m(X)$$

in probability as $n\to\infty$. Note that $0\leqslant P(A)\leqslant1$. By Fatou's Lemma, we have

$$\liminf_n W_n\geqslant\int\liminf_n|g(n,h,x)|\liminf_n P(A)\,dF(x)$$
$$=\int|m(x)|\liminf_n\{1-F(S_{x,hl})\}^n\,dF(x)$$
$$\geqslant\int|m(x)|\cdot\exp\{-\limsup_n[nF(S_{x,hl})/$$
$$(1-F(S_{x,hl}))]\}\,dF(x). \tag{2.5}$$

Because the density f of X exists, we get

$$F(S_{x,hl})/\lambda(S_{x,hl})\to f(x) \text{ as } n\to\infty, \text{ a. s. } x(\lambda),$$

where λ is a Lebesgue measure on R^d. Noting that $nh^d\to b$ as $n\to\infty$ and $\lambda(S_{x,hl})=Ml^dh^d$, we have

$$nF(S_{x,hl})\to Ml^d b\cdot f(x) \text{ as } n\to\infty, \text{ a. s. } x(\lambda). \tag{2.6}$$

From (2.5),

$$\liminf_n W_n \geqslant \int |m(x)| \exp\{-Ml^d\cdot b\cdot f(x)\} dF(x). \tag{2.7}$$

Also
$$V_n \leqslant \int E\left|I_A \sum_{i=1}^n m(X_i) K\left(\frac{X_i - x}{h_n}\right)\Big/ V(n, h_n, x)\right| dF(x)$$
$$\leqslant \int n\left\{E\left[Z \prod_{i=1}^{n-1} I_{S^c_{x,hl}}(X_i)\right]\int_{hl\leqslant\|u-x\|<h\beta} |m(u)| dF(u)\right\} dF(x),$$

where $Z=\min\{1, C_2/V(n-1, h, x)\}$, and $(\cdot)^c$ denotes the complement of a set. Clearly, $E(Z^2)\leqslant E(Z)$. By Cauchy-Schwarz's inequality

$$E\left[Z\prod_{i=1}^{n-1} I_{S^c_{x,hl}}(X_i)\right] \leqslant \left\{EZ^2\cdot E\left[\prod_{i=1}^{n-1} I_{S^c_{x,hl}}(X_i)\right]\right\}^{1/2}$$
$$\leqslant (EZ)^{1/2}[1-F(S_{x,hl})]^{(n-1)/2}.$$

It follows from the proof of Lemma 2.1 in [3] that

$$EZ \leqslant \frac{6C_2}{C_1}\{(n-1)F(S_{x,h\beta})\}^{-1}.$$

Hence
$$V_n \leqslant \frac{6C_2}{C_1}\int\left\{\left[\frac{1}{(n-1)F(S_{x,h\beta})}\right]^{1/2} n\exp\left[-\frac{n-1}{2}F(S_{x,hl})\right]\right.$$
$$\left.\int_{hl\leqslant\|u-x\|<h\beta} |m(u)| dF(u)\right\} dF(x)$$
$$= \frac{6C_2}{C_1}\left[\frac{n}{n-1}\right]^{1/2}\int\left\{[nF(S_{x,h\beta})]^{1/2}\exp\left[-\frac{n-1}{2}\cdot F(S_{x,hl})\right]\right.$$
$$\left.\cdot\frac{1}{F(S_{x,h\beta})}\int_{hl\leqslant\|u-x\|<h\beta} |m(u)| dF(u)\right\} dF(x).$$

Note that
$$\frac{1}{F(S_{x,h\beta})}\int_{hl\leqslant\|u-x\|<h\beta} |m(u)| dF(u)$$
$$\leqslant \frac{1}{F(S_{x,h\beta})}\int_{S_{x,h\beta}} |m(u)| dF(u)$$
$$\leqslant \sup_{r>0}\frac{1}{F(S_{x,r})}\int_{S_{x,r}} E(|Y|\,|X=u) dF(u).$$

From (2.6),

$$F(S_{x,hl})/F(S_{x,h\beta}) \to (1/\beta)^d \text{ as } n\to\infty, \text{ a.s. } x(\lambda). \tag{2.8}$$

Therefore $\left\{[nF(S_{x,h\beta})]^{1/2}\exp\left[-\frac{n-1}{2}F(S_{x,hl})\right]\right\}$ is bounded. By Lemma 1 and Fatou's Lemma

$$\limsup_n V_n \leqslant \frac{6C_2}{C_1}\int \limsup_n\left\{[nF(S_{x,h\beta})]^{1/2}\exp\left[-\frac{n-1}{2}F(S_{x,hl})\right]\right\}$$
$$\cdot\limsup_n\left\{\frac{1}{F(S_{x,h\beta})}\left[\int_{S_{x,h\beta}} |m(u)| dF(u)\right.\right.$$
$$\left.\left.-\int_{S_{x,hl}} |m(u)| dF(u)\right]\right\} dF(x).$$

From [6] we know that when (1.4) holds, for every $r>0$, there exists a Borel set B, such that $F(B)=1$ and $x\in B$ implies

$$\frac{1}{F(S_{x,hr})}\int_{S_{x,hr}}|m(u)|dF(u)\to m(x)$$

as $n\to\infty$. From (2.8) and the Lebesgue dominated convergence theorem, we conclude that

$$\limsup_n V_n\leqslant\frac{6C_2}{C_1}M^{\frac{1}{2}}\beta^{\frac{d}{2}}b^{\frac{1}{2}}\int f^{\frac{1}{2}}(x)\exp[-1/2Ml^d bf(x)]$$
$$\cdot|m(x)|dF(x)\cdot[1-(1/\beta)^d].$$

Because $\sup_x\left\{f^{\frac{1}{2}}(x)\exp\left[-\frac{1}{2}Ml^d bf(x)\right]\right\}$ is bounded, there exist positive numbers ε_0 and $l\in(0,\beta)$, such that the right hand side is bounded by ε_0, but $\int|m(x)|\exp[-Ml^d bf(x)]dF(x)\geqslant 2\varepsilon_0$. From (2.4) and (2.7), we get

$$\liminf_n E\int|m_n(x)-g(n,h,x)|dF(x)\geqslant\varepsilon_0>0,$$

which contradicts (2.3). Thus, no subsequence of nh^d can tend to a finite limit b, and (1.5) must holds. The proof of Theorem 1 is completed.

Lemma 2. *Suppose that $E|Y|\log^+|Y|<\infty$, K satisfies (1.3) and the density function f of X is bounded. Then*

$$\int|g(n,h,x)-g(n,b,x)|dF(x)\to 0 \text{ as } n\to\infty \tag{2.9}$$

whenever $\lim_n h_n=b\in(0,\infty)$.

Proof From Lemma 1, we get

$$\int|g(n,h,x)-g(n,b,x)|dF(x)\leqslant 2C_3Em^*(X)<\infty.$$

By the Lebesgue dominated convergence theorem, it suffices to prove that

$$|g(n,h,x)-g(n,b,x)|\to 0 \text{ as } n\to\infty \text{ a.s. } x(F). \tag{2.10}$$

Note that

$$\begin{aligned}
&|g(n,h,x)-g(n,b,x)|\\
&=n\left|E\left\{\frac{m(X_n)K((X_n-x)/h)}{K((X_n-x)/h)+V(n-1,h,x)}-\frac{m(X_n)K((X_n-x)/b)}{K((X_n-x)/b)+V(n-1,b,x)}\right\}\right|\\
&\leqslant nE|m(X_n)I_{(V(n-1,h,x)=0)}|+nE|m(X_n)I_{(V(n-1,b,x)=0)}|\\
&\quad+n\int_{V(n-1,b,x)>0}\int_{R^d}|m(u)|\frac{|K((u-x)/h)-K((u-x)/b)|}{K((u-x)/b+V(n-1,b,x)}dF(u)\prod_{i=1}^{n-1}dF(x_i)\\
&\quad+n\int_{R^{(n-1)d}}\int_{R^d}|m(u)|\left|\frac{K((u-x)/h)I_{(V(n-1,h,x)>0)}}{K((u-x)/h)+V(n-1,h,x)}\right.\\
&\quad\left.-\frac{K((u-x)/h)I_{(V(n-1,b,x)>0)}}{K((u-x)/b)+V(n-1,b,x)}\right|dF(u)\prod_{i=1}^{n-1}dF(x_i)\\
&\triangleq I_{n1}+I_{n2}+I_{n3}+I_{n4}.
\end{aligned}\tag{2.11}$$

If x is in the support of X, i.e. $F(S_{x,r})>0$ for every $r>0$, then

$$P(K(X_1-x)/r)>0)=P(X_1\in S_{x,r\beta})=F(S_{x,r\beta})>0.$$

Hence

$$I_{n1} \leqslant nE|Y|[1-P(K((X_1-x)/h)>0)]^{n-1}$$
$$\leqslant nE|Y|(1-F(S_{x,h\beta}))^{n-1} \to 0 \text{ as } n\to\infty. \tag{2.12}$$

Similarly

$$I_{n2} \leqslant n\, E|Y|(1-F(S_{x,b\beta}))^{n-1} \to 0 \text{ as } n\to\infty. \tag{2.13}$$

Since $E|m(X)|\leqslant E|Y|<\infty$, from [7], for any $\varepsilon>0$ there exists a bounded continuous function g vanishing outside a compact set, such that

$$\int |g(u)-m(u)|\,dF(u)\leqslant \varepsilon. \tag{2.14}$$

Denote $C_g=\sup_u |g(u)|$. Then

$$I_{n3}\leqslant nE\left\{\frac{1}{V(n-1,b,x)}I_{(V(n-1,b,x)>0)}\right\}\int |m(u)|\left|K\left(\frac{u-x}{h}\right)-K\left(\frac{u-x}{b}\right)\right|dF(u)$$
$$\leqslant nE\left\{\frac{1}{V(n-1,b,x)}I_{(V(n-1,b,x)>0)}\right\}\int |m(u)-g(u)|\left|K\left(\frac{u-x}{h}\right)\right.$$
$$\left.-K\left(\frac{u-x}{b}\right)\right|dF(u)$$
$$+C_g nE\left\{\frac{1}{V(n-1,b,x)}I_{(V(n-1,b,x)>0)}\right\}\int\left|K\left(\frac{u-x}{h}\right)-K\left(\frac{u-x}{b}\right)\right|dF(u).$$

Denote $p_r=F(S_{x,r})$, for $r>0$. Then for n sufficiently large

$$nE\left\{\frac{1}{V(n-1,b,x)}I_{(V(n-1,b,x)>0)}\right\}$$
$$\leqslant \frac{n}{C_1}E\left\{\left[\sum_{i=1}^{n-1}I_{S_{x,b}}(X_i)\right]^{-1}I_{\left[\sum_{i=1}^{n-1}I_{S_{x,b}}(X_i)>0\right]}\right\}$$
$$=\frac{n}{C_1}\sum_{i=1}^{n-1}\frac{1}{i}\cdot\frac{(n-1)!}{i!\,(n-1-i)!}p_{b\beta}^i(1-p_{b\beta})^{n-1-i}$$
$$\leqslant\frac{2}{C_1p_{b\beta}}\sum_{i=2}^{n}\frac{n!}{i!\,(n-i)!}p_{b\beta}^i(1-p_{b\beta})^{n-i}$$
$$=\frac{2}{C_1p_{b\beta}}\{1-(1-p_{b\beta})^n-np_{b\beta}(1-p_{b\beta})^{n-1}\}$$
$$\leqslant 4/(C_1p_{b\beta}). \tag{2.15}$$

In the last inequality, we use $0<p_{b\beta}\leqslant 1$.

Since $K\leqslant C_2$, we get

$$I_{n3}\leqslant\frac{4C_2}{C_1p_{b\beta}}\varepsilon+\frac{4C_g}{C_1p_{b\beta}}\int\left|K\left(\frac{u-x}{h}\right)-K\left(\frac{u-x}{b}\right)\right|dF(u). \tag{2.16}$$

Also

$$n\int_{V(n-1,h,x)=0}\int_{R^d}|m(u)|\cdot\left|\frac{K\left(\frac{u-x}{h}\right)I_{(V(n-1,h,x)>0)}}{K\left(\frac{u-x}{h}\right)+V(n-1,h,x)}\right.$$
$$\left.-\frac{K\left(\frac{u-x}{h}\right)I_{(V(n-1,b,x)>0)}}{K\left(\frac{u-x}{b}\right)+V(n-1,b,x)}\right|dF(u)\prod_{i=1}^{n-1}dF(x_i)$$

$$=n\int_{\substack{V(n-1,h,x)=0\\V(n-1,b,x)>0}}\int_{R^d}\frac{|m(u)|K\left(\frac{u-x}{h}\right)}{K\left(\frac{u-x}{b}\right)+V(n-1,b,x)}dF(u)\prod_{i=1}^{n-1}dF(x_i)$$

$$\leqslant C_2E|Y|nE\left\{\frac{1}{V(n-1,b,x)}I_{[V(n-1,h,x)=0,\ V(n-1,b,x)>0]}\right\}$$

$$\leqslant\frac{C_2}{C_1}E|Y|nE\{I_{(V(n-1,h,x)=0)}\}$$

$$\leqslant\frac{C_2}{C_1}E|Y|n(1-h_{h\beta})^{n-1}\to 0$$

as $n\to\infty$. Similarly,

$$n\int_{V(n-1,b,x)>0}\int_{R^d}|m(u)|\left|\frac{K\left(\frac{u-x}{h}\right)I_{(V(n-1,h,x)>0)}}{K\left(\frac{u-x}{h}\right)+V(n-1,h,x)}-\frac{K\left(\frac{u-x}{h}\right)I_{(V(n-1,b,x)>0)}}{K\left(\frac{u-x}{b}\right)+V(n-1,b,x)}\right|dF(u)\prod_{i=1}^{n-1}dF(x_i)\to 0$$

as $n\to\infty$. Hence

$$I_{n4}\leqslant n\int_{\substack{V(n-1,b,x)>0\\V(n-1,h,x)>0}}\int|m(u)|K\left(\frac{u-x}{h}\right)\left|\frac{1}{K\left(\frac{u-x}{h}\right)+V(n-1,h,x)}-\frac{1}{K\left(\frac{u-x}{b}\right)+V(n-1,b,x)}\right|dF(u)\prod_{i=1}^{n-1}dF(x_i)$$

$$\leqslant n\int_{\substack{V(n-1,b,x)>0\\V(n-1,h,x)>0}}|m(u)-g(u)|\cdot K\left(\frac{u-x}{h}\right)\cdot$$

$$\left|\frac{K\left(\frac{u-x}{b}\right)+V(n-1,b,x)-\left[K\left(\frac{u-x}{h}\right)+V(n-1,h,x)\right]}{\left[K\left(\frac{u-x}{b}\right)+V(n-1,h,x)\right]\left[K\left(\frac{u-x}{b}\right)+V(n-1,b,x)\right]}\right|dF(u)\prod_{i=1}^{n-1}dF(x_i)$$

$$+nC_gC_2\int_{\substack{V(n-1,b,x)>0\\V(n-1,h,x)>0}}\frac{\left|K\left(\frac{u-x}{b}\right)+V(n-1,b,x)-K\left(\frac{u-x}{h}\right)-V(n-1,h,x)\right|}{\left[K\left(\frac{u-x}{h}\right)+V(n-1,h,x)\right]\left[K\left(\frac{u-x}{b}\right)+V(n-1,b,x)\right]}$$

$$\cdot dF(u)\prod_{i=1}^{n-1}dF(x_i)$$

$$\leqslant C_2^2n^2\int_{\substack{V(n-1,b,x)>0\\V(n-1,h,x)>0}}\frac{1}{V(n-1,h,x)V(n-1,b,x)}\prod_{i=1}^{n-1}dF(x_i)\int|m(u)-g(u)|dF(u)$$

$$+C_2C_gn\int_{\substack{V(n-1,b,x)>0\\V(n-1,h,x)>0}}\frac{1}{V(n-1,h,x)V(n-1,b,x)}\prod_{i=1}^{n-1}dF(x_i)$$

$$\cdot\int\left|K\left(\frac{u-x}{b}\right)-K\left(\frac{u-x}{h}\right)\right|dF(u)$$

$$+C_2C_gn(n-1)\int_{\substack{V(n-1,b,x)>0\\V(n-1,h,x)>0}}\frac{\left|K\left(\frac{u-x}{h}\right)-K\left(\frac{u-x}{b}\right)\right|}{V(n-1,h,x)V(n-1,b,x)}dF(u)\prod_{i=1}^{n-2}dF(x_i)$$

$$+C_2C_gn(n-1)\int_{\substack{V(n-2,b,x)>0\\V(n-2,h,x)>0}}\frac{1}{V(n-2,h,x)V(n-2,b,x)}\prod_{i=1}^{n-2}dF(x_i)$$

$$\cdot\int\left|K\left(\frac{u-x}{h}\right)-K\left(\frac{u-x}{b}\right)\right|dF(u)+\frac{C_2C_g}{C_1^2}n(n-1)(1-F(S_{x,b\delta}))^{n-2}.$$

An argument similar to that leading to (2.15) gives

$$n^2E\{I_{(V(n-1,b,x)>0)}/V(n-1,b,x)\}^2\leqslant C_4<\infty,$$
$$n^2E\{I_{(V(n-1,h,x)>0)}/V(n-1,h,x)\}^2\leqslant C_4<\infty$$

for n large enough. By Cauchy-Schwarz inequality

$$n^2\int_{\substack{V(n-1,b,x)>0\\V(n-1,h,x)>0}}\frac{1}{V(n-1,h,x)V(n-1,b,x)}\prod_{i=1}^{n-1}dF(x_i)$$
$$\leqslant\left(n^2E\left\{\frac{I_{(V(n-1,\ b,\ x)>0)}}{V(n-1,b,x)}\right\}^2\right)^{1/2}\cdot\left(n^2E\left\{\frac{I_{(V(n-1,\ h,\ x)>0)}}{V(n-1,h,x)}\right\}^2\right)^{1/2}.$$

Observe that the above also holds when $n-1$ is replaced by $n-2$. From (2.14) and $n(n-1)(1-F(S_{x,b}))^{n-2}\to0$ as $n\to\infty$, we get

$$I_{n4}\leqslant C_2^2C_4\varepsilon+o(1)+C\int\left|K\left(\frac{u-x}{h}\right)-K\left(\frac{u-x}{b}\right)\right|dF(u),$$

where C is a positive constant. From (2.11)—(2.13), (2.15) and the arbitrariness of ε, it suffices to prove

$$\int\left|K\left(\frac{u-x}{h}\right)-K\left(\frac{u-x}{b}\right)\right|dF(u)\to0\text{ as }n\to\infty.\tag{2.17}$$

Since $\int K(u)du<\infty$, there exists a continuous function $K^*(\leqslant C_2)$ on R^d, such that

$$\int|K(u)-K^*(u)|du<\varepsilon.\tag{2.18}$$

Note that X has a bounded probability density

$$\int|K((u-x)/h)-K((u-x)/b)|dF(u)$$
$$\leqslant\int|K((u-x)/h)-K^*((u-x)/h)|f(u)du$$
$$+\int|K((u-x)/b)-K^*((u-x)/b)|f(u)du$$
$$+\int|K^*((u-x)/h)-K^*((u-x)/b)|f(u)du.$$

The first two terms on the right hand side of the inequality can be arbitrarily small by (2.18) and the boundedness of f, and the last term tends to 0 as $n\to\infty$ by the Lebesgue dominated convergence theorem, and this implies (2.17). The Lemma is thus proved.

Lemma 3. *Suppose that* $E|Y|<\infty$, K *satisfies* (1.3) *and* f *is continuous at zero. If*

$$h \to \infty \text{ as } n \to \infty, \tag{2.19}$$

then for every $x \in R^d$, we have

$$|Em_n(x) - EY| \to 0 \quad as\ n \to \infty. \tag{2.20}$$

Proof For any $x \in R^d$,

$$|Em_n(x) - EY| \leqslant J_1 + J_2 + J_3 + J_4, \tag{2.21}$$

where
$$J_1 = n \left| E \frac{K((X_n - x)/h) m(X_n)}{V(n, h, x)} I_{(V(n-1, h, x)=0)} \right|,$$

$$J_2 = n \left| E \left\{ \frac{m(X_n)[K((X_n - x)/h) - K(0)]}{V(n, h, x)} I_{(V(n-1, h, x)>0)} \right\} \right|,$$

$$J_3 = nK(0) \left| E \left\{ \left[\frac{m(X_n)}{V(n, h, x)} - \frac{m(X_n)}{K(0) + V(n-1, h, x)} \right] I_{(V(n-1, h, x)>0)} \right\} \right|,$$

$$J_4 = nK(0) \left| E \left\{ m(X_n) \left[\frac{I_{(V(n-1, h, x)>0)}}{K(0) + V(n-1, h, x)} - \frac{1}{nK(0)} \right] \right\} \right|.$$

An argument similar to that leading to (2.15) gives

$$J_1 \to 0 \text{ as } n \to \infty. \tag{2.22}$$

Since $E|m(X)| < \infty$, for any $\varepsilon > 0$ there exists a large positive number A, such that

$$\int_{S^c_{x,A}} |m(u)| dF(u) < \varepsilon. \tag{2.23}$$

Also, K is continuous at zero. We can find $\delta > 0$ such that $\|u\| < \delta$ implies $|K(u) - K(0)| < \varepsilon$. From (2.19), we know $h_n > A/\delta$ for n large enough, so that

$$|K(u/h) - K(0)| < \varepsilon \quad \text{as } \|u\| \leqslant A.$$

From (2.23), we get

$$\int |K((u-x)/h) - K(0)| \cdot |m(u)| dF(u)$$
$$\leqslant C_2 \int_{S^c_{x,A}} |m(u)| \cdot dF(u) + \int_{S_{x,A}} |K((u-x)/h) - K(0)| \cdot |m(u)| dF(u)$$
$$\leqslant C_2 \varepsilon + \varepsilon \int_{S_{0,A}} |m(u+x)| dF(u+x) \leqslant \varepsilon (C_2 + E|Y|). \tag{2.24}$$

Similarly

$$\int |K((u-x)/h) - K(0)| dF(u) \leqslant \varepsilon (C_2 + 1). \tag{2.25}$$

An argumeut similar to that leading to (2.15) gives that there exists a positive constant C such that

$$E\{n[V(n-1, h, x)]^{-1} I_{(V(n-1,h, x)>0)}\}^i < C, \quad i = 1, 2. \tag{2.26}$$

From (2.24) and (2.26)

$$J_2 \leqslant nE \left\{ \frac{I_{(V(n-1, h, x)>0)}}{V(n-1, h, x)} \right\} E \left| K\left(\frac{X_n - x}{h}\right) - K(0) \right| \cdot |m(X_n)| \to 0$$

and

$$J_3 \leqslant n^2 K(0) E \left\{ \frac{I_{(V(n-1, h, x)>0)}}{V(n-1, h, x)} \right\}^2 E \left| K\left(\frac{X_n - x}{h}\right) - K(0) \right| \cdot |m(X_n)| \to 0$$

as $n \to \infty$. From (2.25) and (2.26)

$$J_4 \leqslant E\left\{\frac{|V(n-1, h, x)-(n-1)K(0)|}{V(n-1, h, x)+K(0)} \cdot m(X_n) I_{(V(n-1, h, x)>0)}\right\}$$
$$\leqslant E|m(X_n)|(n-1)E\left\{\frac{I_{(V(n-2, h, x)>0)}}{K(0)+V(n-2, h, x)}\right\} \cdot E\left|K\left(\frac{X_n-x}{h}\right)-K(0)\right|+o(1)$$
$$\to 0$$

as $n\to\infty$. From (2.21) and (2.22), we get (2.20). Hence we have proved the lemma.

Proof of Theorem 2 By the proof of Theorem 1, we know (2.2) holds. We first assume that $\lim_n h=\infty$. Then by Fatou's Lemma,

$$\liminf_n |g(n, h, x)-m(x)|=0 \text{ a. s. } x(F).$$

By Lemma 3, we have

$$|g(n, h, x)-EY|\to 0 \text{ as } n\to\infty$$

for every $x\in R^d$. Thus, we get $m(x)=EY$, a.s. $x(F)$, i. e.,

$$g(n, r, x)=EY \text{ a.s. } x(F)$$

for every $r>0$. This, however, contradicts (2.2).

Next, we assume that $\lim_n h=b\in(0, \infty)$. Then

$$\int|g(n, h, x)-m(x)|dF(x)$$
$$\geqslant\int|g(n, b, x)-m(x)|dF(x)-\int|g(n, h, x)-g(n, b, x)|dF(x).$$

Since $F\{X:g(n, r, X)=m(X)\}<1$, we have

$$\int|g(n, b, x)-m(x)|dF(x)>0.$$

From (2.9), we get

$$\liminf_n \int|g(n, h, x)-m(x)|dF(x)>0,$$

which contradicts (2.2). Thus, we conclude that no subsequence of h_n can tend to a finite limit b. Therefore, $\lim_n h_n=0$. The proof of Theorem 2 is completed.

Acknowlegment. I shall like to thank Prof. We Zongshu and Zhao Lincheng for their enthusiastic help.

References

[1] Watson, O. S., Smooth Regression Analysis, *Sankhya, Series A*, **26** (1964), 359—372.
[2] Nadaraya, E. A., On Estimating Regression, *Theor. Probability Appl.*, **9** (1964), 141—142.
[3] Devroye, L. P., On the Almost Everywhere Convergence of Nonparametric Regression Function Estimates, *Ann. Statist*, **9** (1981), 1310—1319.
[4] Devroye, L. P. & Wagner, T. J., Distribution-free Consistency Results in Nonparametric Discrimination and Regression Function Estimation, *Ann. Statist*, **8** (1980), 231—239.
[5] Härdle, W., A Law of the Iterated Logarithm for Nonparametric Regression Function Estimates, *Ann. Statist*, **12** (1984), 624—635.
[6] Wheeden, R. L. & Zygmund, A., Measure and Integral, Marcel Dekker, New York (1977).
[7] Dunford, N. & Schwartz, J. T., Linear Operators, Part I, General Theory, Interscience, New York (1957).