非参数回归函数核估计的一致强收敛速度*

李　金　平

（中国科学技术大学）

一、引言及主要结果

设X为d维随机向量，Y为一维$r.v.$ (x_i, y_i)，$i=1,2,\cdots$，n为 (X, Y) 的独立随机样本．如果$E|Y|<\infty$，则$m(x)\hat{=}E(Y|X=x)$对几乎所有的x存在，称Y对X的回归函数．Watson (1964)[1], Nadaraya (1964)[2]提出用

$$m_n(x)=\frac{\sum_{i=1}^{n} Y_i K\left(\frac{x-X_i}{h_n}\right)}{\sum_{i=1}^{n} K\left(\frac{x-X_i}{h_n}\right)}$$

估计$m(x)$，其中$K(x)$为R^d上的概率密度，$h_n>0$，$h_n\to 0$ $(n\to\infty)$．这种估计称为核估计．

假定 $(X, Y)\sim F(x, y)$，引入记号：$w(x)\hat{=}\int_{R^1} yF(x, dy)$，$g(x)\hat{=}\int_{R^1} F(x, dy)$，

又令：$w_n(x)=\frac{1}{nh_n^d}\sum_{i=1}^{n} Y_i K\left(\frac{x-X_i}{h_n}\right)$，$g_n(x)=\frac{1}{nh_n^d}\sum_{i=1}^{n} K\left(\frac{x-X_i}{h_n}\right)$，它们分别是$w(x)$和$g(x)$的估计．则有$m(x)=\frac{w(x)}{g(x)}$，$m_n(x)=\frac{w_n(x)}{g_n(x)}$，(约定$\frac{0}{0}=0$)．

当$d=1$时，E. Schuster and S. yakowitz (1979)[3]证明了如下结果：

定理　假设：i)．$K(u)$为一概率密度，$uK(u)\to 0$ $(|u|\to\infty)$，$\int|u|K(u)du<\infty$，$K^{(j)}(u)$ $j=0,1,\cdots$，N在R^1连续，有界变差．ii)．$w^{(j)}(x)$，$j=0,1,\cdots$，N在R^1可积，有界变差，又：$-\infty<a<b<+\infty$，$(c,d)\supset[a,b]$，$w^{(N)}(x)$ 在 (c,d) 上连续．iii) iii)．$g^{(j)}(x)$，$j=0,1,\cdots$，N在R^1有界，在(c,d)连续，且$\inf_{x\in[a,b]} g(x)>0$．iv)．$\psi(u)$为$K(x)$的特征函数，满足$\int|u|^N|\psi(u)|du<\infty$．v)．$E(Y^2)<\infty$．则：存在常数$c$，使对$\forall\varepsilon>0$，当$n$充分大时，有：

$$P\left\{\sup_{a<x<b}\left|m_n^{(j)}(x)-m^{(j)}(x)\right|\geq\varepsilon\right\}\leq c/(nh_n^{2j+2}\varepsilon^2)\quad.$$

本文的目的是在一般$d\geq 1$的情况下考虑$m(x)$及其导数的估计问题，为得出强收敛的结果，需要对Y的矩加上较强些的条件，其它条件则与〔3〕相当．因此在一定程度上，可以说本文是将〔3〕中弱相合结果改为强相合的一种改进．

令B为含于X的支撑的某个紧集，并引进记号：$a^{(j)}(j_1,\cdots,j_d;x)=\frac{\partial^j a(x)}{\partial x_1^{j_1}\cdots\partial x_d^{j_d}}$，简记

* 1983年1月28日收到.

为$a^{(j)}(x)$，$\|X\|$表示R^d中的范数．我们证明了：

定理Ⅰ．设 i）$K(x)$为R^d上的概率密度，且存在$\rho\in(0,1)$，常数$c<\infty$（以下各次出现现的c可不相同），使对充分小的$\tau>0$，有 $\int_{\{\|x\|>\frac{1}{\tau}\}}K(x)dx\leqslant c\cdot\tau^{\frac{\rho}{1-\rho}}$ （1）

ii）$K(x)$的各混合偏导数$K^{(j)}(x)$，$j=1,2,\cdots,N$在R^d上为有界变差． （2）

iii）$w^{(j)}(x)$，$g^{(j)}(x)$，$j=0,1,\cdots,N$在R^d有界，在开集$B'\supset B$上满足Lipschitz 条件，即存在常数c，使：

$$|w^{(j)}(y)-w^{(j)}(x)|\leqslant c\cdot\|y-x\|,\quad |g^{(j)}(y)-g^{(j)}(x)|\leqslant c\cdot\|y-x\|.$$

$$\inf_{x\in B}g(x)>0. \tag{3}$$

iv）$\psi(u)$为$K(x)$的特征函数，满足$\int_{R^d}|u^N\psi(u)|du<\infty$． （4）

v）$h_n^\rho=o(\varepsilon_n)$，$\varepsilon_n>0$，$\varepsilon_n\to0$，$(n\to\infty)$． （5）

vi）$E|Y|^r<\infty$，$r>2$， （6）

则：$P\{\sup_{x\in B}|m_n^{(j)}(x)-m^{(j)}(x)|\geqslant\varepsilon_n\}\leqslant c/(nh_n^{2(j+d)}\varepsilon_n^2)^{r/2}$，$j=0,1,\cdots,N$（$n$充分大）（7）

适当地选取h_n，有对$o<a<\dfrac{\rho(r-2)}{2r(N+d+\rho)}$．

$$h^a\cdot\{\sup_{x\in B}|m_n^{(N)}(x)-m^{(N)}(x)|\}\to0,\ a\cdot s\cdot\quad(n\to\infty). \tag{8}$$

二、两 个 引 理

引理Ⅰ．设条件（1）成立，$a(x)$在R^d有界，在B'上满足L—条件．则：

$$\sup_{x\in B}\left|\int_{R^d}\frac{1}{h_n^d}K\left(\frac{x-u}{h_n}\right)a(u)du-a(x)\right|\leqslant c\cdot h_n^\rho\text{，对充分大的 }n. \tag{9}$$

证明：$\sup_{x\in B}\left|\int_{R^d}\frac{1}{h_n^d}K\left(\frac{x-u}{h_n}\right)a(u)du-a(u)\right|=\sup_{x\in B}\left|\frac{1}{h_n^d}\int_{R^d}K\left(\frac{u}{h_n}\right)\cdot[a(x-u)-a(x)]du\right|$

$$\leqslant\sup_{x\in B}\left|\int_{\{\|u\|<h_n^\rho\}}\right|+\sup_{x\in B}\left|\int_{\{\|u\|\geqslant h_n^\rho\}}\right|\triangleq I_1+I_2. \tag{10}$$

$I_1\leqslant\sup_{x\in B}c\cdot h_n^\rho\cdot\int_{R^d}\frac{1}{h_n^d}K\left(\frac{u}{h_n}\right)du\leqslant c\cdot h_n^\rho$．设$|a(x)|\leqslant M$，则：当$n$充分大时．

$I_2\leqslant2M\cdot\int_{\{\|u\|>h_n^\rho\}}\frac{1}{h_n^d}K\left(\frac{u}{h_n}\right)du=2M\cdot\int_{\{\|v\|<\frac{1}{h_n^{1-\rho}}\}}K(v)dv\leqslant c\cdot h_n^\rho$，以$I_1,I_2$代入（10）即得（9）成立．

推论Ⅰ：设条件（1），（2），（3）中关于$w(x)$的假定成立，则当n充分大时，

$$\sup_{x\in B}|Ew_n^{(j)}(x)-w^{(j)}(x)|\leqslant c\cdot h_n^\rho,\ j=0,1,\cdots,N, \tag{11}$$

证明：$Ew_n(x)=\frac{1}{h_n^d}E\left\{Y_1K\left(\frac{x-X_1}{h_n}\right)\right\}=\frac{1}{h_n^d}E\left\{E\left[Y_1K\left(\frac{x-X_1}{h_n}\right)\Big|X_1\right]\right\}$

$=\frac{1}{h_n^d}\int_{R^d}K\left(\frac{x-u}{h_n}\right)w(u)du$．由引理1知，当$j=0$时，（11）成立．又

$$Ew_n^{(1)}(1,0,\cdots,0;x)=\frac{1}{h_n^{d+1}}\int_{R^d}w(u)\cdot K^{(1)}(1,0,\cdots,0;\frac{x-u}{h_n})\mathrm{d}u$$

$$=\frac{1}{h_n^d}\cdot w(x-u)\cdot K(\frac{u}{h_n})\Big|_{-\infty}^{+\infty}+\frac{1}{h_n^d}\int_{R^d}K(\frac{u}{h_n})\cdot w^{(1)}(1,0,\cdots,o_j x-u)\mathrm{d}u$$

$$=\frac{1}{h_n^d}\int_{R^d}K(\frac{u}{h_n})\cdot w^{(1)}(1,0,\cdots,o;x-u)\mathrm{d}u\ .$$

由引理1知，对该种偏导数，当n充分大时（11）成立．同理证得 $j=1$ 的其它情况，由条件（2）知当 $|x|\to\infty$时，$K^{(j)}(x)\to 0$，$j=2,3,\cdots,N-1$．于是，对$j=2,3,\cdots,N$时的各种混合偏导数情况，也可同法证得（11）成立．

由（4）式，知$K(x)=\frac{1}{(2\pi)^d}\int_{R^d}e^{-iu'x}\cdot\psi(u)\mathrm{d}u$，故

$w_n(x)=\frac{1}{(2\pi)^d}\int_{R^d}e^{-iu'x}\cdot\psi(h_nu)\cdot\phi_n(u)\mathrm{d}u$， 其中

$$\phi_n(u)=\frac{1}{n}\sum_{j=1}^{n}Y_je^{iu'X_j}\ , \tag{12}$$

$$w_n^{(j)}(x)=w_n^{(j)}(j_1,\cdots,j_d;x)=\frac{1}{(2\pi)^d}\int_{R^d}(-i)^j\cdot(u_1^{j_1}\cdots u_d^{j_d})\cdot e^{-iu'x}\cdot\psi(h_nu)\phi_n(u)\mathrm{d}u$$

引理2：设推论1的条件及（4）、（5）、（6）成立，则当n充分大时，

$$P\{\sup_{x\in B}|w_n^{(j)}(x)-w^{(j)}(x)|\geqslant c/(nh_n^{2(j+d)}\varepsilon_n^2)^{r/2},\ j=0,1,\cdots,N. \tag{13}$$

证：$$J\triangleq\sup_{x\in B}|w_n^{(j)}(x)-w^{(j)}(x)|\leqslant\sup_{x\in B}|w_n^{(j)}(x)-Ew_n^{(j)}(x)|+\sup_{x\in B}|Ew_n^{(j)}(x)-w^{(j)}(x)|$$

$$\triangleq I_1+I_2\ . \tag{14}$$

$$I_1=\sup_{x\in B}|w_n^{(j)}(x)-Ew_n^{(j)}(x)|\leqslant\sup_{x\in B}\frac{1}{(2\pi)^d}\int_{R^d}|u_1^{j_1}\cdots u_d^{j_d}|\cdot|\psi(h_nu)|\cdot|\phi_n(u)$$

$-E\phi_n(u)|\mathrm{d}u$．对任意给定的$(j_1,\cdots,j_d)$，记$\Lambda=\int_{R^d}|u_1^{j_1}\cdots u_d^{j_d}|\cdot|\psi(h_nu)|\mathrm{d}u$，则

$$[(2\pi)^dI_1]^r\leqslant\Lambda^r[\int_{R^d}\frac{|u_1^{j_1}\cdots u_d^{j_d}|\cdot|\psi(h_nu)|}{\Lambda}\cdot|\phi_n(u)-E\phi_n(u)|\mathrm{d}u]^r$$

$$\leqslant\Lambda^r\cdot\int_{R^d}\frac{|u_1^{j_1}\cdots u_d^{j_d}|\cdot|\psi(h_nu)|}{\Lambda}\cdot|\phi_n(u)-E\phi_n(u)|^r\mathrm{d}u$$

$$\leqslant\Lambda^r2^{r-1}\int_{R^d}\frac{|u_1^{j_1}\cdots u_d^{j_d}|\cdot|\psi(h_nu)|}{\Lambda}\cdot|R_e\{\phi_n(u)-E\phi_n(u)|^r\mathrm{d}u+\Lambda^r2^{r-1}$$

$$\int_{R^d}\frac{|u_1^{j_1}\cdots u_d^{j_d}|\cdot|\psi(h_nu)|}{\Lambda}\cdot|I_M\{\phi_n(u)-E\phi_n(u)\}|^r\mathrm{d}u\triangleq I_1'+I_1''\ .$$

令$z_j=Re(Y_je^{iu'x_j}-EY_je^{iu'x_j})$，$j=1,2,\cdots,n$，则$\{z_1\}$独立，$Ez_j=0$．由Marcinkiewicz 不等式；

$$E\left|\frac{1}{n}\sum_{j=1}^{n}z_j\right|^r\leqslant CE\left(\frac{1}{n^2}\sum_{j=1}^{n}z_j^2\right)^{r/2}=\frac{c}{n^{r/2}}E\left(\frac{1}{n}\sum_{j=1}^{n}z_j^2\right)^{r/2}\leqslant\frac{c}{n^{r/2}}E\left(\frac{1}{n}\sum_{j=1}^{n}|z_j|^r\right)$$

$$=\frac{c}{n^{r/2}}E|z_1|^r\leqslant\frac{c}{n^{r/2}}E|Y|^r\leqslant c\cdot\frac{1}{n^{r/2}},\quad 故 E(I_1')\leqslant c\cdot\frac{\Lambda^r}{n^{r/2}},$$

$$\Lambda=\int_{\mathrm{R}^d}|u_1^{j_1}\cdots u_d^{j_d}|\cdot|\psi(h_n u)|\,\mathrm{d}u=\frac{1}{h_n^{j+d}}\int_{\mathrm{R}^d}|v_1^{j_1}\cdots v_d^{j_d}|\cdot|\psi(v)|\,\mathrm{d}v \tag{15}$$

由（4）知存在常数 C，使 $\Lambda^r\leqslant\frac{c}{h_n^{(j+d)r}}$. 代入（15）得：

$$E(I_1')\leqslant c\cdot\frac{1}{n^{r/2}\cdot h_n^{(j+d)r}} \tag{16}$$

同理知 $E(I_1'')$ 也有如（16）的界．故 $E[(2\pi)^r\cdot I_1]^r\leqslant c\cdot\frac{1}{[nh_n^{2(j+d)}]^{r/2}}$，由推论 1 及（5）知当 n 充分大时，$P(J\geqslant\varepsilon_n)\leqslant P\left(I_1\geqslant\frac{\varepsilon_n}{2}\right)=P[(2\pi)^d I_1\geqslant 2^{d-1}\pi^d\varepsilon_n]$

$$\leqslant c\cdot\frac{E[(2\pi)^d I_1]^r}{\varepsilon_n^r}\leqslant c\cdot\frac{1}{(nh_n^{2(j+d)}\varepsilon_n^2)^{r/2}}.$$ （13）得证．

推论 2 在引理 2 中易 w 为 g，并令 $\overline{Y}\equiv 1$，知在定理的条件下，当 n 充分大时，有：

$$P\{\sup_{x\in B}|g_n^{(j)}(x)-g^{(j)}(x)|\geqslant\varepsilon_n\}\leqslant\frac{c}{(nh_n^{2(j+d)}\varepsilon_n^2)^{r/2}}. \tag{17}$$

三、定 理 的 证 明

设 $m^{(N)}(x)=m^{(N)}(i_1,\cdots,i_d;x)=\sum\limits_{\substack{j_h+k_h=i_h\\1\leqslant h\leqslant d}}\frac{i_1!\cdots i_d!}{j_1!\cdots j_d!\,k_1!\cdots k_d!}\cdot\frac{\partial^j[g^{-1}(x)]}{\partial x_1^{j_1}\cdots\partial x_d^{j_d}}\cdot\frac{\partial^{N-j}[w(x)]}{\partial x_1^{k_1}\cdots\partial x_d^{k_d}}$，其中 $i_1+\cdots+i_d=N$，$j_1+\cdots j_d=j$，简记 $\frac{i_1!\cdots i_d!}{j_1!\cdots j_d!\,k_1!\cdots k_d!}\triangleq J_{i,j,k}$.

"Σ"号下的项数记为 Υ.

又设 $m_n^{(N)}(x)=m_n^{(N)}(i_1\cdots,i_d;x)=\sum\limits_{\substack{j_h+k_h=i_h,\\1\leqslant h\leqslant d.}}J_{i,j,k}[g_n^{-1}(x)]^{(j)}(j_1,\cdots,j_d;x)$

$w_n^{(N-j)}(k_1,\cdots,k_d;x)$，

则：$m_n^{(N)}(x)-m^{(N)}(x)=\sum\limits_{\substack{j_h+k_h=i_h,\\1\leqslant h\leqslant d,}}J_{i,j,k}\cdot[w_n^{(j)}\cdot(g_n^{-1})^{(N-j)}-w^{(j)}(g^{-1})^{(N-j)}]$，

故 $P\{\sup\limits_{x\in\mathrm{B}}|m_n^{(N)}-m^{(N)}(x)|\geqslant\varepsilon_n\}\leqslant\sum\limits_{j_h+k_h=i_h,}P\{\sup\limits_{x\in\mathrm{B}}|w_n^{(j)}(g_n^{-1})^{(N-j)}-w^{(j)}(g^{-1})^{(N-j)}|$

$\geqslant \frac{\varepsilon_n}{T \cdot J_{i,j,k}}\}$ (18)

故欲证定理，只要对（18）中每一项估计出其有（7）式所示的界（取（7）式中的 j 为 N）.注意到 $[g^{-1}(x)]^{(N-j)}$ 为由 g^{-1}，g，$g^{(1)}$，…，$g^{(N)}$ 进行＋、－、×运算而得，故若记

$\| h(x)\| = \sup\limits_{x\in B}| h(x)|$，则存在与 $\inf\limits_{x\in B} g(x)$ 有关的常数

$\{ a_j(j_1,\cdots,j_d)\}$，使 $\| (g_n^{-1})^{(N-j)} - (g^{-1})^{(N-j)}\| \leqslant \sum\limits_{j=0}^{N} \sum\limits_{j_1+\cdots+j_d=j} a_j(j_1,\cdots,j_d)\cdot\|$

$\| g_n^{(j)}(j_1,\cdots,j_d;x) - g^{(j)}(j_1,\cdots,j_d;x)\|$，对 n 充分大成立.

令 $B(\varepsilon^*) = \{\| w_n^{(j)}(j_1,\cdots,j_d;x) - w^{(j)}(j_1,\cdots,j_d;x)\| < \varepsilon^*,$

$\| g_n^{(j)}(j_1,\cdots,j_d;x) - g^{(j)}(j_1,\cdots,j_d;x)\| < \varepsilon^*, j=0, 1, \cdots, N, j_1+\cdots+j_d=j\}$.

由（13），(17)，在 n 充分大时，$P(B(\varepsilon^*)^c) \leqslant \frac{C}{(nh_n^{2(j+d)}\varepsilon^{*2})^{r/2}}$. (19)

当 $B(\varepsilon^*)$ 发生时，存在常数 M，使

$I \triangleq \| w_n^{(j)}(j_1,\cdots,j_d;x)[g_n^{-1}]^{(N-j)}(k_1,\cdots,k_d;x) - w^{(j)}(j_1,\cdots,k_d;x)[g^{-1}]^{(N-j}(k_1,\cdots,k_d;x)\|$

$\leqslant M\varepsilon^*$.

设 n 充分大，使 $0 < \varepsilon_n < \frac{1}{2}\inf\limits_{x\in B} g(x)$，取 $\varepsilon_n^* = \frac{\varepsilon_n}{M\cdot T\cdot J_{i,j,k}}$，则：

$\{ I \geqslant \frac{\varepsilon_n}{T\cdot J_{i,j,k}}\} \subset B(\varepsilon_n^*)^c$，由（19）即得所要证的结果，因而（7）式得证.

取 $h_n = n^{-\frac{r-2}{2r(N+d+\rho)}}$，$\forall \varepsilon > 0$，取 $\varepsilon_n = \varepsilon\cdot n^{-a}$，其中 $a\in\left(0, \frac{\rho(r-2)}{2r(N+d+\rho)}\right)$ 则有

$P\{ n^a \sup\limits_{x\in B}| m_n^{(N)}(x) - m^{(N)}(x)| \geqslant \varepsilon\} = P(\sup\limits_{x\in B}| m_n^{(N)}(x) - m^{(N)}(x)| \geqslant \varepsilon_n)$

$\leqslant \frac{C}{(nh_n^{2(N+d)}\varepsilon_n^2)^{r/2}}$, (20)

上式右端分母中 n 的指数为 $\left[1 - 2(N+d)\frac{r-2}{2r(N+d+\rho)} - 2a\right]\cdot\frac{r}{2} >$

$\left[1 - 2(N+d)\frac{r-2}{2r(N+d+\rho)} - \frac{2\rho(r-2)}{2r(N+d+\rho)}\right]\cdot\frac{r}{2} = 1$，故（20）右端为收敛级数的一般项. 由Borel—Cantelli 引理，(8）式得证.

四、X为非随机的情况

设 $d=1$，$x\in[0,1]$ 为非随机量，$0 = x_0 \leqslant x_1 < \cdots < x_n \leqslant x_{n+1} = 1$ 取定，(x_1,Y_1)，…，(x_n,Y_n) 为 (x,Y) 的 $i.i.d.$ 样本，$Y_i \sim Y_i | x_i$，分布密度函数为 $f(x,y)$，$y\in R^1$.

$w(x)=\int_{R^1} yf(x,y)\mathrm{d}y$, Priestley and chao (1972)[4]提出用

$w_n(x)=\frac{1}{h_n}\sum_{i=1}^{n} Y_i\ (x_i-x_{i-1})\cdot K(\frac{x-x_i}{h_n})$ 估计$w(x)$．用〔3〕并结合本文定理1的方法、不难证明如下结果:

定理2：设 i）$K(u)$为R^1上的概率密度，且满足：a)(1)式成立．b) $K^{(j)}(u)$在R^1有界，$j=0, 1, \cdots, N+1$．c) 当$|u|$充分大时，$K^{(j)}(u)\leqslant c\cdot(\frac{1}{|u|})^{j+1+\rho}$. ii)

ii) $w^{(N)}(x)$ 在〔0, 1〕上存在且连续．iii) $\psi(u)$为$K(v)$的特征函数且(4)式成立．

iv) $\int_{R^1}|y|^r f(x,y)\mathrm{d}y$ 对$x\in$〔0,1〕一致有界，$r>2$．v) $h_n=n^{-\frac{r-2}{2r(N+1+\rho)}}$. vi)

vi) $\|\Gamma_n\|=\max\limits_{0\leqslant j\leqslant n}(x_{j+1}-x_j)=0(\frac{1}{n})$．则：对任〔$a,b$〕$\subset$(0, 1)，及

$0<\alpha<\frac{\rho(r-2)}{2r(N+1+\rho)}$，有：$n^{\alpha}\cdot\{\sup\limits_{a<x<b}|w_n^{(N)}(x)-w^{(N)}(x)|\}\to 0$, a.s. $(n\to\infty)$.

在结束本文时，作者特对陈希孺教授的指导，对赵林城博士和白志东博士的热情帮助深表谢意。

参 考 文 献

〔1〕Watson, G. S. (1964), Smooth regression analysis. Sankhyā Ser. A26 359—372。

〔2〕Nadaraya, E. A. (1964), On estimating regression. Theor. Probability Appl. 9 141—142.

〔3〕E. Schuster and S. Yakowitz (1979), Contributions to the theory of non nonparametric regression, with application to system identification. A Ann. Stat. Vol. 7. No. 1, 139—149.

〔4〕Priestley, M. B. and Chao, M. T. (1972), Nonparametric function fitting. J. Roy. Statist. Soc. Ser. B34 385—392.