Chin. Ann. of Math.
11B: 4 (1990), 536—545.

CONVERGENCE RATE OF MULTIVARIATE K-NEAREST NEIGHBOR DENSITY ESTIMATES**

YANG ZHENHAI (杨振海)*

Abstract

Let $\mathscr{F}$ be the collection of m-times continuously differentiable probability densities f on R^d such that $|D^\alpha f(x_1)-D^\alpha f(x_2)|\leqslant M\|x_1-x_2\|^\beta$ for $x_1, x_2\in R^d$, $[\alpha]=m$, where D^α denotes the differential operator defined by $D^\alpha=\frac{\partial^{[\alpha]}}{\partial x_1^{\alpha_1}\cdots\partial x_d^{\alpha_d}}$. Under rather weak conditions on $K(x)$, the necessary and sufficient conditions for sup $|\hat{f}_n(x)-f(x)|=O\left(\left(\frac{\log n}{n}\right)^{\lambda/(d+3\lambda)}\right)$, $\lambda=m+\beta$, $f\in\mathscr{F}$ are that $\int x^\alpha K(x\mathrm{i})dx=0$ for $0<[\alpha]\leqslant m$. Finally the convergence rate at a point is given.

§1. Introduction

Let $X_1, X_2, \cdots X_n$ be a sample from a one dimensional population with probability density function f and distribution function F. There are many papers discussing how to estimate p. d. f. $f(x)$ based on $X_1, \cdots, X_n$. In 1965, Loftsgarden and Quesenberry[1] proposed the estimator, called nearest neighbor estimation, defined by

$$f_n^*(x)=k_n/2na_n(x), \tag{1}$$

where k_n is an integer chosen in advance, depending on n, and $a_n(x)$ is the smallest number such that the number of X_i's lying in $[x-a, x+a]$ among $X_1, \cdots, X_n$ is equal to or greater than k_n. Many authors have studied the properties of Nearest Neighbor estimation. In recentyears, many authors have been interested in the problem of convergence rate of this estimator. In 1981, Chen Riru[2] proved that

If f satisfies Lipschitz condition on R_1 and k_n is chosen so that $n^{2/3}(\log\log n)^{1/2}/k_n$ has positive finite limit, then

$$\sup_{x\in R^1}|f_n^*(x)-f(x)|=O(n^{-1/6}(\log\log n)^{1/6}) \text{ a.s.}$$

and pointedoutthat the rate of convergence of $\sup\limits_x|f_n^*(x)-f(x)|$ to zero cannot reach

Manuscript receive December 7, 1987. Revised January 11, 1989.

* Department of Applied Mathematics, Beijing Polytechnic University, Beijing, China.

** The project supported by National Natural Science Foundation of China.

$O(n^{-1/4}(\log\log n)^{1/4})$.

Yang Zhenhai[5] has improved the above result to

$$\sup_{x\in R^1}|f_n^*(x)-f(x)|=O(n^{-1/4}(\log n)^{1/4}) \quad \text{a.s.}$$

A little later, Yang Zhenhai and Chao Linchen[8] proved that

$$\sup_x|f_n^*(x)-f(x)|=O\left(\left(\frac{\log n}{n}\right)^{\lambda/(1+3\lambda)}\right) \quad \text{a.s.,} \tag{2}$$

where $\lambda=\delta$ if $f(x)$ satisfies the δth order Lipschitz condition and $\lambda=1+\delta$ if $f'(x)$ satisfies the δth order Lipschitz condition.

Let

$$K(x)=\frac{1}{2}I_{[-1,1]}(x),$$

where I_A is the indicator function of A. Then (1) can be written as

$$f_n^*(x)=\frac{1}{na_n(x)}\sum_{i=1}^n K\left(\frac{x-X_i}{a_n(x)}\right)=\frac{1}{a_n(x)}\int K\left(\frac{x-y}{a_n(x)}\right)dF_n(y). \tag{3}$$

The form of formula (3) is similar to one of the kernel density estimates. For a sample from a d-dimentional population, (3) should become

$$\hat{f}_n(x)=\frac{1}{na_n^d(x)}\sum_{i=1}^n K\left(\frac{x-X_i}{a_n(x)}\right)=\frac{1}{a_n^d(x)}\int K\left(\frac{x-y}{a_n(x)}\right)dF_n(y), \tag{4}$$

where $F_n(y)$ is the empirical distribution function of $X_1, \cdots, X_n$, and $a_n(x)$ is the smallest number a such that sphere $S(x, a)=\{y: \|x-y\|\leqslant a, y\in R^d\}$ at least contains $k_n X_i$'s among $X_2, \cdots, X_n$. Mack and Rosenblatt[4] have studied the properties of the estimate defined by (4). They gave the approximate expansions of mean and variance of the estimate. In this paper, we shall study the convergence rate of the nearest neighbor estimate for d dimensional p. d. f. $f(x)$, which is called Multivariate K-nearest neighbor density estimate by Mack and Rosonblatt.

§ 2. Assumptions and Main Result

First, we make some assumptions on kernel function $K(x)$. Through this paper, we suppose $K(x)$ is a bounded continuous function.

Set

$$Q_t=\{x: K(x)\geqslant t\} \text{ for } t\geqslant 0,\ Q_t=\{x: K(x)\leqslant t\} \text{ for } t<0,$$
$$\boldsymbol{Q}_t=\{x: K(x)=t\}.$$

Assumption 1A:

(1) $K(x)$ is a p. d. f. on R^d.

(2) For any $t\in(0, \overline{K}_0)$, $\overline{K}_0=\max\{K(x): x\in R^d\}$, Q_t is a simple closed surface and

$$Q_t\subset Q_{t'} \text{ if } t>t' \text{ and } \boldsymbol{Q}_t\cap\boldsymbol{Q}_{t'}=\varnothing \text{ if } t\neq t'.$$

Assumption 1B:

(1) $\int_{R^d} K(x)dx=1$.

(2) For any $t\in(\underline{K}_0, \overline{K}_0)$, $\underline{K}_0=\min\{K(x),\ x\in k^d\}$, $\overline{K}_0=\max\{K(x):x\in R^d\}$ Q_t is a union of, at most, l disjoint simple closed surfaces, where l is a fixed integr, and

$$Q_t\supsetneqq Q_{t'} \text{ if } |t|>|t'| \text{ and } tt'>0 \text{ and } Q_t\cap Q_{t'}=\varnothing \text{ if } t\neq t'.$$

Now we make some assumptions on p. d. f. $f(x)$. First we introducethe notation. Let $\alpha=(\alpha_1, \cdots, \alpha_d)$ denote a d-tuple of nonnegative integers and set $[\alpha]=\alpha_1+\cdots+\alpha_d$ and $\alpha!=\alpha_1!\alpha_2!\cdots\alpha_d!$. For $x=(x_1, \cdots, x_d)'\in R^d$, set $\|x\|=(\sum_{i=1}^{d}x_i^2)^{1/2}$ and $x^\alpha=(x_1^{\alpha_1}x_2^{\alpha_2}\cdots x_d^{\alpha_d})$. Let D^α denote the differential operator defined by

$$D^\alpha=\frac{\partial^{[\alpha]}}{\partial x_1^{\alpha_1}\cdots,\ \partial x_d^{\alpha_d}}.$$

Let M, β be real constants such that $M>0$ and $0<\beta\leqslant 1$. Let $\mathscr{F}$ be a collection of m-times continuously differentiable probability densities f on R^d such that

$$|D^\alpha f(x_1)-D^\alpha f(x_2)|\leqslant M\|x_1-x_2\|^\beta \text{ for } x_1, x_2\in R^d\ [\alpha]=m, \tag{5}$$

and $D^\alpha f(x)$ is bounded for $[\alpha]\leqslant m$.

Set $\lambda=m+\beta$.

Our main result is the following

Theorem 1. *Suppose that Assumption 1A or 1B holds and $K(x)$ satisfies*

$$K(x)=0 \quad \textit{for } \|x\|>1 \tag{6}$$

and

$$\int x^\alpha K(x)dx=0 \quad \textit{for } [\alpha']\leqslant m (\textit{If Assumption 1A holds then } m<2). \tag{7}$$

Then

$$\sup_{x\in R^d}|\hat{f}_n(x)-f(x)|=O\left(\left(\frac{\log n}{n}\right)^{\lambda/(d+3\lambda)}\right) \quad \text{a.s.}$$

holdfor $f\in\mathscr{F}$.

Theorem 2. (The rate of convergence at a point). *Suppose that Assumption 1A or 1B holds and $f(x)\in\mathscr{F}$ satisfies at 0 that*

$$\begin{gathered} D^{(\alpha)}f(0)=0 \quad \textit{for } [\alpha]<m, \\ D^{(\alpha)}f(0)\neq 0 \quad \textit{at least for one } \alpha \textit{ with } [\alpha]=m, \\ f(0)\neq 0. \end{gathered} \tag{8}$$

If $\lim_{n\to\infty} k_n \Big\backslash \left(\frac{\log n}{n}\right)^{d/(d+2\lambda)} = >0$, then we have

$$|f(0)-\hat{f}_n(0)|=O\left(\left(\frac{\log n}{n}\right)^{\lambda/(d+2\lambda)}\right) \quad \text{a.s.} \tag{9}$$

Theorem 3. *Suppose that Assumption 1A or 1B holds, and for any $f\in\mathscr{F}$*

$$\sup_{x\in R^d}|\hat{f}_n(x)-f(x)|=O\left(\left(\frac{\log n}{n}\right)^{\lambda/(d+3\lambda)}\right) \quad \text{a.s.}$$

holds. Then (7) holds.

§ 3. The Proof of Theorems

Let B be a set in R^d. For any $h>0$ and $x\in R^d$, we define $S(B, x, h)$ as

$$S(B, x, h)=\left\{y: \frac{y-x}{h}\in B\right\}.$$

Let $\mathscr{S}$ be the class of all sets $S(Q_t, x, h)$, $x\in R^d$ $t\in R^1$ and all spheres centered at $x\in R^d$. For any n point set $\{x_1, x_2, \cdots, x_n\}$ in R^d, set

$$\mathscr{S}(x_1, x_2, \cdots, x_n)=\{B: B=\|x_1, \cdots, x_n\}\cap A,\ A\in\mathscr{S}\}.$$

Let $\#\mathscr{S}$ denote the number of different sets in $\mathscr{S}$ and put

$$s(n)=\max\{\#\mathscr{S}(x_1, x_2, \cdots, x_n): x_i\in R^d,\ 1\leqslant i\leqslant n\}.$$

Based on the fact that a simple closed surfase is homomorphic to a sphere and the Function-Counting Theorem for arbitrary surface [9], we get the following Lemma

Lemma 1. *If Assumption 1A holds, then we have*

$$s(n)\leqslant(n)^{2d}.$$

Further we also have

Lemma 1′. *If Assumption 1B holds, then we have.*

$$s(n)\leqslant l(n)^{2d},$$

where l is the integer defined in Assumption 1b.

For any set B in R^d, let

$$D(B)=\sup_{x,y\in B}\|x-y\|.$$

By Lemma 1 and Lemma 1—2 of Devroye and Wagner[6], the following Lemma follows.

Lemma 2. *Let μ_n be the empirical measure for $X_1, \cdots, X_n$ and μ be the measure on the Borel sets of R^d which corresponds to f. Let r be a positive rerl number such that*

$$\sup_{x\in R^d}\mu(S_{(x,r)})\leqslant b<1/4.$$

Then

$$P\{\sup_{B\in S_r}|\mu_n(B)-\mu(B)|\geqslant\varepsilon\}\leqslant 4(2n)^{2ld}e^{-n\varepsilon^2/(64b+\varepsilon)}+8ne^{-nb/10}$$

holds for all $n\geqslant\max\left(\frac{1}{b}, 8b/\varepsilon^2\right)$ and $\varepsilon>0$, where S_r is the class of all sets B in S with $D(B)\leqslant r$, and $S_{(x,r)}$ is the sphere centered at x with radius r.

Lemma 3. *For any $\lambda>0$ and n large enough, we hove*

$$\sup\{\mu_n(B)-\mu(B)|: B\in S_r,\ \text{and}\ \sup_{x\in k}\mu(S_{(x,r)})\leqslant 3k/n\}\leqslant A\left(\frac{\log n}{n}\right)^{1/2}\sqrt{\frac{k}{n}}\quad\text{a.s.}$$

where $k=k_n$ is an integer depending on n with $1\leqslant k\leqslant n$, and A is a constant.
The proof of this Lemma is similar to that of Lemma 2 of [7].

For $x\in R^d$, function $\tilde{a}_n(x)$ is defined as the smallest number a' which satisfies

$$\frac{k_n}{n}=\int_{S(x,a')} f(u)du. \tag{10}$$

In the same way, we define $\bar{a}_n(x)$ and $\underline{a}_n(x)$ as the smallest numbers, respectively, such that

$$(1+l_n)\frac{k_n}{n}=\int_{S(x,\bar{a}_n(x))} f(u)du$$

and

$$(1-l_n)\frac{k_n}{n}=\int_{S(x,\underline{a}_n(x))} f(u)du,$$

where l_n satisfies

$$\lim_{n\to\infty}\frac{l_n\left(\frac{\log n}{n}\right)^{\lambda/(d+3\lambda)}}{(k_n/n)}=+\infty \quad\text{and}\quad \lim_{n\to\infty} l_n=0. \tag{11}$$

Let

$$B=\left\{x: f(x)\geqslant L\left(\frac{\log n}{n}\right)^{\lambda/(d+3\lambda)},\ x\in R^d\right\}, \tag{12}$$

where L is a constant to be determined (From now on we shall omit subscript n if it is possible, for example, write k instead of k_n). It is obvious that if $x\in B$, then we have

$$\frac{1}{2M_0C_0}\,\frac{k}{n}\leqslant \underline{a}_n^d(x)\leqslant \tilde{a}_n^d(x)\leqslant \bar{a}_n^d(x)\leqslant \frac{3}{2LC_0}\left(\frac{k}{n}\right)\left(\frac{\log n}{n}\right)^{-\lambda/(d+3\lambda)}, \tag{13}$$

where

$$M_0=\sup_{x\in R^d} f(x),\ C_0=\int_{S(0,1)} dx.$$

If we put $k=k_n=n^{2\lambda/(d+3\lambda)}(\log n)^{(d+\lambda)/(d+3\lambda)}$, then (13) becomes

$$\frac{1}{2M_0C_0}\left(\frac{\log n}{n}\right)^{(d+\lambda)/(d+3\lambda)}\leqslant \underline{a}_n^d(x)\leqslant \bar{a}_n^d(x)\leqslant \frac{3}{2LC_0}\left(\frac{\log n}{n}\right)^{d/(d+3\lambda)}. \tag{14}$$

Lemma 4. *For $f\in\mathscr{F}$ and any $a>0$,* we have

$$\left|\int_{S(x,a)} f(u)du-\psi(x,a)\right|\leqslant R_0a^{d+\lambda},$$

where R_0 is a constant depending on both M and m.

$$\psi(x,a)=C_0a^d f(x)+a^d\phi(x,a)$$

and

$$\phi(x,a)=\begin{cases}\sum_{0<[\alpha]<m} C_\alpha a^{[\alpha]}\frac{D^\alpha f}{\alpha!} & \text{if } \lambda>2,\\ 0 & \text{if } \lambda\leqslant 2,\end{cases}$$

$$C_\alpha=\int_{S(0,1)} u^\alpha du.$$

Proof By Taylor expanssion we have

$$f(u)=\sum_{0<[\alpha]<m}\frac{D^\alpha f(x)}{\alpha!}(u-x)^\alpha+\sum_{[\alpha]=m}\frac{(D^\alpha f(x+\theta'(u-x))-D^\alpha f(x))}{\alpha!}(u-x)^\alpha.$$

Then the lemma follows from (5) and the above formula.

Lemma 5. *For any $\varepsilon>0$, and $\varepsilon<a$ we have*

$$\left|\int_{S(x,(1+\varepsilon)a)\setminus S(x,a)} f(u)du-\varepsilon\psi(x,a)\right|\leqslant A_2\varepsilon^2a^{d+1}+A_4\varepsilon a^{\lambda+d}.$$

Proof The Lemma can be proved by Taylor expansion and the following relation

$$\int_{S(0,1+\varepsilon)\setminus S(0,1)} u^\alpha du=\varepsilon C_\alpha+\varepsilon^2\sum_{i=2}^{[\alpha]}[\alpha]i\varepsilon^{i-2}C_\alpha.$$

Lemma 6

$$(1-A_1h_n)a_n(x)\leqslant a_n(x)\leqslant(1+A_1h_n)\bar a_n(x)\quad \text{a.s.}$$

holds for all $x\in B$ and n large enough, where $h_n=\left(\frac{\log n}{n}\right)^{\lambda/(d+3\lambda)}$ *and A_1 is constant.*

Proof By Lemma 5 and (13),

$$\begin{aligned}\mu(S_{(x,(1+A_1h_n)\bar a_n(x))})&=\frac{k}{n}(1+ln)+\int_{S(x,(1+A_1h_n)\bar a_n(x))\setminus S(x,\bar a_n(x))}f(u)du\\
&\leqslant 2\frac{k}{n}+C_0A_1h_n\bar a_n^d(x)f(x)+A_2'A_1h_n\bar a_n(x)\bar a_n^d(x)\\
&\quad+A_2A_1^2h_n^2\bar a_n^{d+1}(x)+A_4A_1h_n\bar a_n^{\lambda+d}(x)\\
&\leqslant\frac{k}{n}\left(2+\frac{3C_0A_1M_0}{2LC_0}+\frac{3A_1A_2'}{2LC_0}\bar a_n(x)+\frac{3A_1^2A_2}{2LC_0}\bar a_n(x)h_n+\frac{3A_4A_1}{2LC_0}\bar a_n^\lambda(x)\right)\\
&\leqslant 3\frac{k}{n}. \qquad (15)\end{aligned}$$

fi we choose L and A_1 such that $L\geqslant 2A_1M_0$ and n is large enough. So we can apply Lemma 3 to this case.

$$\begin{aligned}u_n(S_{x,(1+A_1h_n)\bar a_n(x))}&\geqslant\mu(S_{(x,(1+A_1h_n)\bar a_n(x)}-A\left(\frac{k}{n}\left(\frac{\log n}{n}\right)\right)^{1/2}\\
&\geqslant\frac{k}{n}+l_n\frac{k}{n}+A_1h_n\bar a_n^d(x)-A_2'A_1h_n\bar a_n^{d+1}(x)-A_2A_1^2h_n^2\bar a_n^{d+1}(x)\\
&\quad-A_4A_1h_n\bar a_n^{\lambda+d}(x)-A\left(\frac{k}{n}\left(\frac{\log n}{n}\right)\right)^{1/2}\\
&\geqslant\frac{k}{n}+\frac{k}{n}\left(l_n-A_5\bar a_n(x)-A_6\bar a_n(x)h_n-A_7\bar a_n^\lambda(x)-\left(\frac{\log n}{n}\right)^{\lambda/(d+3\lambda)}\right)\\
&\geqslant\frac{k}{n}+\Delta_n.\end{aligned}$$

It is easy to see $\Delta_n\geqslant 0$, so the right hand side$\geqslant k/n$. Hence we have

$$a_n(x)\leqslant(1+A_1h_n)\bar a_n(x)\leqslant 2a_n(x)\quad\text{for sufficiently large } n$$

The rest of the lemma can be proved via the same way.

Proof of Theorem 1 Set

$$f_n(x)=\frac{1}{a_n^d(x)}\int K\left(\frac{x-y}{a_n(x)}\right)f(y)dy. \qquad (16)$$

Then

$$f_n(x)-f(x)=\int_{R^d}K(u)[f(x-a_n(x)u)-f(x)]du$$

$$= \sum_{0<[\alpha]\leq m} \frac{D^\alpha f(x)}{\alpha!}(-a_n(x))^{[\alpha]}\int_{R^d} u^\alpha K(u)du$$
$$+\sum_{[\alpha]=m} a_n^m(x)\int_{R^d} u^m k(u)\frac{D^\alpha f(x+\theta a_n(x)u)-D^\alpha f(x)}{\alpha!}du, \quad \theta=\theta(u),\ |\theta|\leq 1.$$

Hence, if $x\in B$,

$$|f_n(x)-f(x)|\leq \sum_{[\alpha]=m}\frac{a_n^m(x)}{\alpha!}\int u^m a_n^\beta(x)\|\mu\|^\beta K(u)du\leq Aa_u^{m+\beta}(x)$$
$$=Aa_n^\lambda(x)\leq A\left(\frac{\log n}{n}\right)^{\lambda/(d+3\lambda)}.$$

For any n, choose N_n so large that

$$N_n>(n/\log n)^{\lambda/(d+3\lambda)}\left(\frac{n}{k}\right) \tag{17}$$

and

$$\left|\sum_{i\in I_n}\frac{iK_0}{N_n}L_b(Q_i^*)-1\right|\leq\left(\frac{\log n}{n}\right)^{\lambda/(d+3\lambda)}, \tag{18}$$

where

$$Q_i^*=Q_i-Q_{i+\mathrm{sign}(i)},\ K_0=\overline{K}_0-\overline{K}_0,$$

$$t_i=\frac{i}{N_n}K_0,\ i\in I_n=\{[N_n\underline{K}_0/K_0],\ [N_n\underline{K}_0/K_0],\ \cdots,\ -1,\ -0,\ 0,\ 1,\ \cdots,\ [N_n\overline{K}_0/K_0]\}$$

and

$$Q_i=Q_{t_i}$$

and L_b denotes Lebesgue measu e on R^d; we can always find such an N_n because

$$\lim_{N_n\to\infty}\sum_{i\in I_n}^{N_n}\frac{iK_0}{N_n}L_b(Q_i^*)=\int K(x)dx=1.$$

We also have

$$\sum_{i\in I_n}\frac{iK_0}{N_n}L_b(Q_i^*)=\sum_{i\in I_n}\frac{K_0}{N}L_b(Q_i)\mathrm{sign}(i),\ (\mathrm{sign}(0)=1).$$

Set

$$K_n(x)=\sum_{i\in I_n}\frac{K_0L_b(Q_i)}{N_n}\cdot K_{ni}(x),\ K_{ni}(x)=\frac{1}{L_b(Q_i)}I(Q)\mathrm{sign}(i)$$

and

$$\hat{f}_n^*(x)=\frac{1}{na_n^d(x)}\sum_{i=1}^n K_n\left(\frac{x-X_i}{a_n(x)}\right).$$

It is easy to see that

$$|\hat{f}_n(x)-\hat{f}_n^*(x)|\leq K_0/(N_na_n^d(x))\leq K_0\frac{1}{N_n}\frac{2M_0C_0n}{k}\leq A\left(\frac{\log n}{n}\right)^{\lambda/(d+3\lambda)}$$

and

$$|f_n(x)-f_n^*(x)|=-\frac{1}{a_n^d(x)}\left|\int\left[K\left(\frac{x-y}{a_n(x)}\right)-K_n\left(\frac{x-y}{a_n(x)}\right)\right]f(y)dy\right|$$
$$\leq\frac{1}{a_n^d(x)}\frac{K_0}{N_n}\int f(y)dy\leq\frac{1}{a_n^d(x)N_n}\leq A_2\left(\frac{\log n}{n}\right)^{\lambda/(d+3\lambda)},$$

where

$$f_n^*(x)=\frac{1}{a_n^d(x)}\int K_n\left(\frac{x-y}{a_n(x)}\right)f(y)dy.$$

We also have

$$|\hat{f}_n^*(x)-f_n^*(x)|=\left|\frac{1}{a_n^d(x)}\int K_n\left(\frac{x-y}{a_n(x)}\right)d(F_n(y)-F(y))\right|$$
$$\leqslant\frac{1}{a_n^d(x)}\sum_{i\in I_n}\frac{K_0}{N_n}|\mu_n(S(Q_i,x,a_n(x)))-\mu(S(Q_i,x,a_n(x)))|$$
$$\leqslant\frac{1}{a_n^d}\sum_{i\in I_n}\frac{K_0}{N_n}\left(\frac{\log n}{n}\right)^{\frac{d+2\lambda}{d+3\lambda}}\leqslant A_3\left(\frac{\log n}{n}\right)^{\lambda/(d+3\lambda)}\quad\text{a.s.}$$

Hence

$$\sup_{x\in B}|\hat{f}_n(x)-f(x)|\leqslant A_4\left(\frac{\log n}{n}\right)^{\lambda/(d+3\lambda)}\quad\text{a.s.}\tag{19}$$

Now we prove

$$\sup_{x\in B^c}|\hat{f}_n(x)-f(x)|\leqslant A_5\left(\frac{\log n}{n}\right)^{\lambda/(d+3\lambda)}\quad\text{a.s.}$$

Set

$$B_1=\left\{x\colon D(x,B)\geqslant q_0\left(\frac{\log n}{n}\right)^{1/(d+\lambda 3)}\right\}$$

and

$$B_2=B^cB_1^c,$$

where q_0 is a positive number such that $LC_0q_0<1$. If $x\in B_1$, then

$$\mu\left(S\left(x,\ q_0\left(\frac{\log n}{n}\right)^{1/(d+3\lambda)}\right)\right)\leqslant LC_0q_0\left(\frac{\log n}{n}\right)^{(d+\lambda)/(d+3\lambda)}=C_0Lq_0\frac{k}{n}.$$

By Lemma 3,

$$\mu_n\left(S\left(x,\ q_0\frac{\log n}{n}\right)^{1/(d+3\lambda)}\right)\leqslant\mu\left(S\left(x,\ q_0\left(\frac{\log n}{n}\right)^{1/(d+3\lambda)}\right)\right)+A\left(\frac{\log n}{n}\right)^{(d+2\lambda)/(d+3\lambda)}$$
$$<\frac{k}{n}\quad(n\text{ large enough}).$$

Hence

$$a_n(x)\geqslant q_0\left(\frac{\log n}{n}\right)^{1/(d+3\lambda)}\quad\text{a.s. for }x\in B_1.$$

Therefore

$$\sup_{x\in B_1}|\hat{f}_n(x)-f(x)|\leqslant\sup_{x\in B_1}\hat{f}_n(x)+\sup_{x\in B_1}f(x)\leqslant\frac{K}{na_n^d(x)}\cdot K_0+L\left(\frac{\log n}{n}\right)^{\lambda/(d+3\lambda)}$$
$$\leqslant A\left(\frac{\log n}{n}\right)^{\lambda/(d+3\lambda)}\quad\text{a. s.,}\tag{20}$$

because

$$\hat{f}_n(x)=\frac{1}{a_n^d(x)}\int_{S(x,a_n(x))}K\left(\frac{x-y}{a_n(x)}\right)dF_n(x)\leqslant K_0\frac{1}{a_n^d(x)}\frac{k}{n}.$$

If $x\in B_2$, there is a point $y\in B$, such that $D(x,y)\leqslant q_0\left(\frac{\log n}{n}\right)^{1/(d+3\lambda)}$, put

$$p=q_0\left(\frac{\log n}{n}\right)^{1/(d+2\lambda)}+\bar{a}_n(y),$$

then

$$\int_{S(x,p)}f(\ddot{u})du=\mu(S(x,p))\geqslant\mu(S_{(y,\bar{a}_n(y))})=k(1+l_n)/n.$$

Hence

$$\bar{a}_n(x)\leqslant p\leqslant A_5\left(\frac{\log n}{n}\right)^{1/(d+3\lambda)}.$$

It follows that (14) holds if we replace the constant by A_5. Therefore

$$\sup_{x\in B_2}|\hat{f}_n(x)-f(x)|\leqslant A_6\left(\frac{\log n}{n}\right)^{\lambda/(d+3\lambda)}\quad\text{a.s.}$$

can be proved in the same way as that for $x\in B$. This completees the proof.

Proof of Theorem 2 The outline of the proofis the sa me as that of Theorem 1. We only point out the key point for the proof of Theorem 3.

First: for a fixed point, (13) becomes

$$A_{11}\left(\frac{k}{n}\right)<\tilde{a}_n^d(0)<A_{10}\left(\frac{k}{n}\right),\tag{21}$$

where $0<A_{11}<A_{10}$, and maybe depends on point 0 (precisely on $f(0)$).

Second: Let $f_n(x)$ be defined by (16). Then we have under condition (8)

$$|f_n(0)-f(0)|\leqslant A_{12}\left(\frac{k}{n}\right)^{\lambda}\quad\text{a.s.,}$$

where $\lambda=m+\beta$.

Third: From the proof of Theorem 1, it follows that

$$|f_n(0)-f_n(0)|\leqslant\max\left(\left(\frac{\log n}{n}\right)^{1/2}\sqrt{\frac{n}{k}},\left(\frac{k}{n}\right)^{\lambda/d}\right).$$

Hence $$|f_n(0)-\hat{f}_n(0)|=0\left(\max\left(\left(\frac{\log n}{n}\right)^{1/2}\sqrt{\frac{n}{y}},\left(\frac{k}{n}\right)^{\lambda/d}\right)\right),$$

and (9) follows.

Proof of Theorem 3 Take $f\in F$such that

$$f(0)>0,\ \text{and}\ \frac{\partial f(0)}{\partial x^i}\neq 0,\ \frac{\partial f(0)}{\partial x^j}=0,\quad j\neq i,$$

where x^i is the ith component of $x\in R^d$. Then we have

$$|f_n(0)-f(0)|\leqslant\max|f_n(x)-\hat{f}(x)|\leqslant A\left(\frac{\log n}{n}\right)^{\lambda/(d+3\lambda)},\tag{22}$$

but

$$\begin{aligned}|f_n(0)-f(0)|&=\left|\frac{1}{a_n^d(0)}\int K\left(\frac{-y}{a_n(0)}\right)f(y)dy-f(0)\right|\\&=\left|\frac{1}{a_n^d(0)}\int K\left(-\frac{y}{a_n(0)}\right)\left[f(0)+\sum_{[\alpha]=1}\frac{D^\alpha f(0)}{\alpha!}y^\alpha dy\right]-f(0)\right|+O(a_n^2(0))\\&=\left|\int K(y)y_i\,dy\right|a_n(0)+O(a_n^2(0)).\end{aligned}\tag{23}$$

From (22) and (23), it follows that

$$\int y^\alpha K(y)dy=0\quad\text{for }[\alpha]=1.$$

By induction and taking particular f, we can prove

$$\int y^\alpha K(y)dy=0\quad\text{for}[\alpha]\leqslant m.$$

References

[1] Loftsgarden, D. O. & Quesenberry, C. P., A nonparametric estimate of a multivariate density function, *Ann. Math. Statis*, **36** (1965), 1049.

[2] Chen Xiru, The rate of convergence of nearest neighbor estimate. *Scienta Sinica*, **12** (1981), 1419.

[3] Stone, C. J., *Optimal uniform rate of convergence for nonparametric estimators of a density function or its derivatives*, Recent Advances in Statistics: Papers Presented in Honor of Herman Chernoff's Sixtieth Birthday, Rizvi, Rustagi, & Siegmund (eds.), Academic Press, New York (1983).

[4] Mack, Y. P. & Rosenblatt, Multivariate k-nearest neighbor density estimate, *J. Multivariate Analysis*, **9** (1979), 1—15.

[5] Yang Zhenhai, The uniform convergence rate of nearest neighbor estimate, *Koxue Tengbao* (*Science Report*), *China*, (1983), 1355—1358

[6] Devroye, L. P. & Wagner, T. J., The strong uniform consistency of kernel density estimation, *Multivariate Analysis V.*, (1980), 55—77.

[7] Yang Zhenhai, The uniform convergence rate of kernel density estimate, *Chin. Ann. of Math.*, **6B**: 3 (1985), 335—343.

[8] Yang Zhenhai & Chao Linchen, Uniform strong convergence rate of nearest neighbor density estimation. *Chin. Ann. of Math.*, **5B**: 3 (1984), 325—332.

[9] Thomas, M. Cover, Geometrical and statistical properties of systems of linear Inequalitics with applicrtion in pattern recognition *IEEE, Transaction on Electronic Computers*, **10** (1965), 321—334.