ELSEVIER

Available online at www.sciencedirect.com

SciVerse ScienceDirect

Procedia Engineering 31 (2012) 1012 – 1018

Procedia Engineering

www.elsevier.com/locate/procedia

International Conference on Advances in Computational Modeling and Simulation

The Convergence Rate of Multidimensional Density Kernel Estimation with Bootstrap

Dewang Li[a*], Meilan Qiu[a]

[a]*Department of Mathematics , Hechi University, Yizhou Guangxi, 546300, China.*

Abstract

There have important applications of density kernel estimation in statistics. In certain conditions, we obtain the convergence rate of multidimensional density kernel estimation by exploiting Bootstrap.

Keyword: Bootstrap, multidimensional probability density, kernel estimation, convergence rate.

---

1. Introduction

In 1979, Efron proposed a re-sampling procedure, called Bootstrap [1]. By using this method, we re-sample the empirical distribution of samples at random, and get Bootstrap sub-samples. Then, we re-estimate the amount of statistics. In this paper, basing on references [2-3], we study the Bootstrap convergence rate of multidimensional density kernel estimation. Let X be a d -dimensional random variable, $X_1,\cdots,X_n$ be the sample of X . Suppose that $F_n(X)$ is the empirical distribution function based on following samples $X_1,\cdots,X_n$ which with observed values $x_1,\cdots,x_n$. $X_1^*,\cdots,X_n^*$ are independent identically distributed samples deriving from $F_n(X)$. The kernel estimates $f(X)$ about

* Corresponding author. Tel.:+ 86-13558384315.
E-mail address: LDWLDW1976@126.COM.

1877-7058 © 2011 Published by Elsevier Ltd. Open access under CC BY-NC-ND license.
doi:10.1016/j.proeng.2012.01.1135

probability density function of X defined as $f_n(X) = \frac{1}{nh^d \det(S)^{1/2}} \sum_{i=1}^{n} K[\frac{(X - X_i)^T S^{-1} (X - X_i)}{h^2}]$. Where $X = (X_1, \cdots, X_d)^T$, $X_i = (X_{i1}, \cdots X_{id})^T$ $(i = 1, \cdots, n)$, $K(\cdot)$ is a given probability density kernel function, h is a bandwidth coefficient , n is the sample capability, S denotes the symmetric sample covariance matrix for $d \times d$ -dimension. $f_n^*(X) = \frac{1}{nh^d \det(S)^{1/2}} \sum_{i=1}^{n} K[\frac{(X - X_i^*)^T S^{-1} (X - X_i^*)}{h^2}]$ is estimated by $f(X)$ with Bootstrap.

2 Main results

Theorem If $K(u)$ and $f(X)$ satisfy the following conditions:

a) $f(X) \neq 0$, $f''(X)$ is continuous around everywhere and bounded at R^d ;

b) $K(u)$ is the probability density function which is bounded at R^d , and $\int_{R^d} uK(u)du$ =0, $\int_{R^d} u^2 K(u)du < +\infty$;

c) $\lim_{|u| \to \infty} |uK(u)|$ =0 or $f(X)$ is bounded from above on R^d ;

d) $h = \left(\frac{(\log n)^{1/2}}{n} \right)^{\frac{1}{d(1+\lambda)}}$, $\lim_{n \leftarrow \infty} nh^d = +\infty$, $\lim_{n \to \infty} \frac{\log n}{nh^d} = 0$.

Then if $n \to +\infty$, there is

$$\left\| P^* \left\{ \sqrt{nh^d} \det(s)^{1/2} (f_n^*(X) - f_n(X)) \leq Z \right\} - P \left\{ \sqrt{nh^d} \det(s)^{1/2} (f_n(X) - f(X)) \leq Z \right\} \right\|_\infty$$

$$= o\left(\frac{(\log n)^{\frac{\lambda(1+d)+d}{4d(1+\lambda)}}}{n^{\frac{\lambda}{2d(1+\lambda)}}} M_n\right), a.s.$$

.

3 Several Lemmas

Almost the same as in Chen –Fang[4],we can get the following lemmas.

Lemma 1 If $K(u)$ and $f(X)$ satisfy the parts b), c) of Theorem1, and let

$g_n(X) = \frac{1}{h^d \det(S)^{1/2}} \int_{R^d} K^j \left[\frac{(X-u_i)^T S^{-1} (X-u_i)}{h^2} \right] f(u) du$, then we have

$$\lim_{n\to\infty} g_n(X) = f(X) \int_{R^d} K^j(u) du, \text{ j=1,2,3.}$$

Lemma 2 If $K(u)$ and $f(X)$ satisfy the parts b), c), d) of Theorem1, then if $n \to +\infty$, there have

$$\sqrt{nh^d \det(S)^{1/2}} (Ef_n(X) - f(X)) \to 0.$$

Lemma 3 If $K(u)$ and $f(X)$ satisfy the parts b), c), d) of Theorem1, and X is the continuous point of $f(X)$, we have

$$\lim_{n\to\infty} \frac{1}{nh^d \det(S)^{1/2}} \sum_{i=1}^{n} K^j \left[\frac{(X-X_i)^T S^{-1} (X-X_i)}{h^2} \right] = f(X) \int_{R^d} K^j(u) du, \text{ (j=2, 3)}.$$

Lemma 4 For $p > 0$, there have[5]

(1) $\sup_x |\Phi(x+q) - \Phi(x)| \le \frac{|q|}{\sqrt{2\pi}}$,

(2) $\sup_x |\Phi(px) - \Phi(x)| \le \frac{1}{\sqrt{2\pi e}} \{ |p-1| + |p^{-1} - 1| \}$.

Lemma 5 If $K(u)$ and $f(X)$ satisfy the parts a),b), c), d) of Theorem1, there have[3]

$$\sup_z \left| P\left\{ \frac{\sqrt{nh^d \det(S)^{1/2}} (f_n(X) - f(X))}{(f(X) \int_{R^d} K^2(u) du)^{1/2}} \le Z \right\} - \Phi(Z) \right| \to 0,$$ where $\Phi(Z)$ is the standard normal distribution function.

4 Proofs of Theorem

Proof. Note that $V = \frac{(X-X_i)^T S^{-1} (X-X_i)}{h^2}$, $V^* = \frac{(X-X_i^*)^T S^{-1} (X-X_i^*)}{h^2}$, denote $\sup_z |\ |$ by $\|\ \|_\infty$.

$$\left\| P^*\left\{ \sqrt{nh^d \det(s)^{1/2}} (f_n^*(X) - f_n(X)) \le Z \right\} - P\left\{ \sqrt{nh^d \det(s)^{1/2}} (f_n(X) - f(X)) \le Z \right\} \right\|_\infty$$

$$\le \left\| P^*\left\{ \sqrt{nh^d \det(s)^{1/2}} (f_n^*(X) - f_n(X)) \le Z \right\} - \Phi\left(\frac{Z}{\left(\frac{1}{h^d \det(s)^{1/2}} Var^* K[V^*] \right)^{1/2}} \right) \right\|_\infty$$

$$+ \left\| P\left\{ \sqrt{nh^d \det(s)^{1/2}} (f_n(X) - Ef_n(X)) \le Z + \sqrt{nh^d \det(s)^{1/2}} (f(X) - Ef_n(X)) \right\} \right.$$

$$-\Phi\left(\frac{Z+\sqrt{nh^d \det(s)^{1/2}}(f(X)-Ef_n(X))}{\left(\frac{1}{h^d \det(s)^{1/2}} Var^* K[V^*]\right)^{1/2}}\right)\Bigg\|_\infty$$

$$+\left\|\Phi\left(\frac{Z}{\left(\frac{1}{h^d \det(s)^{1/2}} Var^* K[V^*]\right)^{1/2}}\right)-\Phi\left(\frac{Z+\sqrt{nh^d \det(s)^{1/2}}(f(X)-Ef_n(X))}{\left(\frac{1}{h^d \det(s)^{1/2}} Var^* K[V^*]\right)^{1/2}}\right)\right\|_\infty$$

$$\hat{=} L_1 + L_2 + L_3 \ . \qquad (1)$$

Almost analogy the same proof of I_1 as Lemma 5, we obtain

$$L_1 = O((nh^d \det(s)^{1/2})^{-1/2}) \quad a.s. \qquad (2)$$

Almost analogy the same proof of J_{1n} as Lemma 5, we obtain

$$L_2 = O((nh^d \det(s)^{1/2})^{-1/2}) \quad a.s. \qquad (3)$$

$$L_3 \le \left\|\Phi\left(\frac{Z}{\left(\frac{1}{h^d \det(s)^{1/2}} Var^* K[V^*]\right)^{1/2}}\right)-\Phi\left(\frac{Z}{\left(\frac{1}{h^d \det(s)^{1/2}} VarK[V]\right)^{1/2}}\right)\right\|_\infty$$

$$+\left\|\Phi\left(\frac{Z}{\left(\frac{1}{h^d \det(s)^{1/2}} VarK[V]\right)^{1/2}}\right)-\Phi\left(\frac{Z+\sqrt{nh^d \det(s)^{1/2}}(f(X)-Ef_n(X))}{\left(\frac{1}{h^d \det(s)^{1/2}} Var^* K[V^*]\right)^{1/2}}\right)\right\|_\infty$$

$$\hat{=} L_3^{(1)} + L_3^{(2)} \ . \qquad (4)$$

By Lemma 4, we have

$$L_3^{(1)} \le \frac{1}{\sqrt{2\pi e}}\left|\left(\frac{VarK[V]}{Var^* K[V^*]}\right)^{1/2}-1\right| + \frac{1}{\sqrt{2\pi e}}\left|\left(\frac{Var^* K[V^*]}{VarK[V]}\right)^{1/2}-1\right|$$

$$=\frac{1}{\sqrt{2\pi e}}\left\{\frac{\frac{1}{h^d \det(s)^{1/2}} Var^* K[V^*]-\frac{1}{h^d \det(s)^{1/2}} VarK[V]}{\left(\frac{1}{h^d \det(s)^{1/2}} Var^* K[V^*]\right)^{1/2}\left(\frac{1}{h^d \det(s)^{1/2}} VarK[V]\right)^{1/2}}\right\} \cdot \qquad (5)$$

With marks $b_n(X), a_n(X)$

$$\frac{1}{h^d \det(s)^{1/2}} Var^* K[V^*] = b_n^2(X) \to f(X)\int_{R^d} K^2(u)du, a.s.$$

$$\frac{1}{h^d \det(s)^{1/2}} VarK[V] = a_n^2(X) \to f(X)\int_{R^d} K^2(u)du, a.s.$$

Since $K(u)$ is bounded on R^d, there exist a constant $M > 0$, which satisfy $|K(u)| \le M$,

$$\left| \frac{1}{h^d \det(s)^{1/2}} Var^* K[V^*] - \frac{1}{h^d \det(s)^{1/2}} VarK[V] \right|$$

$$\le \left| \frac{1}{nh^d \det(s)^{1/2}} \sum_{i=1}^{n} (K^2[V] - EK^2[V]) \right| + 2M \left| \frac{1}{nh^d \det(s)^{1/2}} \sum_{i=1}^{n} (K[V] - EK[V]) \right|.$$

For $h = (\frac{(\log n)^{1/2}}{n})^{\frac{1}{d(1+\lambda)}}$, by the inequality of Bernstein, there is

$$\sum_{n=N}^{\infty} P\left\{ \left| \frac{1}{nh^d \det(s)^{1/2}} \sum_{i=1}^{n} (K^2[V] - EK^2[V]) \right| \ge \varepsilon (\frac{(\log n)^{\frac{\lambda(1+d)+d}{4d(1+\lambda)}}}{n^{\frac{\lambda}{2d(1+\lambda)}}} M_n) \right\} < +\infty .$$

Therefore, by the lemma of Borel-Cantelli, we get

$$\left| \frac{1}{nh^d \det(s)^{1/2}} \sum_{i=1}^{n} (K^2[V] - EK^2[V]) \right|$$

$$= o(\frac{(\log n)^{\frac{\lambda(1+d)+d}{4d(1+\lambda)}}}{n^{\frac{\lambda}{2d(1+\lambda)}}} M_n), a.s. \qquad (6)$$

The same proof as above, we get

$$\left| \frac{1}{nh^d \det(s)^{1/2}} \sum_{i=1}^{n} (K[V] - EK[V]) \right|$$

$$= o(\frac{(\log n)^{\frac{\lambda(1+d)+d}{4d(1+\lambda)}}}{n^{\frac{\lambda}{2d(1+\lambda)}}} M_n), a.s. \qquad (7)$$

Combining with (5), (6) and (7), we have

$$L_3^{(1)} = o(\frac{(\log n)^{\frac{\lambda(1+d)+d}{4d(1+\lambda)}}}{n^{\frac{\lambda}{2d(1+\lambda)}}} M_n), a.s. \qquad (8)$$

By Lemma 4, we have

$$L_3^{(2)} \le \frac{1}{\sqrt{2\pi}} \frac{\sqrt{nh^d \det(s)^{1/2}}\,|f(X)-Ef_n(X)|}{(\frac{1}{h^d \det(s)^{1/2}} VarK[V])^{1/2}},$$

$$|f(X)-Ef_n(X)|$$

$$= \left| f(X) - \frac{1}{h^d \det(s)^{1/2}} \int_{R^d} K[\frac{(X-u)^T S^{-1}(X-u)}{h^2}] f(u)du \right|$$

$$\le Ch^{d\lambda} \int_{R^d} |v|^{\lambda} K(v)dv .$$

For $\frac{1}{h^d \det(s)^{1/2}} VarK[V] = a_n^2(X) \to f(X)\int_{R^d} K^2(u)du, a.s.$

$$h = (\frac{(\log n)^{1/2}}{n})^{\frac{1}{d(1+\lambda)}} .$$

So $L_3^{(2)} = O(n^{\frac{1}{2d}} h^{\frac{1}{2}+\lambda})$

$$= O(\frac{(\log n)^{\frac{\lambda(1+d)+d}{4d(1+\lambda)}}}{n^{\frac{\lambda}{2d(1+\lambda)}}}), a.s. \qquad (9)$$

Combining with (4), (8) and (9), we have

$$L_3 = o(\frac{(\log n)^{\frac{\lambda(1+d)+d}{4d(1+\lambda)}}}{n^{\frac{\lambda}{2d(1+\lambda)}}} M_n), a.s. \qquad (10)$$

Together with (1), (2), (3), (10) and $h = (\frac{(\log n)^{1/2}}{n})^{\frac{1}{d(1+\lambda)}}$, we obtain

$$\left\| P^* \left\{ \sqrt{nh^d \det(s)^{1/2}} (f_n^*(X) - f_n(X)) \le Z \right\} - P \left\{ \sqrt{nh^d \det(s)^{1/2}} (f_n(X) - f(X)) \le Z \right\} \right\|_{\infty}$$

$$= o(\frac{(\log n)^{\frac{\lambda(1+d)+d}{4d(1+\lambda)}}}{n^{\frac{\lambda}{2d(1+\lambda)}}} M_n), a.s.$$

This completes the proof of the theorem.

Acknowledgement

The authors gratefully acknowledge the support from department of education-funded projects in Guangxi (201010LX459), the introduction of talent project of He chi University of China (2010QS-N003). School-level issues of He chi University of China (2009A-N006).

References

[1] B. Eforn. Bootstrap methods, Another look at the jackknife[J], Ann.Statist. 1979, 7 (1)p1-26.

[2]Qian Weiming. Bootstrapping kernel estimates [J], Journal of Tongji University. 1992, 20(1)p 87-93.

[3]Li Dewang, Chen xin, Yu Dalei. Bootstrap approximation for the kernel estimates of multivariate density[J], Journal of Southwest University. 2007, 29 (11)p 34-37.

[4]Chen Xiru, Fag Zhaoben, Lie Guoying . NonparametricStatistics [M]. Shanghai Scientific and Technical Publishers. 1989, p 283-305.

[5]Gänssler P.et al. Wahrscheinlichkeitstheorie[M]. Heidelberg: Springer Verlag. 1977, p105-106.