Pure Mathematics 理论数学, 2014, 4, 223-228
Published Online November 2014 in Hans. http://www.hanspub.org/journal/pm
http://dx.doi.org/10.12677/pm.2014.46032

Moment Consistency of Kernel Regression Estimation for Random Design

Xin Yang

Department of Mathematics, Guilin University of Aerospace Technology, Guilin
Email: xinyang_emily@sina.cn

Received: Sep. 19th, 2014; revised: Oct. 15th, 2014; accepted: Oct. 20th, 2014

Abstract

For the nonparametric regression model with random design, we discuss the Nadaraya-Watson type kernel regression estimator for $\tilde{\rho}$ mixing samples, and prove the point moment consistency and the global moment consistency of the kernel regression estimator. The obtained results generalize the Devroye's (1981) conclusions.

Keywords

Random Design, $\tilde{\rho}$ Mixing Sample, Kernel Regression Estimator, Moment Consistency

随机设计核回归估计的矩相合性

杨　昕

桂林航天工业学院数理部，桂林
Email: xinyang_emily@sina.cn

收稿日期：2014年9月19日；修回日期：2014年10月15日；录用日期：2014年10月20日

摘　要

对随机设计非参数回归模型，在$\tilde{\rho}$混合样本下研究Nadaraya-Watson型核回归估计，证明了这种核回归

估计的逐点矩相合性和全局矩相合性，所获结果推广了Devroye (1981)的结论。

关键词

随机设计，$\tilde{\rho}$ 混合样本，核回归估计，矩相合性

1. 引言

设 X 是 d 维随机向量，Y 是一维随机变量。如果 $E|Y|<\infty$，则回归函数 $m(x)=E[Y|X=x]$ 存在。对于来自总体 (X,Y) 的一组样本 $(X_1,Y_1),(X_2,Y_2),\cdots,(X_n,Y_n)$，Nadaraya (1964, 1965) [1] [2]和 Watson [3] (1964)提出了如下核回归估计

$$m_n(x)=\sum_{i=1}^{n}W_{ni}(x)Y_i\,, \tag{1.1}$$

其中

$$W_{ni}(x)=K\left(\frac{x-X_i}{h_n}\right)\Big/\sum_{j=1}^{n}K\left(\frac{x-X_j}{h_n}\right), \tag{1.2}$$

$K(u)$ 为核函数，$h_n>0$ 为窗宽。这种估计通常称为随机设计的 NW 核回归估计。

关于随机设计的 NW 核回归估计的相合性质，有许多学者做过研究。Devroye and Wagner [4]于 1980 年在独立样本条件下证明了 NW 核回归估计的全局矩相合性，并将结果应用于证明相应的判别分析的 Bayes 风险的相合性。Devroye [5] (1981)进一步讨论了这种核回归估计的逐点矩相合性和全局矩相合性，以及强相合性。Greblicki et al. [6] (1984)证明 NW 核回归估计的弱相合性，并在 Y 有界的条件下证明了该估计的强相合性。Nze et al. [7] (2002)在一类弱相依样本下讨论 NW 核回归估计和相应的 Boostrap 估计的渐近性质，如渐近方差、渐近正态性、强相合性及其收敛速度。本文将在 $\tilde{\rho}$ 混合样本下研究 NW 核回归估计的逐点矩相合性和全局矩相合性，所获结果推广了 Devroye [5] (1981)的结论。

2. 主要结论

我们首先明确 $\tilde{\rho}$ 混合的定义。设 $\{\xi_i:i\in N\}$ 是概率空间 $(\Omega,\mathscr{F},P)$ 中的随机变量序列(其中 N 为全体自然数集)，$\mathscr{F}_S=\sigma(X_i:i\in S\subset N)$ 为 σ-域。对 $\mathscr{F}$ 中任意两个子 σ-域 $\mathscr{A}$ 和 $\mathscr{B}$，令

$$\rho(\mathscr{A},\mathscr{B})=\sup\left\{\left|\operatorname{corr}(\xi,\eta)\right|:\xi\in L^2(\mathscr{A}),\eta\in L^2(\mathscr{B})\right\}. \tag{2.1}$$

Bradley [8] [9] (1990, 1992)引进 $\tilde{\rho}$ 混合系数：对 $k\ge 0$，令

$$\tilde{\rho}(k)=\sup\left\{\rho(\mathscr{F}_S,\mathscr{F}_T):\text{有限子集}S,T\subset N,\operatorname{dist}(S,T)\ge k\right\}. \tag{2.2}$$

如果存在 $k\ge 1$ 使得 $\tilde{\rho}(k)<1$，则称随机变量序列 $\{\xi_i:i\in N\}$ 是 $\tilde{\rho}$ 混合的。

我们需要如下假设

(A1) $\{(X_i,Y_i):i\ge 1\}$ 为 $\tilde{\rho}$ 混合的随机向量序列，且 (X_i,Y_i) 与 (X,Y) 同分布。

(A2) 存在正常数 $\lambda\in(0,1)$，使得 $n^{1-\lambda}h_n^d\to\infty$ (当 $n\to\infty$ 时)，其中 d 为随机向量 X 的维数。

(A3) 存在正常数 c_1，c_2 和 r，使得

$$c_1 I(\|u\|\le r)\le K(u)\le c_2 I(\|u\|\le r).$$

本文将使用缩写“$a.e.(\mu)x\in R^d$”表示结论“对 $x\in R^d$ 关于测度 μ 几乎处处”成立，并用 μ_X 表示随

机变量 X 的概率测度。

定理 1：设条件(A1)~(A3)成立。如果 $E|Y|^p<\infty$ (其中 $p>1$)，则

$$E|m_n(x)-m(x)|^p\to 0,\ \ a.e.(\mu_X)\,x\in R^d. \tag{2.3}$$

定理 2：在定理 1 的条件下，有

$$E|m_n(X)-m(X)|^p\to 0. \tag{2.4}$$

3. 定理的证明

为了证明定理的结论，我们需要下面几个引理。

引理 1([4], Lemma 1.1)：设 S_r 表示在 R^d 中以 x 为中心 r 为半径的闭球。如果 $f(x)\in L^1(\mu)$，即 $\int|f(x)|\mu(\mathrm{d}x)<\infty$，则当 $r\to 0$ 时，有

$$\frac{1}{\mu(S_r)}\int_{S_r}f(y)\mu(\mathrm{d}y)\to f(x),\ \ a.e.(\mu_X)\,x\in R^d. \tag{3.1}$$

引理 2([4], Lemma 2.2)：设 S_r 表示在 R^d 中以 x 为中心 r 为半径的闭球，则存在关于测度 μ 几乎处处有限的非负函数 $g(x)$，使得当 $r\to 0$ 时

$$r^d/\mu(S_r)\to g(x),\ \ a.e.(\mu_X)\,x\in R^d. \tag{3.2}$$

从而，如果当 $n\to\infty$ 时 $h_n\to 0$ 且 $nh_n^d\to\infty$，则

$$n\mu(S_{h_n})\to\infty,\ \ a.e.(\mu_X)\,x\in R^d. \tag{3.3}$$

引理 3([10]，定理 2)：设 $\{\xi_j: j\ge 1\}$ 是 $\tilde{\rho}$ 混合的随机序列，$\tilde{\rho}(1)<1$，$E\xi_j=0$ 和 $E|\xi_j|^q<\infty$ $(q>1,\forall j\ge 1)$，则存在与 n 无关的常数 $C>0$，使得当 $1<q\le 2$ 时

$$E\left|\sum_{j=1}^n\xi_j\right|^q\le C\sum_{j=1}^n E|\xi_j|^q, \tag{3.4}$$

当 $q>2$ 时

$$E\left|\sum_{j=1}^n\xi_j\right|^q\le C\left\{\sum_{j=1}^n E|\xi_j|^q+\left(\sum_{j=1}^n E\xi_j^2\right)^{q/2}\right\}. \tag{3.5}$$

引理 4：设条件(A1)~(A3)成立。记

$$u_n=K\left(\frac{x-X}{h_n}\right),\ \ V_n=\sum_{i=1}^n K\left(\frac{x-X_i}{h_n}\right). \tag{3.6}$$

则当 n 充分大时有

$$V_n\ge nu_n/2,\ \ a.s. \tag{3.7}$$

证明：记 $\xi_{n,i}=K\left(\frac{x-X_i}{h_n}\right)-EK\left(\frac{x-X_i}{h_n}\right)$。显然

$$P(V_n<nu_n/2)=P\left(\sum_{i=1}^n\xi_{n,i}<-nu_n/2\right)\le P\left(\left|\sum_{i=1}^n\xi_{n,i}\right|>nu_n/2\right). \tag{3.8}$$

对任意实数 $s>1$，利用条件(A3)，有 $E|\xi_{n,i}|^s\le 2^s EK^s\left(\frac{x-X_i}{h_n}\right)\le 2^s c_2^{s-1}u_n$。选择常数 $q>2/\lambda$。由于 $0<\lambda<1$，

所以 $q>2$。根据 Markov 不等式和引理 3，有

$$\begin{aligned} P\left(V_n<nu_n/2\right) &\le \frac{C}{\left(nu_n\right)^q}E\left|\sum_{i=1}^n \xi_{n,i}\right|^q \le \frac{C}{\left(nu_n\right)^q}\left\{\sum_{i=1}^n E\left|\xi_{n,i}\right|^q+\left(\sum_{i=1}^n E\xi_{n,i}^2\right)^{q/2}\right\} \\ &\le \frac{C}{\left(nu_n\right)^q}\left\{nu_n+\left(nu_n\right)^{q/2}\right\}. \end{aligned} \tag{3.9}$$

由条件(A3)和引理 2 知，$nu_n \ge n\mu_X\left(S_{rh_n}\right)\to\infty$。所以当 n 充分大时，$nu_n \le \left(nu_n\right)^{q/2}$。从而上式变为

$$P\left(V_n<nu_n/2\right)\le\frac{C}{\left(nu_n\right)^q}\le\frac{C}{\left(n\mu_X\left(S_{rh_n}\right)\right)^{q/2}}\le C\left(\frac{1}{n\left(rh_n\right)^d}\frac{\left(rh_n\right)^d}{\mu_X\left(S_{rh_n}\right)}\right)^{q/2}\le Cn^{-\lambda q/2}. \tag{3.10}$$

由于 $\lambda q/2>1$，所以 $\sum_{n=1}^\infty P\left(V_n<nu_n/2\right)<\infty$。由 Borel-Cantelli 引理知，

$$P\left(\left\{V_n<nu_n/2\right\},\ i.o.\right)=0. \tag{3.11}$$

因此，存在 $N>1$，当 $n\ge N$ 时，$V_n\ge nu_n/2,\ a.s.$ 证毕。

引理 5：设条件(A1)-(A3)成立，且 $p\ge1$，如果 $f\in L^p\left(\mu_X\right)$，那么

$$\sum_{i=1}^n E\left[W_{ni}\left(x\right)\left|f\left(X_i\right)-f\left(x\right)\right|^p\right]\to0,\quad a.e.\left(\mu_X\right)x\in R^d \tag{3.12}$$

证明：由条件(A3)和引理 4 知，当 n 充分大时

$$W_{ni}\left(x\right)=V_n^{-1}K\left(\frac{x-X_i}{h_n}\right)\le\frac{2c_2}{nu_n}I\left(\left\|x-X_i\right\|\le rh_n\right),\ a.s. \tag{3.13}$$

由条件(A3)，有 $u_n\ge c_1\mu_X\left(S_{rh_n}\right)$。于是

$$\begin{aligned} \sum_{i=1}^n E\left[W_{ni}\left(x\right)\left|f\left(X_i\right)-f\left(x\right)\right|^p\right] &\le \frac{2c_2}{nu_n}\sum_{i=1}^n E\left[I\left(\left\|x-X_i\right\|\le rh_n\right)\left|f\left(X_i\right)-f\left(x\right)\right|^p\right] \\ &\le \frac{2c_2}{nu_n}\int_{S_{rh_n}}\left|f\left(y\right)-f\left(x\right)\right|^p\mu_X\left(\mathrm{d}y\right)\le\frac{2c_2}{c_1}\frac{1}{\mu_X\left(S_{rh_n}\right)}\int_{S_{rh_n}}\left|f\left(y\right)-f\left(x\right)\right|^p\mu_X\left(\mathrm{d}y\right). \end{aligned} \tag{3.14}$$

由引理 1 知，上式趋向于 0。证毕。

定理 1 的证明：记 $\eta_{n,i}=K\left(\frac{x-X_i}{h_n}\right)\left[Y_i-m\left(X_i\right)\right]$。显然，$\left\{\eta_{n,i}:1\le i\le n\right\}$ 是 $\tilde\rho$ 混合，$E\left(\eta_{n,i}\right)=0$。令

$$G\left(v\right)=E\left\{\left|\sum_{i=1}^n W_{ni}\left(x\right)\left[Y_i-m\left(X_i\right)\right]\right|^p\middle|V_n=v\right\}. \tag{3.15}$$

我们有

$$G\left(v\right)=E\left\{\left|\sum_{i=1}^n v^{-1}K\left(\frac{x-X_i}{h_n}\right)\left[Y_i-m\left(X_i\right)\right]\right|^p\middle|V_n=v\right\}=Cv^{-p}E\left|\sum_{j=1}^n\xi_{n,i}\right|^p. \tag{3.16}$$

当 $p>2$ 时，由引理 3

$$G\left(v\right)\le Cv^{-p}\left\{\sum_{i=1}^n E\left|\eta_{n,i}\right|^p+\left(\sum_{i=1}^n E\eta_{n,i}^2\right)^{p/2}\right\}\le Cv^{-p}\left\{nE\left|\eta_{n,1}\right|^p+\left[nE\left(\eta_{n,1}\right)^2\right]^{p/2}\right\}. \tag{3.17}$$

令 $h_s(x)=E\left[\left|Y-m(X)\right|^s \middle| X=x\right]$。对任意的常实数 $s\in[1,p]$，有

$$E\left|\eta_{n,1}\right|^s \le c_2^s E\left[I\left(\|x-X_1\|\le rh_n\right)\left|Y_1-m(X_1)\right|^s\right]$$
$$=c_2^s E\left[I\left(\|x-X_1\|\le rh_n\right)h_s(X)\right]=c_2^s\int_{S_{rh_n}} h_s(y)\mu_X(\mathrm{d}y) \tag{3.18}$$

由于 $E|Y|^p<\infty$，所以 $E\left|h_s(X)\right|<\infty$。利用引理 1，得 $\int_{S_{rh_n}} h_s(y)\mu_X(\mathrm{d}y)\Big/\mu_X\left(S_{rh_n}\right)\to h_s(x)$。从而

$$E\left|\eta_{n,i}\right|^s \le C\mu_X\left(S_{rh_n}\right). \tag{3.19}$$

将此代入(3.17)式，且注意到 $n\mu_X\left(S_{rh_n}\right)\to\infty$ (引理 2)，得

$$G(v)\le Cv^{-p}\left[n\mu_X\left(S_{rh_n}\right)+\left(n\mu_X\left(S_{rh_n}\right)\right)^{p/2}\right]\le Cv^{-p}\left(n\mu_X\left(S_{rh_n}\right)\right)^{p/2}. \tag{3.20}$$

由此和引理 4，得

$$E\left|\sum_{i=1}^n W_{ni}(x)\left[Y_i-m(X_i)\right]\right|^p = EG(V_n)\le C\left(n\mu_X\left(S_{rh_n}\right)\right)^{p/2}E\left(V_n^{-p}\right)$$
$$\le C\left(n\mu_X\left(S_{rh_n}\right)\right)^{p/2}\Big/\left(n\mu_X\left(S_{rh_n}\right)\right)^p \tag{3.21}$$
$$\le C\left(n\mu_X\left(S_{rh_n}\right)\right)^{-p/2}\to 0.$$

当 $1<p\le 2$ 时，由(3.16)式和引理 3

$$G(v)\le Cv^{-p}\sum_{i=1}^n E\left|\eta_{n,i}\right|^p \le Cv^{-p}nE\left|\eta_{n,1}\right|^p\le Cv^{-p}n\mu_X\left(S_{rh_n}\right). \tag{3.22}$$

因此

$$E\left|\sum_{i=1}^n W_{ni}(x)\left[Y_i-m(X_i)\right]\right|^p \le Cn\mu_X\left(S_{rh_n}\right)E\left(V_n^{-p}\right)$$
$$\le Cn\mu_X\left(S_{rh_n}\right)\Big/\left(n\mu_X\left(S_{rh_n}\right)\right)^p\le C\left(n\mu_X\left(S_{rh_n}\right)\right)^{-p+1}\to 0. \tag{3.23}$$

联合(3.21)式和(3.23)式得结论。证毕。

定理 2 的证明：令 $\zeta_n(x)=E\left[\left|m_n(X)-m(X)\right|^p \middle| X=x\right]$。由定理 1 知，对 x 关于概率测度 μ_X 几乎处处有 $\zeta_n(x)\to 0$。也就是，$\zeta_n(X)\to 0,\ a.s.(\mu_X)$。如果存在可积函数 $\varphi(X)\in L^1(\mu_X)$ 满足 $0\le\zeta_n(X)\le\varphi(X)$，则由控制收敛定理得

$$\lim_{n\to\infty}E\left|m_n(X)-m(X)\right|^p=\lim_{n\to\infty}E\left[\zeta_n(X)\right]=0, \tag{3.24}$$

这就是定理的结论。所以余下我们只需证明存在这样的可积函数 $\varphi(X)$。

令 $g(x)=E\left[|Y|^p \middle| X=x\right]$。显然，$\zeta_n(x)\ge 0$，且

$$\zeta_n(x)\le 2^{p-1}\left\{E\left[\left|m_n(X)\right|^p \middle| X=x\right]+E\left[\left|m(X)\right|^p \middle| X=x\right]\right\}$$
$$\le 2^{p-1}\left\{E\left[\left|m_n(X)\right|^p \middle| X=x\right]+E\left[|Y|^p \middle| X=x\right]\right\} \tag{3.25}$$
$$\le 2^{p-1}\left\{E\left|m_n(x)\right|^p+g(x)\right\}.$$

利用 Holder 不等式

$$\begin{aligned}E\left|m_n(x)\right|^p &= E\left|\sum_{i=1}^{n}W_{ni}(x)Y_i\right|^p = E\left|\sum_{i=1}^{n}W_{ni}^{1-1/p}(x)\left(W_{ni}^{1/p}(x)Y_i\right)\right|^p \\ &\le E\left\{\left(\sum_{i=1}^{n}W_{ni}(x)\right)^{p-1}\left(\sum_{i=1}^{n}W_{ni}(x)\left|Y_i\right|^p\right)\right\} = \sum_{i=1}^{n}E\left[W_{ni}(x)\left|Y_i\right|^p\right]\end{aligned} \tag{3.26}$$

而由条件(A3)和引理 4

$$\begin{aligned}E\left[W_{ni}(x)\left|Y_i\right|^p\right] &\le \left\{2c_2/(nu_n)\right\}E\left[I\left(\left\|x-X_i\right\|\ge rh_n\right)\left|Y_i\right|^p\right] \\ &= \left\{2c_2/(nu_n)\right\}E\left\{E\left[I\left(\left\|x-X_i\right\|\ge rh_n\right)\left|Y_i\right|^p \mid X\right]\right\} \\ &= \left\{2c_2/(nu_n)\right\}E\left\{I\left(\left\|x-X_i\right\|\ge rh_n\right)g(X)\right\} \le \frac{2c_2}{c_1 n}\frac{1}{\mu_X\left(S_{rh_n}\right)}\int_{S_{rh_n}}g(y)\mu_X(\mathrm{d}y).\end{aligned} \tag{3.27}$$

联合上面两式，且利用引理 1，有

$$E\left|m_n(x)\right|^p \le \frac{2c_2}{c_1}\frac{1}{\mu_X\left(S_{rh_n}\right)}\int_{S_{rh_n}}g(y)\mu_X(\mathrm{d}y) \to \frac{2c_2}{c_1}g(x). \tag{3.28}$$

从而存在正常数 c_3 使得 $E\left|m_n(x)\right|^p \le c_3\left[g(x)+1\right]$。因此

$$\zeta_n(x) \le 2^{p-1}\left\{c_3\left[g(x)+1\right]+g(x)\right\} \le c_4\left[g(x)+1\right]. \tag{3.29}$$

令 $\varphi(X)=c_4\left[g(X)+1\right]$。由于 $E\varphi(X)=c_4\left[Eg(X)+1\right]=c_4\left[E|Y|^p+1\right]<\infty$，所以 $\varphi(X)\in L^1(\mu_X)$。因此，函数 $\varphi(X)$ 满足要求。证毕。

基金项目

国家自然科学基金项目(11461009)，广西自然科学基金项目(2011GXNSFA018133)。

参考文献 (References)

[1] Nadaraya, E.A. (1964) On estimating regression. *Theory of Probability and Its Applications*, **9**, 141-142.

[2] Nadaraya, E.A. (1965) On non-parametric estimates of density functions and regression curves. *Theory of Probability and Its Applications*, **10**, 186-190.

[3] Watson, G.S. (1964) Smooth regression analysis. *Sankhya*: *The Indian Journal of Statistics*, *Series A*, **26**, 359-372.

[4] Devroye, L. (1981) On the almost everywhere convergence of nonparametric regression function estimates. *The Annals of Statistics*, **9**, 1310-1319.

[5] Devroye, L.P. and Wagner, T.J. (1980) Distribution-free consistency results in nonparametric discrimination and regression function estimation. *The Annals of Statistics*, **8**, 231-239.

[6] Greblicki, W., Krzyzak, A. and Pawlak, M. (1984) Distribution-free pointwise consistency of kernel regression estimate. *The Annals of Statistics*, **12**, 1570-1575.

[7] Nze, P.A., Buhlmann, P. and Doukhan, P. (2002) Weak dependence beyond mixing and asymptotics for nonparametric regression. *Annals of Statistics*, **30**, 397-430.

[8] Bradley, R.C. (1990) Equivalent mixing conditions for random fields. Technical Report No. 336, Center for Stochastic Processes, University of North Carolina, Chapel Hill.

[9] Bradley, R.C. (1992) On the spectral density and asymptotic normality of weakly dependent random fields. *Journal of Theoretical Probability*, **5**, 355-374.

[10] 杨善朝 (1998) 一类随机变量部分和的矩不等式及其应用. *科学通报*, **17**, 1823-1828.