第24卷第2期
2010年3月

Vol.24 No.2
Mar. 2010

一类变系数模型局部最小二乘估计的重对数律

张荣强，余 波，董 宁

（湖南工业大学 理学院，湖南 株洲 412008）

摘 要：考虑一类变系数模型，保持第 p 项系数函数不动而对前 $p-1$ 项系数函数在 t_0 点展开，得到了部分线性模型的近似形式。基于半参数回归模型的方法，建立了局部最小二乘估计量，在一定合适的条件下讨论了估计量收敛到真实参数速度的重对数律。

关键词：重对数律；局部最小二乘；变系数模型

中图分类号：O212.1　　**文献标志码**：A　　**文章编号**：1673-9833(2010)02-0027-03

The Laws of Iterated Logarithm of Local Least Square Estimators in A Class of Varing-Cofficient Model

Zhang Rongqiang，Yu Bo，Dong Ning

（School of Science, Hunan University of Technology，Zhuzhou Hunan 412008，China）

Abstract：Considers a class of varing-coefficient model. By keeping the p-th function coefficient remained，expands first $p-1$ function coefficients at t_0 and obtains the approximate form of the partly linear model. Based on the method of semi-parametric regression model, establishes the local least square estimators. Then discusses the laws of iterated logarithm of the proposed estimators under some conditions.

Keywords：laws of iterated logarithm; local least square method; varing-coefficient model

0 引言

考虑一类变系数模型

$$y_i = x_{i1}\beta_1(t_i)+\cdots+x_{i,p-1}\beta_{p-1}(t_i)+\beta_p(t_i)+\sigma_i e_i, \qquad (1)$$

式中：$i \leqslant n$；$\{x_{ij}\}$和$\{t_i\}$为回归变量；$\{y_i\}$为响应变量；$\{e_i\}$是独立同分布的随机变量，且满足 $Ee_i=0$，$Ee_i^2=1$，$\sup_{i\leqslant n} E|e_i|^{2+\varepsilon}<\infty \ (\varepsilon>0)$；$\{\beta_i(t_i)\}$是未知的具有有界二阶导数的函数。

该模型由Cleveland等针对高维数据的“维数祸根”问题提出，既部分继承了非参数回归的稳定性特点，又保留了经典线性模型直观、容易解释的特点。在现代数理统计中高维数据处理方面得到了广泛应用。

近年来，许多学者对模型（1）中系数函数的估计提出了不同的方法，如核函数[1]、多项式样条函数[2]等。J. Q. Fan 等[3]利用局部最小二乘标准的2步估计量来估计系数函数，得到了2步估计的渐近均方误差和最优收敛律[2]。但在模型（1）中，如果对每个系数函数展开，则忽略了模型本身的特点，必然导致原有模型的一些信息损失。因为第 p 个系数函数是1个与解释变量无关的函数，受部分线性模型中最小二乘估计量的启发，可以想象用已得到的数据暂时不必估计第 p 个系数函数，即先把它当作是已知的，并保留第 p 个分量的函数，仅对前 $p-1$ 个函数系数展开，利用局部最小二乘方法得到前 $p-1$ 个函数的估计量 $\hat{\beta}(t_0)$，再代回原模型中求得第 p 个函数的估计 $\hat{\beta}_p(t)$。这样，在估计前就减少了原来模型的损失，使得模型最大限度地保留了原有具体信息量，从而得到的估计量具有更优良的性质。

在本文中，假设系数函数足够光滑（至少二阶可导），对前 $p-1$ 个系数函数运用局部最小二乘的标准，

收稿日期：2009-08-06

通信作者：张荣强（1983-），男，山东临沂人，湖南工业大学教师，主要研究方向为应用数学，E-mail：zrq1983518@sina.com

对所给的点 t_0 局部近似函数

$$\beta_j(t)=\beta_j(t_0)+\beta_j'(t_0)(t-t_0)+\beta_j''(\xi_0)(t-t_0)^2/2，\quad (2)$$

于是有: $y_i\approx\dot{\boldsymbol{x}}_i^{\mathrm T}\boldsymbol{\beta}(t_0)+\beta_p(t_i)+\sigma_i e_i$，（3）

式（3）中: $\dot{\boldsymbol{x}}_i^{\mathrm T}=(x_{i1},\cdots,x_{ip})\otimes(1,t_i-t_0)$;

$\boldsymbol{\beta}(t_0)^{\mathrm T}=(\beta_1(t_0),\beta_1'(t_0),\cdots,\beta_{p-1}(t_0),\beta_{p-1}'(t_0))$;

⊗表示Kronecker积。

如前述，为了减少原模型的信息损失，由式（3），要求 $\beta_p(t_i)=\mathrm E(y_i-\dot{\boldsymbol{x}}_i^{\mathrm T}\boldsymbol{\beta}(t_0))$，即可得 $\beta_p(t_i)$ 表达式为:

$$\beta_p(t_i)=\sum_{k=1}^n w_{nk}(t_i)(y_k-\dot{\boldsymbol{x}}_k^{\mathrm T}\boldsymbol{\beta}(t_0))。\quad (4)$$

式（4）中: $w_{nk}(t_i)$ $(1\le i\le n)$ 是权函数，因此由模型式（3）产生的最小二乘问题通过最小化

$$\min\sum_{i=1}^n\left(y_i-\dot{\boldsymbol{x}}_i^{\mathrm T}\boldsymbol{\beta}(t_0)-\beta_p(t_i)\right)^2,$$

可得: $$\hat{\boldsymbol{\beta}}(t_0)=\left(\sum_{i=1}^n\tilde{\boldsymbol{x}}_i\tilde{\boldsymbol{x}}_i^{\mathrm T}\right)^{-1}\left(\sum_{i=1}^n\tilde{\boldsymbol{x}}_i\tilde{y}_i\right),\quad (5)$$

其中 $\tilde{\boldsymbol{x}}_i=\dot{\boldsymbol{x}}_i-\sum_{k=1}^n w_{nk}(t_i)\dot{\boldsymbol{x}}_k$, $\tilde{y}_i=y_i-\sum_{k=1}^n w_{nk}(t_i)y_k$。

1 预备知识

为得到本文的主要结果，给出如下假设，并且这些假设对于非参数回归是合理的[4-5]。

假设 A_1 存在连续函数 $h_j(t_i)$ 满足 $x_{ij}=h_j(t_i)+u_{ij}$ $(1\le i\le n,\ 1\le j\le p-1)$，$t_i$ 是定义于[0,1]上的设计点列，设 $\dot{\boldsymbol{u}}_i^{\mathrm T}=(u_{i1},\cdots,u_{i,p-1})\otimes(1,t_i-t_0)$, 假定 $\lim_{n\to\infty}\sum_{i=1}^n\dot{\boldsymbol{u}}_i\dot{\boldsymbol{u}}_i^{\mathrm T}/n=\boldsymbol{\Sigma}_0$，且对 $m=1,\cdots,2p-2$ 和所有 $(1,\cdots,n)$ 的排列 $(j_1,\cdots,j_n)$ 都有 $\limsup_{n\to\infty}\max_{k\le n}\left|\sum_{i=1}^k\dot{u}_{j_i m}\right|/a_n<\infty$，其中 $\boldsymbol{u}_i=(u_{i1},\cdots,u_{i,p-1})^{\mathrm T}$，$a_n=n^{1/2}\log n$，且 $\boldsymbol{\Sigma}_0$ 是一正定矩阵。

假设 A_2 β_p，h_j 是一阶Lipschits连续函数。

假设 A_3 正权函数 $w_{ni}(t_i)$ 满足下列条件:

$$\max_{i\le n}\sum_{k=1}^n w_{nk}(t_i)=O(1),\ \max_{k\le n}\sum_{i=1}^n w_{nk}(t_i)=O(1),$$

$$\max_{i\le n}\sum_{k=1}^n w_{nk}(t_i)\left(I_{(|t_i-t_k|>c_n)}\right)=O(c_n),$$

$$\max_{i\le n}\left|\sum_{k=1}^n w_{nk}(t_i)\dot{u}_{kq}\right|=O(d_n),\ \max_{k\le n}\left|\sum_{i=1}^n w_{nk}(t_i)\dot{u}_{iq}\right|=O(d_n)。$$

这里 c_n，d_n 满足 $\liminf_{n\to\infty}c_n^2n>0$, $\limsup_{n\to\infty}nc_n^4\log n<\infty$, $\limsup_{n\to\infty}nd_n^4\log n<\infty$。

假设 A_4 二阶导数 $\beta_j''(t_0)=O(n^{-s_0})$ 且 $\sup_{i,q}|\dot{u}_{iq}|^{3}<\infty$，其中 $s_0\ge1/2,\ j\le p,\ i\le n,\ q\le2p-2$。

引理 1[6] 设 $V_1,\cdots,V_n$ 为独立同分布的随机变量，$\mathrm EV_i=0$ 且方差有限，并满足 $\sup_{i\le n}\mathrm E|V_i|^r\le c<\infty\ (r\ge2)$，$\{a_{ni}\}$ 是一实序列，满足 $\sup_{1\le i\le n}|a_{ni}|=O(n^{-p_1})$，式中 $0<p_1<1$，且 $\sum_{i=1}^n a_{ni}=O(n^{p_2})$，其中 $p_2\ge\max(0,2/(r-p))$，则有:

$$\max_{k\le n}\left|\sum_{1\le i\le n}a_{ni}V_i\right|=O(n^{-s}\log n)，其中 s=\frac{p_1-p_2}{2}。$$

引理 2[6] 若假设 $A_1\sim A_3$ 成立，则对 $j=0,\cdots,p-1$ 有

$$\max_{i\le n}\left|G_j(t_i)-\sum_{1\le k\le n}w_{nk}(t_i)G_j(t_k)\right|=O(c_n)，其中$$

$G_0(\cdot)=\beta_p(\cdot)$, $G_l(\cdot)=h_l(\cdot)$, $l=0,\cdots,p-1$。

引理 3[6] 设 $\{V_i\}$ 是零均值随机变量，若存在常数 $\delta_0>0$ 使得 $\max_{i\le n}\mathrm E|V_i|^{2+\delta_0}<\infty$ 且 $\liminf_{n\to\infty}\left(\sum_{i=1}^n\mathrm{Var}(V_i)\right)/n>0$，则有 $\limsup_{n\to\infty}|S_n|/\sqrt{2s_n^2\log\log s_n^2}=1$，其中

$$S_n=\sum_{i=1}^nV_i,\quad s_n^2=\sum_{i=1}^n\mathrm EV_i^2。$$

给定假设条件下，由引理1~3可知：当 n 充分大时，$\frac1n\sum_{i=1}^n\tilde{\boldsymbol{x}}_i\tilde{\boldsymbol{x}}_i^{\mathrm T}$ 收敛到一正定矩阵，因此是可逆的。把估计量式（5）代入式（4）中，得到 $\beta_p(t_i)$ 的估计量 $\hat\beta_p(t_i)=\sum_{k=1}^n w_{nk}(t_i)\left(y_k-\dot{\boldsymbol{x}}_k^{\mathrm T}\hat{\boldsymbol{\beta}}(t_0)\right)$。这样，模型（1）中的所有系数函数都能估计出来。文中得到的局部估计量 $\hat{\boldsymbol{\beta}}(t_0)$，以确切的 $(2(\log\log n)/n)^{1/2}$ 阶的速度收敛到真实值的重对数律。

2 重对数率及证明

定理 1 在条件 $A_1\sim A_4$ 下，若 $\mathrm E|\varepsilon_i|^3<\infty$，且 $c_n=n^{-1/3}\log n$，则有: $\limsup_{n\to\infty}(n/2\log\log n)^{1/2}\left|\hat\beta_j(t_0)-\beta_j\right|=(\sigma^2\sigma^{jj})^{1/2}$，

这里 $\hat\beta_j(t_0)$ 和 β_j 分别为 $\hat{\boldsymbol{\beta}}(t_0)$ 和 $\boldsymbol{\beta}$ 的第 j 个分量，$\sigma^{jj}=(\boldsymbol{a}^j)^{\mathrm T}\boldsymbol{\Sigma}_0\boldsymbol{a}^j$，$\boldsymbol{a}^j$ 为 $\boldsymbol{\Sigma}_0^{-1}$ 的第 j 列。

证明 首先给出估计量与真实值之差的分解:

$$\sqrt n\left(\hat{\boldsymbol{\beta}}(t_0)-\boldsymbol{\beta}\right)=\sqrt n\left(\sum_{i=1}^n\tilde{\boldsymbol{x}}_i\tilde{\boldsymbol{x}}_i^{\mathrm T}\right)^{-1}\left\{\sum_{i=1}^n\tilde{\boldsymbol{x}}_i\tilde\beta_p(t_i)+\sum_{i=1}^n\tilde{\boldsymbol{x}}_i\varepsilon_i+\sum_{i=1}^n\tilde{\boldsymbol{x}}_i\boldsymbol{R}_{ni}\boldsymbol{\beta}''(\xi_0)\right\}=\boldsymbol{A}^{-1}(\boldsymbol{H}_1+\boldsymbol{H}_2+\boldsymbol{H}_3)，\quad (6)$$

其中 $\boldsymbol{A}^{-1}=n\left(\sum_{i=1}^n\tilde{\boldsymbol{x}}_i\tilde{\boldsymbol{x}}_i^{\mathrm T}\right)^{-1}$，$\boldsymbol{H}_1=\sum_{i=1}^n\tilde{\boldsymbol{x}}_i\tilde\beta_p(t_i)/\sqrt n$，

$\boldsymbol{H}_2=\sum_{i=1}^n\tilde{\boldsymbol{x}}_i\varepsilon_i/\sqrt n$，$\boldsymbol{H}_3=\sum_{i=1}^n\tilde{\boldsymbol{x}}_i\boldsymbol{R}_{ni}\boldsymbol{\beta}''(\xi_0)/\sqrt n$，

$$\boldsymbol{R}_{ni}=\left\{(t_i-t_0)^2\boldsymbol{x}_i^{\mathrm T}-\sum_{k=1}^n w_{nk}(t_i)(t_k-t_0)^2\boldsymbol{x}_k^{\mathrm T}\right\}/2。$$

下面证明 $\boldsymbol{H}_1$ 和 $\boldsymbol{H}_3$ 几乎处处收敛到零。

考虑 $\boldsymbol{H}_1$ 的第 j 个分量，因为 $\sqrt nH_{1j}=$

$$\sum_{i=1}^n\tilde h_j(t_i)\tilde\beta_p(t_i)+\sum_{i=1}^n\dot u_{ij}\tilde\beta_p(t_i)-\sum_{i=1}^n\left\{\sum_{k=1}^n w_{nk}(t_i)\dot u_{kj}\right\}\tilde\beta_p(t_i),$$

所以对于上式右边各项，由引理2可知:

$$\left|\sum_{i=1}^{n}\tilde{h}_j(t_i)\tilde{\beta}_r(t_i)\right| = O(nc_n^2),$$

且由假设 A_2 可知：

$$\left|\sum_{i\le n}\dot{\mu}_{ij}\tilde{\beta}_r(t_i)\right| \le \max_{i\le n}\left|\tilde{\beta}_r(t_i)\right|\cdot\max_{i\le n}\sum_{i=1}^{k}\dot{\mu}_{i,m}\Big| = O(a_n c_n)。$$

同时由假设 A_3 可知：$\left|\sum_{i=1}^{n}\left(\sum_{k=1}^{n} w_{nk}(t_i)\dot{\mu}_{kj}\right)\tilde{\beta}_r(t_i)\right|\le$

$$n\max_{i\le n}\sum_{k=1}^{n}w_{nk}(t_i)\dot{\mu}_{kj}\cdot\left|\tilde{\beta}_r(t_i)\right| = O(n^{2/3}\log n d_n),$$

因此有: $\sqrt{n}H_{1j} = O\left(n^{1/2}(\log n)^{-1/2}\right)$。 （7）

考虑 $\sqrt{n}H_{3j}$，由条件 A_3、A_4 得 $R_{ni}\beta''_{\xi_{ni}}$ 第 $j\ (j\le p)$ 元：

$$(R_{ni}\beta''_{\xi_{ni}})_j \le \max_j\left|x_{ij}\beta_{\xi ij}\right| = \max_i\left|h_j(t_i)\beta_{\xi ij}\right| + \max_i\left|\mu_{ij}\beta_{\xi ij}\right| = O(n^{-r_0})。$$

因此由 Cauthy-Schwartz 不等式有：

$$\left|H_{3j}\right|\le n^{-1}\left(\sum_{i=1}^{n}\tilde{x}^2\right)^{\frac{1}{2}}\left(\sum_{i=1}^{n}\left(R_{ni}\beta''_{\xi_{ni}}\right)^2\right)^{\frac{1}{2}} = O\left(n^{-r_0}\right) = o(1)。$$

（8）

接下来考虑 $\sqrt{n}H_{2j}$，可分解为：

$$\sqrt{n}H_{2j} = \sum_{i=1}^n\tilde{x}_{ij}\varepsilon_i - \sum_{i=1}^n\tilde{x}_{ij}\left(\sum_{k=1}^n w_{nk}(t_i)\varepsilon_k\right) \triangleq H_{2j_1} - H_{2j_2}。 \quad (9)$$

而由引理1中的公式和引理2易得：

$$\left|H_{2j_2}\right| \le \left|\sum_{k=1}^n\sum_{i=1}^n w_{nk}(t_i)\tilde{h}_j(t_i)\varepsilon_k\right| + \left|\sum_{k=1}^n\sum_{i=1}^n w_{nk}(t_i)\dot{u}_{ij}\varepsilon_k\right| + \left|\sum_{i=1}^n\left(\left(\sum_{l=1}^n w_{nl}(t_i)\dot{u}_{lj}\right)\left(\sum_{k=1}^n w_{nk}(t_i)\varepsilon_k\right)\right)\right|,$$

由 Cauthy-Schwartz 不等式，上式第一项具有阶 $O(nc_n^2)$；引理1中取 $r=2$，$1/4<p_1<1/3$，$p_2=1-p_1$，则第二项为 $O\left(n^{(2p_1-1)/2}\log n\right)$。此外，由假设 A_3 和引理1中公式知，第三项为 $O\left(n^{2/3}\log n d_n\right)$，从而 $\left|H_{2j_2}\right| = o\left(n^{1/2}\right)$。因此，由式(6)~(9)知定理成立只须用引理3表明：

$$\limsup_{n\to\infty}\left|A^{-1}\left(\sum_{i\le n}\tilde{x}_i\varepsilon_i\right)\right|_j\Big/(2n\log\log n)^{1/2} = \left(\sigma^2\sigma^{jj}\right)^{1/2} \text{ a.s.,} \quad (10)$$

这样只需验证引理3的条件，为此对式（10）分子作分解，记 A^{-1} 中的第 (j,k) 元为 a^{jk}，则有

$$n^{-1/2}\left(A^{-1}\left(\sum_{i\le n}\tilde{x}_i\varepsilon_i\right)\right)_j = n^{-1/2}\left\{\sum_{i=1}^n\sum_{k=1}^p a^{jk}\tilde{h}_k(t_i)\varepsilon_i - \sum_{i=1}^n\left(\sum_{k=1}^p a^{jk}\sum_{q=1}^n w_{nq}(t_i)\dot{u}_{qk}\right)\varepsilon_i + \sum_{i=1}^n\left(\sum_{k=1}^p a^{jk}u_{ik}\right)\varepsilon_i\right\} \triangleq n^{-1/2}(L_1-L_2+L_3)。 \quad (11)$$

由引理2知

$$\left|n^{-1/2}\sum_{k=1}^p a^{jk}\tilde{h}_k(t_i)\right| = O\left(n^{-1/4}(\log n)^{-1/2}\right) = o\left(n^{-1/4}\right),$$

因此，在引理1中可取 $r=2$，$5/6<p_1<1$，$p_2=1-p_1$，则有：

$$n^{-1/2}\left|L_1\right| = O\left(n^{-(2p_1-1)/2}\log n\right) = o(1)。 \quad (12)$$

由假设 A_3 得：$\left|n^{-1/2}\sum_{k=1}^p b^{jk}\sum_{q=1}^n w_{nq}(t_i)\dot{u}_{qk}\right| = O\left(n^{-1/6}\log n\right)$，

同样，在引理1中取 $r=2$，$p_1=3/4$，$p_2=1/4$ 可得

$$n^{-1/2}\left|L_2\right| = O\left(n^{-1/4}\log n\right) = o(1)。 \quad (13)$$

令 $V_i = \left(a^j\right)^T u_i\varepsilon_i$，由假设 A_5 及定理条件有：

$$\max_{i\le n} E\left|V_i\right|^{2+\delta_0} \le CE\left|\varepsilon_i\right|^{2+\delta_0} < \infty。$$

同时可验证

$$\liminf_{n\to\infty}\frac{1}{n}\sum_{i=1}^n \mathrm{Var}(V_i) - \sigma^2\liminf_{n\to\infty}\left(b^j\right)^T u_i u_i^T b^j - \sigma^2\sigma^{jj} > 0。$$

从而由引理3有：

$$\limsup_{n\to\infty}\left(\sum_{i\le n}\left(a^j\right)^T u_i\varepsilon_i \Big/ \left(2\sigma^2\left(a^j\right)^T\sum_{i\le n}u_iu_i^T b^j\log\log\left\{\sum_{i\le n}u_iu_i^Tb^j\right\}\right)^{\frac{1}{2}}\right) = 1。 \quad (14)$$

由 $\sigma^2\left(a^j\right)^T\sum_{i=1}^n u_iu_i^Tb^j$ 的收敛性知式（14）相当于

$$\limsup_{n\to\infty}\frac{\left|\sum_{i\le n}\left(a^j\right)^T u_i\varepsilon_i\right|}{\left(2\sigma^2\sigma^{jj}n\log\log\left\{\sigma^2\sigma^{jj}n\right\}\right)^{1/2}} = 1。 \quad (15)$$

由式（11）~（15）表明，式（10）成立，定理得证。

参考文献：

[1] Wu C O，Chiang C T. Kernel Smoothing on Varing-Coefficient Models with Longitudinal Dependant Variable[J]. Statistica Sinica，2000，10：433−456.

[2] Huang J H，Wu C O，Zhou L. Varing-Coefficient Models and Basic Function Approximations for the Analysis of Repeated Measurements[J]. Biometrika，2003，89：111−128.

[3] Fan J Q，Zhang M. Statistical Estimation in Varing-Coeffcient Models[J]. Ann. Statist.，1999，27：1491−1518.

[4] 高集体，陈希儒，赵林城. 部分线性模型中估计的渐近正态性[J]. 数学学报，1994(2)：96−104.
Gao Jiti，Chen Xiru，Zhao Lincheng. Asymptotic Normality of A Class of Estimators in Partial Linear Models[J]. Acta Mathematica Sinica，1994(2)：96−104.

[5] 高集体，洪圣岩，梁　华. 部分线性模型中估计的收敛速度[J]. 数学学报，1995(5)：658−669.
Gao Jiti，Hong Shengyan，Liang Hua. Convergence Rates of A Class of Estimators in Partly Linear Models[J]. Acta Mathematica Sinica，1995(5)：658−669.

[6] Hardle W，Liang H，Gao J T. Partially Linear Models[M]. Heidelberg：Physica-Varlag，2000.

[7] Liang Hua. The LIL for the Estimates of the Parameters in A Partly Linear Regression Model[J]. Taiwaness Journal of Mathemaics，1999，3，517−528.

（责任编辑：廖友媛）