

具有外生变量部分线性自回归模型的样条估计 ***

武新乾* 田 铮** 韩四儿*

摘要 考虑自回归模型 $Y_t = \theta^T X_t + g(Z_t) + \varepsilon_t$, $t = 1, \dots, n$, 其中 $X_t = (Y_{t-1}, \dots, Y_{t-d})^T$, Z_t 为实值外生随机变量, $\theta = (\theta_1, \dots, \theta_d)^T$ 为待估参数向量, g 为未知非参数光滑函数. 基于多项式样条方法, 在一定的条件下, 给出了 θ 的估计的渐近正态性, 得到了 g 的估计的收敛速度. 模拟例子验证了所得的理论结果.

关键词 外生变量, 部分线性自回归模型, 样条估计, 渐近正态性, 收敛速度

MR (2000) 主题分类 62G05, 62G08, 62M10

中图法分类 O212.7

文献标识码 A

文章编号 1000-8314(2007)03-0377-10

1 引 言

考虑回归模型

$$Y_t = \theta^T X_t + g(Z_t) + \varepsilon_t, \quad t = 1, \dots, n, \quad (1.1)$$

其中 $X_t = (X_{t1}, \dots, X_{td})^T = (Y_{t-1}, \dots, Y_{t-d})^T$, Z_t 为实值外生随机变量, $\theta = (\theta_1, \dots, \theta_d)^T$ 为未知待估参数向量, $g(\cdot)$ 为未知非参数光滑函数, ε_t 为独立同分布随机误差, $E\varepsilon_t = 0$, $E\varepsilon_t^2 = \sigma^2 < \infty$; $\{\varepsilon_t\}$ 与 $\{Z_t\}$ 独立, ε_t 与 $X_{t'i}$ ($t' \leq t, i = 1, \dots, d$) 独立. 文中 $(\cdot)^T$ 表示向量或者矩阵的转置.

模型 (1.1) 为一类具有外生变量的部分线性自回归模型, 它反映了一些自然现象和经济现象中变量之间的关系, 如大气污染的程度不仅与过去大气污染的程度有关, 而且也与大气温度和风速有关; 在计量经济学中, 消费是前期消费和收入的函数 (见 [1]). 对于具有外生变量的自回归模型的研究主要集中于线性模型, Hannan 和 Deistler [2] 系统介绍了线性模型的统计理论. 近年来, 对于具有外生变量的非参数或半参数自回归模型的研究已引起了一些学者的兴趣, 但所见文献甚少. Bosq 和 Shen [1] 针对模型 $Y_t = \phi(Y_{t-1}, \dots, Y_{t-d}) + g(Z_t) + \varepsilon_t$ 中 $\phi(\cdot)$ 为未知非线性函数, $g(\cdot)$ 为线性和非线性函数两种情形讨论了核估计的 a.s. 和 L_s ($s \geq 2$) 收敛性. Liang [3] 考虑了模型 (1.1) 中 $d = 1$ 和 Z_t 为服从 $[0, 1]$ 均匀分布的独立同分布随机变量的一种特殊情形, 基于 g 的分段多项式近似构造了未知参数 θ 的一类二阶渐近有效估计量. Zhu 和 An [4] 在一定条件下给出了 θ 的强相合估计和 g 的相合核估计.

模型 (1.1) 放松了文 [3] 中对 Z_t 为服从 $[0, 1]$ 均匀分布的约束, 具有一般性. 本文旨在研究模型 (1.1) 中的参数向量和非参数函数的估计问题. 众所周知, 局部线性估计比常用的 Nadaraya-Watson 核估计有更好的性质 (见 [5]), 然而, 局部多项式估计 (包括局部线性估计) 不能给出所拟合模型的简单的显式表达式, 计算量较大. 多项式样条估计是全局

本文 2006 年 1 月 19 日收到.

*西北工业大学应用数学系, 西安 710072. E-mail: wuxinqian1001@163.com; hansier2003@126.com

**西北工业大学应用数学系, 西安 710072; 中国科学院自动化研究所模式识别国家重点实验室, 北京 100080.
E-mail: zhtian@nwpu.edu.cn

***国家自然科学基金 (No. 60375003) 和航空基础科学基金 (No. 03I53059) 资助的项目.

光滑方法,能较好地克服上述局部多项式估计的弊端(见[6]).本文利用多项式样条与最小二乘方法相结合,得到了模型(1.1)中 θ 和 g 的相应估计量 $\hat{\theta}$ 和 \hat{g} ,然后,在适当的条件下证明了 $\hat{\theta}$ 的渐近正态性,得到了 \hat{g} 的收敛速度.最后,通过模拟例子验证了所得的理论结果.

2 样条估计

本文仅考虑 $Z_t \in D = [a, b]$ ($a < b$) 时非参数函数的估计问题.将区间 $[a, b]$ 予以剖分,结点序列为

$$a = z_0 < z_1 < \cdots < z_{N_n} < z_{N_n+1} = b. \quad (2.1)$$

设 k 为非负整数,如果一个函数满足:在每个区间 $[z_i, z_{i+1})$ ($i = 0, 1, \cdots, N_n - 1$) 和 $[z_{N_n}, z_{N_n+1}]$ 上为 k 次多项式,且在 $[a, b]$ 上 $k-1$ ($k \geq 1$) 阶连续可微,则称该函数为 k 次多项式样条.多项式样条空间由多项式的次数和结点序列所确定,例如,具有结点序列(2.1)的 k 次多项式样条空间 S_{k, N_n} 构成了 $K = N_n + k + 1$ 维线性空间,其基函数可取为

$$1, z, \cdots, z^k, (z - z_1)_+^k, \cdots, (z - z_{N_n})_+^k, \quad (2.2)$$

这里 $(z - z_i)_+ = \max\{0, z - z_i\}$ ($i = 1, \cdots, N_n$).对样条函数的详细了解见文[7].

模型(1.1)中的非参数函数 g 可由多项式样条函数近似,即存在样条函数空间中的一组基函数 $B_s(\cdot)$ 和常数 β_s ($s = 1, \cdots, K$),使得

$$g(z) \approx \sum_{s=1}^K \beta_s B_s(z). \quad (2.3)$$

令

$$m(\theta, \beta) = \sum_{t=1}^n \left\{ Y_t - \theta^T X_t - \sum_{s=1}^K \beta_s B_s(Z_t) \right\}^2, \quad (2.4)$$

其中 $\beta = (\beta_1, \cdots, \beta_K)^T$.关于 θ 和 β 最小化(2.4),可得它们的估计分别为 $\hat{\theta} = (\hat{\theta}_1, \cdots, \hat{\theta}_d)^T$ 和 $\hat{\beta} = (\hat{\beta}_1, \cdots, \hat{\beta}_K)^T$.

具体地,假定文中涉及的逆矩阵都存在,记 $Y = (Y_1, \cdots, Y_n)^T$, I 为 $n \times n$ 单位矩阵,

$$X = \begin{pmatrix} X_{11} & X_{12} & \cdots & X_{1d} \\ X_{21} & X_{22} & \cdots & X_{2d} \\ \vdots & \vdots & & \vdots \\ X_{n1} & X_{n2} & \cdots & X_{nd} \end{pmatrix}, \quad B = \begin{pmatrix} B_1(Z_1) & B_2(Z_1) & \cdots & B_K(Z_1) \\ B_1(Z_2) & B_2(Z_2) & \cdots & B_K(Z_2) \\ \vdots & \vdots & & \vdots \\ B_1(Z_n) & B_2(Z_n) & \cdots & B_K(Z_n) \end{pmatrix},$$

$$A = B(B^T B)^{-1} B^T, \quad H = I - X[X^T(I - A)X]^{-1} X^T(I - A),$$

则 θ 和 β 的估计分别为

$$\hat{\theta} = [X^T(I - A)X]^{-1} X^T(I - A)Y, \quad \hat{\beta} = (B^T B)^{-1} B^T H Y. \quad (2.5)$$

从而可得 g 的估计为

$$\hat{g}(z) = \sum_{s=1}^K \hat{\beta}_s B_s(z), \quad (2.6)$$

这里 $\hat{\theta}$ 和 \hat{g} 分别为 θ 和 g 的样条估计.

3 主要结果

本节主要给出 θ 和 g 的样条估计的大样本性质. 为方便讨论, 给出几个记号.

对于两个正数序列 $\{a_n\}$ 和 $\{b_n\}$, 如果 $\frac{a_n}{b_n}$ 一致有界, 那么记 $a_n \lesssim b_n$; 如果 $a_n \lesssim b_n$ 且 $b_n \lesssim a_n$, 那么记 $a_n \asymp b_n$. 另外, 记 \xrightarrow{P} 和 \xrightarrow{L} 分别表示依概率收敛和依分布收敛; $C^{k+1}([a, b])$ 表示由 $[a, b]$ 上 $k+1$ 阶导数连续的函数构成的空间; S_{k, N_n} 为紧区间 D 上的样条函数空间 (给定多项式次数 k , 结点序列 (2.1) 具有有限网比, 即相邻结点之间距离之比一致有界); $\|g\|_2 = \{\int_D g^2(x)dx\}^{1/2}$ 为紧区间 D 上平方可积函数 $g(\cdot)$ 的 L_2 -范数; $|\cdot|$ 表示向量的 Euclid 范数或者实数的绝对值; $I(\cdot)$ 为示性函数.

下面的一些条件对于证明本文的主要结果是充分的.

(1) $N_n \asymp n^r$, $\frac{1}{2k+2} < r < 1$.

(2) Z_t 的密度函数 f_Z 绝对连续, 并且 f_Z 在 D 上恒正有界.

(3) $\{(X_t, Z_t, Y_t)\}$ 是严平稳 α -混合过程, 其 α -混合系数 $\alpha(l)$ 满足 $\alpha(l) \leq c_0 l^{-\alpha}$, $\alpha > \frac{(5/2)r}{(1-r)}$ 并且 $(1 - \frac{2}{s_0})\alpha > 1$, 其中 $s_0 > 2$, $c_0 > 0$ 为常数. α -混合系数 $\alpha(l)$ 定义为

$$\alpha(l) = \sup\{|P(A \cap B) - P(A)P(B)|, A \in \mathcal{F}_1^p, B \in \mathcal{F}_{p+l}^\infty, p \in N\},$$

其中 N 为自然数集, $\mathcal{F}_{t_1}^{t_2}$ 为由 $\{(X_t, Z_t, Y_t) : t_1 \leq t \leq t_2\}$ 生成的 σ -代数.

(4) 对充分大的 $m > 0$, 有 $E(|Y_t|^m) < \infty$.

(5) $E\{g(Z_t)\} = 0$.

注 3.1 条件 (1) 意味着 $K \asymp n^r$, $h = \max_{j=0, \dots, N_n} |z_{j+1} - z_j| \asymp n^{-r}$. 条件 (2) 的假设是合理的 (见 [8]). 条件 (3) 和 (4) 是一些较为基本的假设 (见 [6, 9, 10]).

注 3.2 满足条件 (5) 的 $g(Z_t)$ 显然是存在的, 如 Z_t 为服从 $[-1, 1]$ 均匀分布的独立同分布随机变量且 $g(z) = \sin z$. 又如 $\{Z_t\}$ 满足 $Z_t = 0.5Z_{t-1} + e_t$, 其中 $\{e_t\}$ 为独立同分布序列, 且 e_t 与 $\{Z_{t'}, t' < t\}$ 独立, 若 e_t 服从二点分布, 即 $P(e_t = 0) = P(e_t = \frac{1}{2}) = \frac{1}{2}$, 则由文 [11] 中例 4.2.12 知, Z_t 的分布为 $[0, 1]$ 上的均匀分布, 亦即 $[0, 1]$ 上的均匀分布是该模型的不变概率分布, 也称为该模型的平稳分布; 同时, 该模型还是几何遍历的, 进而 $\{Z_t\}$ 是 α -混合的, 此时, 取 $g(z) = \sin(z - \frac{1}{2})$, 则条件 (5) 满足. 当条件 (5) 不成立时, 需附加其它条件, 证明更具技术性, 本文不予讨论.

注 3.3 若加强条件 (2) 为 $\{Z_t\}$ 为独立同分布随机变量序列, 则条件 (3) 中的 α -混合系数可改为满足 $\alpha(l) \leq c_0 \rho^l$, 其中 $0 < \rho < 1$, $c_0 > 0$ 为常数.

本文的主要结果论述如下.

定理 3.1 如果 $g \in C^{k+1}([a, b])$, 并假定上述条件 (1)–(5) 成立, 则有

$$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{L} N(0, \sigma^2 \Gamma_d^{-1}),$$

其中 $\Gamma_d = \text{Cov}(X_t)$ 为向量 $X_t = (X_{t1}, \dots, X_{td})^T$ 的协方差阵.

定理 3.2 如果 $g \in C^{k+1}([a, b])$, 并假定上述条件 (1)–(5) 成立, 那么

$$\|\hat{g} - g\|_2^2 = O_P(n^{r-1}).$$

特别地, 当 $r = \frac{2}{4k+3}$ 时, $\|\hat{g} - g\|_2^2 = O_P(n^{-(4k+1)/(4k+3)})$.

4 定理的证明

由于样条估计 $\hat{\theta}$ 和 \hat{g} 由函数空间 S_{k,N_n} 所唯一确定, 不同的基函数可以给出相同的估计, 因而, 本文主要结果并不依赖特定基的选择. 于是, 为了方便证明, 本节采用基函数 $B_s = K^{1/2}b_s$, $s = 1, \dots, K$, 其中 b_s 为文 [12] 中定义的 B-样条.

为了证明定理, 首先给出几个引理.

引理 4.1 存在正常数 M_1 和 M_2 , 使得 $M_1|\beta|^2 \leq \int \left\{ \sum_s \beta_s B_s(z) \right\}^2 dz \leq M_2|\beta|^2$.

引理 4.1 的证明见文 [12] 中的第 5 章定理 4.2.

引理 4.2 若条件 (1)–(3) 成立, 则存在区间 $[M_3, M_4]$, $0 < M_3 < M_4$, 使得

$$P\{(B^T B)/n \text{ 的特征值落入 } [M_3, M_4]\} \rightarrow 1 \quad (n \rightarrow \infty).$$

引理 4.2 是文 [6] 中引理 2 的特殊情形, 证明略.

引理 4.3 假定 M_m^n 是由一个平稳强混合过程生成的 σ -代数. 令 $\eta_i \in M_{s_i}^{t_i}$, 其中 $s_1 < t_1 < s_2 < t_2 < \dots < t_m$, m 为正整数, 并假定对所有的 i , 都有 $t_i - s_i > \tau$. 进一步假定 $\|\eta_i\|_{p_i}^{p_i} = E|\eta_i|^{p_i} < \infty$, 其中 $p_i > 1$, $Q = \sum_{i=1}^m (\frac{1}{p_i}) < 1$. 那么

$$\left| E \left[\prod_{i=1}^m \eta_i \right] - \prod_{i=1}^m E[\eta_i] \right| \leq 10(m-1)\alpha(\tau)^{1-Q} \prod_{i=1}^m \|\eta_i\|_{p_i},$$

其中 $\alpha(\tau)$ 为混合系数.

证 见 Roussas 和 Ioannides [13], 对于 $m = 2$ 时的结论是众所周知的, 例如, 见 Hall 和 Heyde [14] 中的定理 A.5, 也可参见 Kim 和 Cox [15] 中的引理 5.1.

引理 4.4 假定条件 (1)–(5) 成立, 那么 $\frac{X^T(I-A)X}{n} \xrightarrow{P} \Gamma_d$.

证 令 $\tilde{X}_j = (X_{1j}, X_{2j}, \dots, X_{nj})^T$ ($j = 1, \dots, d$), 则 $X = (\tilde{X}_1, \dots, \tilde{X}_d)$, 于是

$$(X^T(I-A)X)_{ij} = \tilde{X}_i^T(I-A)\tilde{X}_j = \tilde{X}_i^T\tilde{X}_j - \tilde{X}_i^T A \tilde{X}_j \quad (i, j = 1, \dots, d).$$

因为 A 是一个非负定矩阵 (除了当 $n \rightarrow \infty$ 时, 概率趋于零的一个事件以外), 并且 A 也是一个幂等阵, 于是, 有

$$\begin{aligned} E(\tilde{X}_i^T A \tilde{X}_i | A) &= E(\text{tr}(\tilde{X}_i^T A \tilde{X}_i) | A) = \text{tr}(A E(\tilde{X}_i \tilde{X}_i^T | A)) \\ &= O_P(1) \text{tr}(A) \text{Var}(X_{ti}) = O_P(1)K = o_P(n). \end{aligned} \quad (4.1)$$

进而有

$$E(|\tilde{X}_i^T A \tilde{X}_j| | A) \leq \frac{1}{2} \{E(\tilde{X}_i^T A \tilde{X}_i | A) + E(\tilde{X}_j^T A \tilde{X}_j | A)\} = o_P(n). \quad (4.2)$$

因此, 要证引理 4.4 的结论成立, 只需证明

$$(\tilde{X}_i^T \tilde{X}_j)/n \xrightarrow{P} \text{Cov}(X_{ti}, X_{tj}), \quad (4.3)$$

即证 $n^{-1} \sum_{t=1}^n X_{ti} X_{tj} \xrightarrow{P} E(X_{ti} X_{tj})$. 令 $e_{t,ij} = X_{ti} X_{tj} - E(X_{ti} X_{tj})$, 易知 $E e_{t,ij} = 0$. 由引理 4.3 知,

$$\begin{aligned}
 E\left(\sum_{t=1}^n e_{t,ij}\right)^2 &= E\left(\sum_{t=1}^n e_{t,ij}^2\right) + E\left(2 \sum_{1 \leq t_1 < t_2 \leq n} e_{t_1,ij} e_{t_2,ij}\right) \\
 &\leq n E e_{1,ij}^2 + 2n \sum_{t=1}^{\infty} |E(e_{1,ij} e_{t+1,ij})| \leq n E e_{1,ij}^2 + C_1 n \sum_{t=1}^{\infty} \{\alpha(t)\}^{1-2/s_0},
 \end{aligned}$$

其中 C_1 为正常数, $s_0 > 2$. 由条件 (3) 中 $(1 - \frac{2}{s_0})\alpha > 1$ 知, $\sum_{t=1}^{\infty} \{\alpha(t)\}^{1-2/s_0} < \infty$, 所以 $E\left(\frac{1}{n} \sum_{t=1}^n e_{t,ij}\right)^2 = O(n^{-1})$. 于是, 当 $n \rightarrow \infty$ 时, $E\left\{\frac{1}{n} \sum_{t=1}^n X_{ti} X_{tj} - E(X_{ti} X_{tj})\right\}^2 \rightarrow 0$. 从而知, $n^{-1} \sum_{t=1}^n X_{ti} X_{tj} \xrightarrow{P} E(X_{ti} X_{tj})$, 即 (4.3) 成立, 也即

$$\frac{1}{n} (X^T X)_{ij} \xrightarrow{P} \text{Cov}(X_{ti}, X_{tj}). \quad (4.4)$$

于是引理 4.4 得证.

引理 4.5 令 $\{Z_{nk}, k \geq 0\}$ 为一随机变量序列, $\{\mathcal{G}_{n,k-1}, k = 1, \dots, n\}$ 是一个递增 σ -代数序列, 使得 $\{Z_{nk}\}$ 关于 $\mathcal{G}_{n,k}$ 可测, 并且对于 $1 \leq k \leq n$ 有 $E(Z_{nk} | \mathcal{G}_{n,k-1}) = 0$. 如果当 $n \rightarrow \infty$ 时,

(i) $\sum_{k=1}^n E(Z_{nk}^2 | \mathcal{G}_{n,k-1}) \xrightarrow{P} \sigma_n^2$, 其中 $\sigma_n^2 > 0$ 为常数;

(ii) 对每个 $a_0 > 0$, $\sum_{k=1}^n E(Z_{nk}^2 I(|Z_{nk}| > a_0) | \mathcal{G}_{n,k-1}) \xrightarrow{P} 0$. 那么当 $n \rightarrow \infty$ 时,

$$\sum_{k=1}^n Z_{nk} \xrightarrow{L} N(0, \sigma_n^2).$$

证 参见 Pollard [16] 中的第 VIII 章定理 1.

定理 3.1 的证明 令 $\Phi = (g(Z_1), \dots, g(Z_n))^T$, $\mathbb{E} = (\varepsilon_1, \dots, \varepsilon_n)^T$, 显然, 有

$$Y = X\theta + \Phi + \mathbb{E}.$$

于是, 由 (2.5) 知, $\hat{\theta} - \theta = [X^T(I - A)X]^{-1} X^T(I - A)\Phi + [X^T(I - A)X]^{-1} X^T(I - A)\mathbb{E}$. 因而, 为了证明定理 3.1, 只需证明

$$\sqrt{n} [X^T(I - A)X]^{-1} X^T(I - A)\mathbb{E} \xrightarrow{L} N(0, \sigma^2 \Gamma_d^{-1}), \quad (4.5)$$

$$\sqrt{n} [X^T(I - A)X]^{-1} X^T(I - A)\Phi \xrightarrow{P} 0. \quad (4.6)$$

因为对于 $g \in C^{k+1}([a, b])$, 存在样条函数 $g^* \in S_{k, N_n}$ 和常数 $C_2 > 0$, 使得

$$\|g^* - g\|_{\infty} \leq C_2 h^{k+1}, \quad (4.7)$$

其中 $\|g\|_{\infty} = \sup_{x \in D} |g(x)|$ (见 [7, p. 149]). 记 $g^*(z) = \sum_{s=1}^K \beta_s^* B_s(z)$, $\beta^* = (\beta_1^*, \dots, \beta_K^*)^T$, 有 $\Phi^* = B\beta^*$, 其中 $\Phi^* = (g^*(Z_1), \dots, g^*(Z_n))^T$. 注意到 $(I - A)\Phi = (I - A)(\Phi + B\beta^* - B\beta^*) = (I - A)(\Phi - \Phi^*)$. 因此, 由引理 4.4 以及 $|g^* - g| \leq \|g^* - g\|_{\infty} \leq C_2 h^{k+1}$ 和条件 (1) 知,

$$\begin{aligned}
 |n^{-1/2} (X^T(I - A)\Phi)_j| &= |n^{-1/2} \tilde{X}_j^T (I - A)\Phi| \\
 &\leq n^{-1/2} [\tilde{X}_j^T (I - A)\tilde{X}_j]^{1/2} \{(\Phi - \Phi^*)^T (\Phi - \Phi^*)\}^{1/2}
 \end{aligned}$$

$$= \left[\frac{\tilde{X}_j^T(I-A)\tilde{X}_j}{n} \right]^{1/2} \left\{ \sum_{t=1}^n [g^*(Z_t) - g(Z_t)]^2 \right\}^{1/2} \xrightarrow{P} 0,$$

其中 $\tilde{X}_j = (X_{1j}, X_{2j}, \dots, X_{nj})^T$ ($j = 1, \dots, d$), 从而有

$$\sqrt{n} \left[X^T(I-A)X \right]^{-1} X^T(I-A)\Phi = \left[\frac{X^T(I-A)X}{n} \right]^{-1} n^{-1/2} X^T(I-A)\Phi \xrightarrow{P} 0.$$

另一方面, 由于 $\sqrt{n} \left[X^T(I-A)X \right]^{-1} X^T(I-A)\mathbb{E} = \left[\frac{X^T(I-A)X}{n} \right]^{-1} n^{-1/2} X^T(I-A)\mathbb{E}$.

因而, 由引理 4.4 知, 要证明 (4.5) 成立, 只需证明

$$n^{-1/2} X^T(I-A)\mathbb{E} \xrightarrow{L} N(0, \sigma^2 \Gamma_d). \quad (4.8)$$

注意到

$$X^T(I-A)\mathbb{E} = X^T\mathbb{E} - X^T A \mathbb{E} \triangleq I_1 + I_2. \quad (4.9)$$

于是, 只需证明当 $n \rightarrow \infty$ 时,

$$\begin{aligned} n^{-1/2} I_1 &= n^{-1/2} X^T \mathbb{E} \\ &= \left(n^{-1/2} \sum_{t=1}^n X_{t1} \varepsilon_t, \dots, n^{-1/2} \sum_{t=1}^n X_{td} \varepsilon_t \right)^T \xrightarrow{L} N(0, \sigma^2 \Gamma_d), \end{aligned} \quad (4.10)$$

$$I_2 = X^T A \mathbb{E} = o_P(n^{1/2}). \quad (4.11)$$

为证 (4.10), 只需证明对于任意向量 $c = (c_1, \dots, c_d)^T$, 有

$$n^{-1/2} c^T X^T \mathbb{E} \xrightarrow{L} N(0, \sigma^2 c^T \Gamma_d c). \quad (4.12)$$

又由于 $n^{-1/2} c^T X^T \mathbb{E} = n^{-1/2} \sum_{t=1}^n (c_1 X_{t1} + \dots + c_d X_{td}) \varepsilon_t = n^{-1/2} \sum_{t=1}^n W_t \varepsilon_t$, 其中 $W_t = c^T X_t = c_1 X_{t1} + \dots + c_d X_{td}$. 因而, 证明 (4.12) 成立, 等价于证明

$$n^{-1/2} \sum_{t=1}^n W_t \varepsilon_t \xrightarrow{L} N(0, \sigma^2 c^T \Gamma_d c). \quad (4.13)$$

记 \mathcal{G}_t 为由 $\{Y_{1-d}, \dots, Y_0; Z_s, \varepsilon_s, 1 \leq s \leq t\}$ 生成的 σ -代数, 显然, $\{W_t \varepsilon_t\}$ 关于 \mathcal{G}_{t-1} 是鞅差. 注意到 $\frac{1}{n} \sum_{t=1}^n W_t^2 = c^T \left(\frac{1}{n} \sum_{t=1}^n X_t X_t^T \right) c = c^T \left(\frac{X^T X}{n} \right) c$. 由 (4.4) 知, 当 $n \rightarrow \infty$ 时,

$$\frac{1}{n} \sum_{t=1}^n E \{ (W_t \varepsilon_t)^2 \mid \mathcal{G}_{t-1} \} \xrightarrow{P} \sigma^2 c^T \Gamma_d c.$$

另外, 对于任何给定的 $b_0 > 0$,

$$\begin{aligned} & \frac{1}{n} \sum_{t=1}^n E \{ (W_t \varepsilon_t)^2 I(|W_t \varepsilon_t| > b_0 n^{1/2}) \mid \mathcal{G}_{t-1} \} \\ & \leq \frac{1}{n} \sum_{t=1}^n E \{ (W_t \varepsilon_t)^2 I(\varepsilon_t^2 > b_0 n^{1/2}) \mid \mathcal{G}_{t-1} \} + \frac{1}{n} \sum_{t=1}^n E \{ (W_t \varepsilon_t)^2 I(W_t^2 > b_0 n^{1/2}) \mid \mathcal{G}_{t-1} \} \\ & = \frac{1}{n} \sum_{t=1}^n W_t^2 E \{ \varepsilon_t^2 I(\varepsilon_t^2 > b_0 n^{1/2}) \} + \frac{1}{n} \sum_{t=1}^n \sigma^2 W_t^2 I(W_t^2 > b_0 n^{1/2}). \end{aligned}$$

由 (4.4) 知, 上式第一项依概率收敛于 0. 同时, 由条件 (3) 和 (4) 知, 上式第二项也依概率收敛于 0. 从而得 $\frac{1}{n} \sum_{t=1}^n E \{ (W_t \varepsilon_t)^2 I(|W_t \varepsilon_t| > b_0 n^{1/2}) \mid \mathcal{G}_{t-1} \} \xrightarrow{P} 0$. 于是, 由引理 4.5 知, (4.13) 成立.

要证 (4.11) 成立, 只需证明对于 $j = 1, \dots, d$, 有

$$\left(X^T A \mathbb{E}\right)_j = \tilde{X}_j^T A \mathbb{E} = o_P(n^{1/2}). \quad (4.14)$$

令 $A\tilde{X}_j = (r_{j1}, \dots, r_{jn})^T$, 则有

$$E(\tilde{X}_j^T A \mathbb{E})^2 = E\left(\sum_{t=1}^n r_{jt} \varepsilon_t\right)^2 = \sum_{t=1}^n E(r_{jt} \varepsilon_t)^2 = \sigma^2 \sum_{t=1}^n E r_{jt}^2 = \sigma^2 E(\tilde{X}_j^T A \tilde{X}_j).$$

于是, 利用 (4.1) 可得 (4.14). 这就完成了定理 3.1 的证明.

定理 3.2 的证明 记 Φ 与 \mathbb{E} 同定理 3.1, 显然, 有 $Y = X\theta + \Phi + \mathbb{E}$. 由 (2.5) 易知,

$$\hat{\beta} = (B^T B)^{-1} B^T (Y - X\hat{\theta}).$$

令 $\hat{\Phi} = (\hat{g}(Z_1), \dots, \hat{g}(Z_n))^T$, 则 $\hat{\Phi} = B\hat{\beta}$, 其中 $\hat{\beta} = (\hat{\beta}_1, \dots, \hat{\beta}_K)^T$.

令 $\tilde{Y} = (\tilde{Y}_1, \dots, \tilde{Y}_n)^T$, 其中 $\tilde{Y}_t = (\theta - \hat{\theta})^T X_t + g(Z_t)$, 从而, $\tilde{Y} = X(\theta - \hat{\theta}) + \Phi$. 定义

$$\tilde{\beta} = (B^T B)^{-1} B^T \tilde{Y}, \text{ 其中 } \tilde{\beta} = (\tilde{\beta}_1, \dots, \tilde{\beta}_K)^T. \text{ 并令 } \tilde{g}(z) = \sum_{s=1}^K \tilde{\beta}_s B_s(z).$$

下面分别求 $\|\hat{g} - \tilde{g}\|_2$ 和 $\|\tilde{g} - g\|_2$.

由上易知, $\hat{\beta} - \tilde{\beta} = (B^T B)^{-1} B^T \mathbb{E}$. 因为 $\{\varepsilon_t\}$ 为独立同分布序列, 均值为 0, 并且 $\{\varepsilon_t\}$ 与 $\{Z_t\}$ 独立, 所以有 $E\{B_s(Z_t) \varepsilon_t B_s(Z_{t'}) \varepsilon_{t'}\} = 0$ ($t' \neq t$). 于是,

$$E(\mathbb{E}^T B B^T \mathbb{E}) = E\left\{\sum_{s=1}^K \left[\sum_{t=1}^n B_s(Z_t) \varepsilon_t\right]^2\right\} = \sum_{s=1}^K \sum_{t=1}^n E\{B_s(Z_t) \varepsilon_t\}^2 \lesssim nK.$$

因而, $\mathbb{E}^T B B^T \mathbb{E} = O_P(nK)$. 从而由引理 4.2, 有

$$|\hat{\beta} - \tilde{\beta}|^2 = \mathbb{E}^T B (B^T B)^{-1} (B^T B)^{-1} B^T \mathbb{E} \lesssim \frac{1}{n^2} \mathbb{E}^T B B^T \mathbb{E} = O_P\left(\frac{K}{n}\right).$$

又注意到 (2.6), 由引理 4.1 得 $\|\hat{g} - \tilde{g}\|_2^2 \asymp |\hat{\beta} - \tilde{\beta}|^2 = O_P\left(\frac{K}{n}\right)$. 令 $g^* \in S_{k, N_n}$, 满足 (4.7).

此时, $\|g^* - g\|_2 \lesssim \|g^* - g\|_\infty \lesssim h^{k+1}$. 记 $g^*(z) = \sum_{s=1}^K \beta_s^* B_s(z)$, $\beta^* = (\beta_1^*, \dots, \beta_K^*)^T$, $\Phi^* = (g^*(Z_1), \dots, g^*(Z_n))^T$, 显然, 有 $\Phi^* = B\beta^*$. 由引理 4.1 和引理 4.2 知, $\|\tilde{g} - g^*\|_2^2 \asymp |\tilde{\beta} - \beta^*|^2 \asymp \frac{1}{n} (\tilde{\beta} - \beta^*)^T B^T B (\tilde{\beta} - \beta^*)$. 因为 $B\tilde{\beta} = B(B^T B)^{-1} B^T \tilde{Y}$ 为正交投影, 所以,

$$\begin{aligned} \frac{1}{n} (\tilde{\beta} - \beta^*)^T B^T B (\tilde{\beta} - \beta^*) &\leq \frac{1}{n} |\tilde{Y} - B\beta^*|^2 = \frac{1}{n} |X(\theta - \hat{\theta}) + \Phi - \Phi^*|^2 \\ &\leq \frac{2}{n} [|X(\theta - \hat{\theta})|^2 + |\Phi - \Phi^*|^2] = \frac{2}{n} \sum_{t=1}^n \left[\sum_{i=1}^d (\theta_i - \hat{\theta}_i) X_{ti} \right]^2 + \frac{2}{n} \sum_{t=1}^n [g(Z_t) - g^*(Z_t)]^2. \end{aligned}$$

由条件 (2) 知, $E[g(Z_t) - g^*(Z_t)]^2 \asymp \|g - g^*\|_2^2$, 进而, 有

$$\frac{1}{n} \sum_{t=1}^n [g(Z_t) - g^*(Z_t)]^2 = O_P(\|g - g^*\|_2^2).$$

另外, 由于 $\frac{1}{n} \sum_{t=1}^n \left[\sum_{i=1}^d (\theta_i - \hat{\theta}_i) X_{ti} \right]^2 \leq d \sum_{i=1}^d \left\{ (\theta_i - \hat{\theta}_i)^2 \left(\frac{1}{n} \sum_{t=1}^n X_{ti}^2 \right) \right\}$, 又由条件 (4) 知,

$E(X_{ti}^2) < \infty$, 于是, 由定理 3.1 可得 $\frac{1}{n} \sum_{t=1}^n \left[\sum_{i=1}^d (\theta_i - \hat{\theta}_i) X_{ti} \right]^2 = O_P(n^{-1})$. 因而, $\|\tilde{g} - g^*\|_2^2 = O_P(n^{-1} + h^{2(k+1)})$. 又由于 $\|\hat{g} - g\|_2^2 \leq 3 [\|\hat{g} - \tilde{g}\|_2^2 + \|\tilde{g} - g^*\|_2^2 + \|g^* - g\|_2^2]$, 所以, 由条件 (1), 有 $\|\hat{g} - g\|_2^2 = O_P\left(\frac{K}{n} + h^{2(k+1)}\right) = O_P(n^{-1})$. 后一部分结论的证明是显然的.

5 模拟算例

令模型 (1.1) 中的 $d = 1, \theta = 0.5, g(z) = \sin(z), Z_t$ 服从 $[-\frac{\pi}{2}, \frac{\pi}{2}]$ 上的均匀分布, ε_t 服从正态分布 $N(0, 0.2^2)$, 初始值 $Y_0 = 0$. 对不同的样本容量 $n = 100, 200$ 和 400 , 分别产生 201 个样本. 为了便于衡量非参数函数的估计量, 引入 RASE (the square-Root of Average Squared Error) 为

$$\text{RASE} = \left\{ n_{\text{grid}}^{-1} \sum_{m=1}^{n_{\text{grid}}} [\hat{g}(z_m) - g(z_m)]^2 \right\}^{1/2}, \tag{5.1}$$

其中 $\{z_m, m = 1, \dots, n_{\text{grid}}\}$ 为等距格点集, $z_1 = -0.475\pi, z_{n_{\text{grid}}} = 0.475\pi, n_{\text{grid}} = 101$. 在模拟计算中, 取 $N_n = \lfloor n^{2/(4k+3)} \rfloor$, 即 N_n 为 $n^{2/(4k+3)}$ 的整数部分, 在区间 $[-\frac{\pi}{2}, \frac{\pi}{2}]$ 上等距放置结点, 并利用第 2 节所给的样条函数空间的基函数 (2.2).

表 5.1 给出了不同样本容量 n 和 $k = 1, 3$ 的条件下, θ 的样条估计的均值和标准差以及 $g(z)$ 的样条估计的 RASE 的均值和标准差 (括号内为标准差). 由是可见, 参数 θ 的估计量的标准差和非参数函数 $g(z)$ 的估计量的 RASE 随着 n 的增加而减小, 这与上述理论结果是吻合的. 另外, 当样本容量较大 ($n = 200, 400$) 时, 三次样条拟合稍好于一次样条拟合.

表 5.1 θ 的样条估计的均值和标准差与 $g(z)$ 的样条估计的 RASE 的均值和标准差 (括号内)

n	$k = 1$		$k = 3$	
	$\theta = 0.5$	$g(z)$	$\theta = 0.5$	$g(z)$
100	0.4982 (0.0230)	0.0456 (0.0128)	0.4984 (0.0235)	0.0407 (0.0127)
200	0.4996 (0.0182)	0.0339 (0.0091)	0.4994 (0.0166)	0.0308 (0.0091)
400	0.4990 (0.0116)	0.0267 (0.0067)	0.5005 (0.0116)	0.0212 (0.0065)

图 5.1 分别给出了 $n = 100, 200$ 和 400 时非参数函数 $g(z)$ 的 201 条三次样条估计曲线, 图 5.2 为图 5.1 中的典型估计曲线 (对应于 201 个 RASE 中取中值的那组样本数据所得的估计曲线), 其中实线为真实函数曲线, 点线为估计函数曲线. 由图 5.1 和图 5.2 可以看出, 随着样本容量 n 的增加, 估计误差趋于减小, 同时还表明, 甚至在样本容量较小 (如 $n = 100$) 时, 也能得到较为满意的估计.

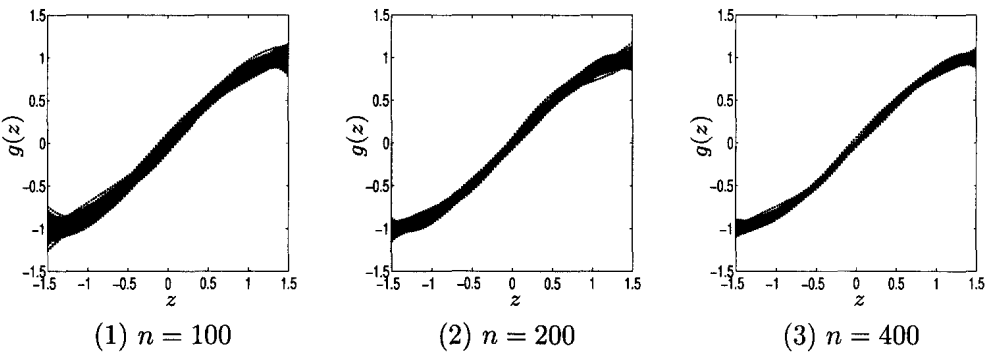


图 5.1 当 $n = 100, 200$ 和 400 时, 函数 $g(z)$ 的 201 条估计曲线

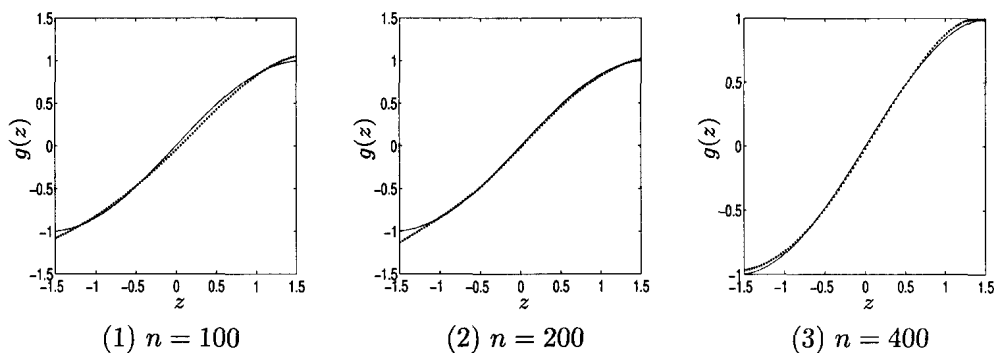


图 5.2 当 $n = 100, 200$ 和 400 时, 函数 $g(z)$ 的典型估计曲线

6 结束语

本文在适当的条件下, 利用多项式样条估计方法给出了具有外生变量的部分线性自回归模型 (1.1) 中参数向量估计的渐近正态性以及非参数函数估计的收敛速度. 模拟例子验证了所述的理论结果. 文中所用样条估计方法具有一定的优点: 它能给出所拟合模型的简单的显式表达式, 计算量较小. 但是, 在具体应用中, 需要考虑结点个数的选择和结点位置的放置等问题, 这有待作进一步的研究.

参 考 文 献

- [1] Bosq D. and Shen J., Estimation of an autoregressive semiparametric model with exogenous variables [J], *J. Statist. Plann. Inference*, 1998, 68(1):105–127.
- [2] Hannan E. J. and Deistler M., The Statistical Theory of Linear Systems [M], New York: Wiley, 1988.
- [3] Liang Hua, Second order asymptotic efficiency in a partly autoregressive model [J], *J. Systems Sci. Math. Sci.*, 1996, 9(2):164–170.
- [4] Zhu Lixing and An Hongzhi, A note on the strong consistency of estimates in partially linear models [J], *Acta Math. Sci.*, 1994, 14(2):146–152.
- [5] Fan J., Local linear regression smoothers and their minimax efficiencies [J], *Ann. Statist.*, 1993, 21(1):196–216.
- [6] Huang J. Z. and Shen H., Functional coefficient regression models for nonlinear time series: A polynomial spline approach [J], *Scand. J. Statist.*, 2004, 31(4):515–534.
- [7] de Boor C., A Practical Guide to Splines [M], New York: Springer-Verlag, 2001.
- [8] Shi Peide, M -type regression splines involving time series [J], *J. Statist. Plann. Inference*, 1997, 61(1):17–37.
- [9] Gao J. and Yee T., Adaptive estimation in partially linear autoregressive models [J], *Canad. J. Statist.*, 2000, 28(3):571–586.
- [10] Härdle W. and Vieu P., Kernel regression smoothing of time series [J], *J. Time Ser. Anal.*, 1992, 13(3):209–232.

- [11] 安鸿志, 陈敏, 非线性时间序列分析 [M], 上海: 上海科学技术出版社, 1998.
- [12] deVore R. A. and Lorentz G. G., Constructive Approximation [M], Berlin Heidelberg: Springer-Verlag, 1993.
- [13] Roussas G. G. and Ioannides D., Moment inequalities for mixing sequences of random variables [J], *Stochastic Anal. Appl.*, 1987, 5(1):61–120.
- [14] Hall P. and Heyde C., Martingale Limit Theory and Its Applications [M], New York: Academic Press, 1980.
- [15] Kim T. Y. and Cox D. D., Asymptotic behaviors of some measures of accuracy in nonparametric curve estimation with dependent observations [J], *J. Multivariate Anal.*, 1995, 53(1):67–93.
- [16] Pollard D., Convergence of Stochastic Processes [M], New York: Springer, 1984.

Spline Estimation for Partially Linear Autoregressive Models with an Exogenous Variable

WU Xinqian* TIAN Zheng** HAN Sier*

*Department of Applied Mathematics, Northwestern Polytechnical University, Xi'an 710072, China. E-mail: wuxinqian1001@163.com; hansier2003@126.com

**Department of Applied Mathematics, Northwestern Polytechnical University, Xi'an 710072, China; National Key Laboratory of Pattern Recognition, and Institute of Automation, Chinese Academy of Sciences, Beijing 100080, China. E-mail: zhtian@nwpu.edu.cn

Abstract Consider the autoregressive model $Y_t = \theta^T X_t + g(Z_t) + \varepsilon_t$, $t = 1, \dots, n$, where $X_t = (Y_{t-1}, \dots, Y_{t-d})^T$, Z_t is a real value exogenous random variable, $\theta = (\theta_1, \dots, \theta_d)^T$ is an unknown parameter vector to be estimated and g is an unknown nonparametric smooth function. Based on polynomial spline estimation, asymptotic normality of estimator of θ is given and convergence rate of estimator of g is obtained under some mild conditions. Main results are illustrated by a simulation example.

Keywords Exogenous variable, Partially linear autoregressive model, Spline estimation, Asymptotic normality, Convergence rate

2000 MR Subject Classification 62G05, 62G08, 62M10