Vol. 45 ( 2025 )
No. 4

数 学 杂 志
J. of Math. (PRC)

# 半参数变系数空间误差回归模型的惩罚经验似然

王以恒, 何帮强
(安徽工程大学数理与金融学院, 安徽 芜湖 241000)

**摘 要**: 文章研究半参数变系数空间误差回归模型下的参数估计和变量选择问题. 利用了局部多项式的方法对变系数函数进行估计, 其次分别构造参数部分和非参数部分的最大经验对数似然比估计, 同时使用惩罚经验似然 (PEL) 进行选择变量, 并且用平方再求和的方法估计空间系数及误差项的方差. 得到了参数与非参数的估计值, 以及 PEL 方法选择变量的优越性的结论. 在合适的条件下, 推广了惩罚经验似然估计具有 Oracle 特征且在零假设下服从渐近卡方分布.

**关键词**: 部分线性模型; 惩罚经验似然; 空间自回归; 变量误差

MR(2010) **主题分类号**: 62G05; 62E20 **中图分类号**: O212.7

**文献标识码**: A **文章编号**: 0255-7797(2025)04-0337-12

## 1 引言

半参数变系数模型在金融, 经济等数据分析凭借良好的解释能力得到广泛应用. 具有非参数模型的灵活和参数模型易解释的优点, 从而成为近年来用以研究变量间非线性关系的最有效模型之一. 经济学、环境科学、地理学、流行病学等领域的数据都存在空间相依性或空间异质性, 这将导致不同空间尺度的响应变量观测值之间的内生相互作用效应. 空间相关性的存在就会让模型变得比较复杂, 所以如何对模型存在的空间相关性进行有效的处理成为后面的考虑对象. 半参数变系数空间自回归模型就是一个非常重要的模型, 由于其灵活性和可解释性, 越来越受到关注. 例如, Wei 等 [1] 利用剖面极大似然方法研究了半参数变系数空间自回归模型的统计推断; Luo 等 [2] 提出的半参数变系数空间自回归模型的经验似然推断; Liang 等 [3] 考虑了在固定效应和时变系数下半参数空间自回归面板数据模型的推断; 以及 Su[4] 研究的半参数空间自回归模型的 GMM 估计.

空间误差未被纳入模型时, 可能导致参数估计的偏误和不一致性. 针对此类问题 Zhang 等 [5] 研究了具有空间误差和未知异方差的部分线性可加高阶空间自回归下的统计推断; Su 和 Jin[6] 提出的半参数滞后回归模型, 在刻画空间滞后因变量引起的空间相关性特征方面意义鲜明; 陈建宝 [7] 的半参数变系数空间误差回归模型在捕捉非线性特征和处理空间效应两方面的表现都就较为优越. 本文考虑带空间误差的半参数变系数模型:

$$\begin{cases} Y_i = Z_i^{\mathrm{T}}\beta + X_i^{\mathrm{T}}\alpha\left(U_i\right) + m_i, \\ m_i = \rho_0 \sum\limits_{j=1}^{n} w_{i,j} m_j + \varepsilon_i, \end{cases} \tag{1}$$

*收稿日期: 2024-10-14 接收日期: 2024-11-04

基金项目: "行政记录人口普查"的数据质量评估框架研究 (21CTJ005); 金融资产泡沫的形成、识别及其价格有效性测度研究 (72271003).

作者简介: 王以恒 (2000–), 男, 安徽亳州, 研究生, 研究方向为: 数理统计.

通讯作者: 何帮强 (1974–), 男, 教授, 研究方向为: 数理统计.

其中 $Y$ 是响应变量,$Z\in R^{P},X\in R^{q},U$ 是协变量,$\beta=(\beta_1,\beta_2,\cdots,\beta_p)^{\mathrm{T}}$ 是 $p$ 维的未知参数向量, $\alpha(\cdot)=(\alpha_1(\cdot),\alpha_2(\cdot),\cdots,\alpha_q(\cdot))^{\mathrm{T}}$ 是 q 维的未知函数向量, 误差 $\varepsilon_i$ 是均值为 0, 方差为 $\sigma^2$ 的独立同分布的随机变量. $\rho_0$ 是真实的空间相关相关系数, 反映的是误差项的自相关关系. $m_i$ 是具有空间误差的误差项.

自从 $Owen$[8] 提出经验似然法进行构造置信区间以来, 该方法受到了广泛的关注. 经验似然是一种基于样本数据构建的似然函数, 为样本提供一个权重分配, 使得在样本的约束条件下, 通过最大化似然函数, 能够反映样本的分布特性. 基于经验似然的置信区域不需要对区域形状施加先验约束. 众所周知, 高维数据分析在许多当代统计研究中经常出现. 各种惩罚方法已被开发用于变量选择. Tang 和 Leng(2010)[9] 首次引入惩罚经验似然 (PEL), 用于分析多变量的平均向量具有发散参数的线性模型的分析和回归系数. Chen 等 [10] 考虑在纵向数据下高维广义线性模型的惩罚经验似然；He 等 [11] 利用 PEL 对带固定效应含误差变量半参数高维面板数据模型进行降维和参数估计. Wang 等 [12] 利用惩罚经验方法研究 GINAR(p) 模型. 本文在陈建宝 [7] 的半参数变系数空间误差回归模型的基础之上构造了参数以及非参数部分的估计, 然后又对参数部分的随机辅助向量构建了经验似然比函数, 最后通过惩罚经验似然的方法进行了变量选择, 并证明在给出的条件下, 惩罚经验似然估计具有 Oracle 特征且在零假设下服从渐近卡方分布.

## 2 模型与方法

将模型 (1) 写成矩阵形式如下：

$$\begin{cases} Y=Z^{\mathrm{T}}\beta+X^{\mathrm{T}}\alpha(u)+m, \\ m=\rho Wm+\varepsilon, \end{cases} i=1,\cdots,n\,, \tag{2}$$

第一步, 利用局部多项式的方法对变系数函数 $\alpha(u_i)$ 进行估计, 在任给的 $U$ 邻域内一点 $u$, 有

$$\alpha_j(U_i)\approx\alpha_j(u)+\alpha_j'(u)(U_i-u)=a_j+b_j(U_i-u)\,,j=1,2,\cdots,q$$

从而系数函数可以通过极小化下式估计

$$\frac{1}{n}\sum_{i=1}^{n}\left\{Y_i-{Z_i}^{\mathrm{T}}\beta-X_i^{\mathrm{T}}\left[a_j+b_j(U_i-u)\right]\right\}^2K_h(U_i-u)\,, \tag{3}$$

其中 $K_h()=K(/h)/h$ 是核函数, $h$ 是带宽, 为了书写方便, 引入以下记号

$$X=\begin{pmatrix} X_1 & (U_1-u)X_1^{\mathrm{T}}/h \\ \vdots & \vdots \\ X_n & (U_n-u)X_n^{\mathrm{T}}/h \end{pmatrix},$$

$K=diag(K_h(U_1-u),K_h(U_2-u),\cdots,K_h(U_n-u))$,$M=(X_1\alpha(U_1),X_2\alpha(U_2),\cdots,X_n\alpha(U_n))$, 则 (3) 式的解为

$$\begin{pmatrix} \hat{a}_j \\ h\hat{b}_j \end{pmatrix}=\left(X^{\mathrm{T}}KX\right)^{-1}X^{\mathrm{T}}K\left(Y-Z^{\mathrm{T}}\beta\right),$$

记 $S=e_1^{\mathrm{T}}\left(X^{\mathrm{T}}KX\right)^{-1}X^{\mathrm{T}}K, e_1=(1,0)^{\mathrm{T}}$ 则

$$\tilde{M}=X^{\mathrm{T}}\hat{\alpha}(u)=S\left(Y-Z^{\mathrm{T}}\beta\right), \tag{4}$$

第二步, 估计 $\beta$, 令 $H=(H_1,H_2,\cdots,H_n)^{\mathrm{T}}$, $\tilde{Y}=(I-S)Y$, $\tilde{Z}=(I-S)Z$, $H$ 是工具变量, 满足 $E\left({H_n}^{\mathrm{T}}m\right)=0$, 我们使用 $G=(I-S)H$ 作为 $\tilde{Z}$ 的工具变量矩阵, 来定义 $\beta$ 的最大经验似然比估计

$$E\left(G^{\mathrm{T}}m\right)=E\left[G^{\mathrm{T}}\left(Y-Z^{\mathrm{T}}\beta-\tilde{M}\right)\right]=E\left[G^{\mathrm{T}}\left(\tilde{Y}-\tilde{Z}^{\mathrm{T}}\beta\right)\right]=0, \tag{5}$$

引入 $\beta$ 的辅助随机向量

$$\xi_i(\beta)=G(\beta)\left(\tilde{Y}_i-\tilde{Z}_i^{\mathrm{T}}\beta\right), \tag{6}$$

定义 $\beta$ 的经验对数似然比函数：

$$L(\beta)=-\max\left\{\sum_{i=1}^n\log(np_i)\middle|p_i\geq 0,\sum_{i=1}^n p_i\xi_i(\beta)=0,\sum_{i=1}^n p_i=1\right\},$$

由 Lagrange 乘数法 $L(\beta)$ 可以表示为

$$L(\beta)=\sum_{i=1}^n\log\left\{1+\lambda^{\mathrm{T}}\xi_i(\beta)\right\}, \tag{7}$$

其中 $\lambda$ 是下面方程的解：

$$\frac{1}{n}\sum_{i=1}^n\frac{\xi_i(\beta)}{1+\lambda^{\mathrm{T}}\xi_i(\beta)}=0, \tag{8}$$

为了达到变量选择的目的, 在经验对数似然比的基础上添加惩罚函数, 得到

$$L_p(\beta)=\sum_{i=1}^n\log\left\{1+\lambda^{\mathrm{T}}\xi_i(\beta)\right\}+n\sum_{j=1}^p p_\tau(|\beta_j|),$$

其中 $p_\tau(\cdot)$ 是惩罚函数, 他的一阶导满足

$${p_\tau}'(t)=\tau\left\{I(t\leq\tau)+\frac{(a\tau-t)}{(a-1)\tau}I(t>\tau)\right\},$$

其中 $t>0$, $a>2$ 且 $\tau>0$ 是调整参数, 本文选择 Fan[13] 建议的 $a=3.7$ .

假设 $\mathrm{A}=\{j:\beta_{0j}\neq 0\}$ 代表由真实参数 $\beta_0$ 的非零元素组成的集合. 我们将参数向量划分为 $\beta=\left({\beta_1}^{\mathrm{T}},{\beta_2}^{\mathrm{T}}\right)^{\mathrm{T}}$, 其中 $\beta_1\in R^d,\beta_2\in R^{p-d}$ 代表非零元素与零元素, 那么真实参数 $\beta_0=\left(\beta_{10}^{\mathrm{T}},\beta_{20}^{\mathrm{T}}\right)=\left(\beta_{10}^{\mathrm{T}},0^{\mathrm{T}}\right)^{\mathrm{T}}$. 类似的, 记 $\hat{\beta}=\left(\hat{\beta}_1^{\mathrm{T}},\hat{\beta}_2^{\mathrm{T}}\right)^{\mathrm{T}}$, 其中 $\hat{\beta}_1,\hat{\beta}_2$ 分别是 $\beta_1,\beta_2$ 的惩罚经验似然估计值. 令

$$\Sigma_1=E\left[H_i{H_i}^{\mathrm{T}}\right]-E\left[\Phi^{\mathrm{T}}(U_i)\Gamma^{-1}(U_i)\Phi(U_i)\right],$$

$$\Sigma_2=\left\{E\left[H_i{H_i}^{\mathrm{T}}\right]-E\left[\Phi^{\mathrm{T}}(U_i)\Gamma^{-1}(U_i)\Phi(U_i)\right]\right\}\sigma^2(I-\rho W)^{-2},$$

其中 $\Gamma(u)=E\left(X_iX_i{}^{\mathrm{T}}|U=u\right)$, $\Phi(u)=E\left(X_iH_i{}^{\mathrm{T}}|U=u\right)$ , $\Sigma=\Sigma_1^{-1}\Sigma_2\Sigma_1^{-1}$, 对应的, 将 $\Sigma$ 划分为矩阵块 $\Sigma_{ij}\,(i=1,2;j=1,2)$.

第三步, 估计空间系数 $\rho$ 和误差项的方差 $\sigma^2$, 记 $\bar{m}_n=W_nm_n,\bar{\bar{m}}_n=W_n^2m_n,\bar{\varepsilon}_n=W_n\varepsilon_n$, 则有 $m_n=\rho\bar{m}_n+\varepsilon_n$, $\bar{m}_n=\rho\bar{m}_n+\bar{\varepsilon}_n$ 所以有

$$m_n-\lambda\bar{m}_n=\varepsilon_n,\quad m_n-\bar{\lambda}m_n=\bar{\varepsilon}_n. \tag{9}$$

对 (9) 式平方再求和得

$$\begin{aligned}
&\tfrac{1}{n}\sum_{i=1}^{n}m_n^2=\tfrac{2\rho}{n}\sum_{i=1}^{n}m_n\bar{m}_n-\tfrac{\rho^2}{n}\sum_{i=1}^{n}\bar{m}_n^2+\tfrac{1}{n}\sum_{i=1}^{n}\varepsilon_n^2,\\
&\tfrac{1}{n}\sum_{i=1}^{n}\bar{m}_n^2=\tfrac{2\rho}{n}\sum_{i=1}^{n}\bar{m}_n\bar{\bar{m}}_n-\tfrac{\rho^2}{n}\sum_{i=1}^{n}\bar{\bar{m}}{}^2{}_n+\tfrac{1}{n}\sum_{i=1}^{n}\bar{\varepsilon}_n^2,\\
&\tfrac{1}{n}\sum_{i=1}^{n}m_n\bar{m}_n=\tfrac{\rho}{n}\sum_{i=1}^{n}\left(m_n\bar{\bar{m}}_n+\bar{\bar{m}}{}^2{}_n\right)-\tfrac{\rho^2}{n}\sum_{i=1}^{n}\bar{m}_n\bar{\bar{m}}_n+\tfrac{1}{n}\sum_{i=1}^{n}\varepsilon_n\bar{\varepsilon}_n,
\end{aligned}$$

且对应的期望分别是 $E\left(\frac{1}{n}\sum\limits_{i=1}^{n}\varepsilon_n^2\right)=\sigma^2,E\left(\frac{1}{n}\sum\limits_{i=1}^{n}\bar{\varepsilon}_n^2\right)=\frac{\sigma^2}{n}tr\left(W_n{}^{\mathrm{T}}W_n\right),E\left(\frac{1}{n}\sum\limits_{i=1}^{n}\varepsilon_n\bar{\varepsilon}_n\right)=0$, 令 $\Theta$=$(\rho,\rho^2,\sigma^2)^{\mathrm{T}}$, 则有

$$\Lambda_n\Theta=C_n, \tag{10}$$

$$\Lambda_n=\frac{1}{n}\begin{pmatrix}2E\left(m_n\bar{m}_n\right) & -E\left(\bar{m}_n^{\mathrm{T}}\bar{m}_n\right) & 1\\ 2E\left(\bar{m}_n\bar{\bar{m}}_n\right) & -E\left(\bar{\bar{m}}_n^{\mathrm{T}}\bar{\bar{m}}_n\right) & tr\left(W_n^{\mathrm{T}}W_n\right)\\ 2E\left(m_n\bar{\bar{m}}_n+\bar{\bar{m}}{}^2{}_n\right) & -E\left(\bar{m}_n^{\mathrm{T}}\bar{\bar{m}}_n\right) & 0\end{pmatrix},C_n=\frac{1}{n}\begin{pmatrix}E\left(m_n^{\mathrm{T}}m_n\right)\\ E\left(\bar{m}_n^{\mathrm{T}}\bar{m}_n\right)\\ E\left(m_n^{\mathrm{T}}\bar{m}_n\right)\end{pmatrix},$$

则 $\Theta$ 的估计

$$\hat{\Theta}=\Lambda_n^{-1}C_n, \tag{11}$$

一般情况下, $\Lambda_nC_n$ 是未知的, 所以用到两步估计方法来估计 $\Theta$, 此时用 $m_n$ 的估计值 $\tilde{m}_n$ 估计 $\Theta$, 其中 $\tilde{m}_n=\tilde{Y}_n-\tilde{Z}_n^{\mathrm{T}}\beta$, 令 $\tilde{\bar{m}}_n=W_n\tilde{m}_n,\tilde{\bar{\bar{m}}}_n=W_n^2\tilde{m}_n$, 则有

$$F_n=\frac{1}{n}\begin{pmatrix}2E\left(\tilde{m}_n\tilde{\bar{m}}_n\right) & -E\left(\tilde{\bar{m}}_n^{\mathrm{T}}\tilde{\bar{m}}_n\right) & 1\\ 2E\left(\tilde{\bar{m}}_n\tilde{\bar{\bar{m}}}_n\right) & -E\left(\tilde{\bar{\bar{m}}}_n^{\mathrm{T}}\tilde{\bar{\bar{m}}}_n\right) & tr\left(W_n^{\mathrm{T}}W_n\right)\\ 2E\left(\tilde{m}_n\tilde{\bar{\bar{m}}}_n+\tilde{\bar{m}}{}^2{}_n\right) & -E\left(\tilde{\bar{m}}_n^{\mathrm{T}}\tilde{\bar{\bar{m}}}_n\right) & 0\end{pmatrix},d_n=\frac{1}{n}\begin{pmatrix}E\left(\tilde{m}_n^{\mathrm{T}}\tilde{m}_n\right)\\ E\left(\tilde{\bar{m}}_n^{\mathrm{T}}\tilde{\bar{m}}_n\right)\\ E\left(\tilde{m}_n^{\mathrm{T}}\tilde{\bar{m}}_n\right)\end{pmatrix},$$

根据 (11) 式, 有

$$d_n=F_n\Theta+v_n, \tag{12}$$

这里 $v_n$ 是估计偏差, 对残差平方和取最小值计算 $\Theta$ 的估计有

$$\tilde{\Theta}=\arg\min_{\Theta}\left(d_n-F_n\Theta\right)^{\mathrm{T}}\left(d_n-F_n\Theta\right)=\left(F_n{}^{\mathrm{T}}F_n\right)^{-1}F_n{}^{\mathrm{T}}d_n, \tag{13}$$

## 3 主要结论

在讨论样本性质前, 先给出一些正则性条件, 记 $a_n=(p/n)^{1/2},c_n=\{\log n/nh\}^{1/2}+h^2$,
C1. 当 $|\rho|<1$ 时, 矩阵 $(I-\rho W)$ 是非奇异矩阵.

C2. 当 $|\rho|<1$ 时, 矩阵 $W$ 和 $(I-\rho W)^{-1}$ 的绝对行和与绝对列和一致有界.
C3. 误差序列 $\varepsilon$ 完全独立, 且满足: $E(\varepsilon)=0, E(\varepsilon)^2=\sigma^2$.
C4. $\Gamma(u)$ 和 $\Phi(u)$ 是 Lipschitz 连续的, $\Gamma(u)=E\left(X_i X_i{}^{\mathrm{T}}|U=u\right)$ 是非奇异的, 以及 $\Phi(u)=E\left(X_i H_i{}^{\mathrm{T}}|U=u\right)$, $\mu_j=\int u^j K(u)\,du$, $v_j=\int u^j K^2(u)\,du$.
C5. $\{\alpha_j(\cdot), j=1,2\cdots,q\}$ 具有连续的二阶导数.
C6. 核密度函数 $K(\cdot)$ 是具有紧支撑的的对称密度函数且 Lipschitz 连续, 存在 $d<2-r^{-1}$, 且窗宽 $h$ 满足: $nh^6\to 0, nh^3/(\log n)^3\to\infty$ 和 $n^{2d-1}h\to\infty$.
C7. $\Sigma_1$ 和 $\Sigma_2$ 为正定矩阵, 且其特征值均有界.
C8. 当 $n\to\infty, \tau(n/p)^{1/2}\to\infty$ 和 $\min_{j\in\mathrm{A}}\beta_{j0}/\tau\to\infty$.
C9. 惩罚函数 $p'_\tau(\cdot)$ 满足 $\max_{j\in\mathrm{A}}p'_\tau(|\beta_{j0}|)=o\left((np)^{-1/2}\right)$ 和 $\max_{j\in\mathrm{A}}p''_\tau(|\beta_{j0}|)=o\left(p^{-1/2}\right)$.

**引理 1** 假设条件 C1-C6 成立, 那么

$$
\begin{aligned}
&X^{\mathrm{T}}KX=nf(u)\,\Gamma(u)\otimes\begin{pmatrix}1&0\\0&\mu_2\end{pmatrix}(1+O_p(c_n)),\\
&X^{\mathrm{T}}KH=nf(u)\,\Phi(u)\otimes(1,0)^{\mathrm{T}}(1+O_p(c_n)),
\end{aligned}
$$

**证** 由 Fan[14] 引理 A.2 可得.

**引理 2** 假设条件 C1-C7 成立, 那么

$$
\frac{1}{n}G^{\mathrm{T}}(\beta)(I-S)Z_i=\Sigma_1+o_p(1), \tag{14}
$$

$$
\frac{1}{n}G^{\mathrm{T}}(\beta)(I-S)(Y-Z\beta)=\Psi_n+o_p\left(c_n^2\right), \tag{15}
$$

**证** 由引理 1, 可以得到

$$
\left(X_i{}^{\mathrm{T}},0_{1\times q}\right)\left\{X^{\mathrm{T}}(U_1)W_{hu_1}X(U_1)\right\}^{-1}X^{\mathrm{T}}(U_1)W_{hu_1}H=X_i{}^{\mathrm{T}}\Gamma^{-1}(U_i)\,\Phi(U_i)\{1+O_p(c_n)\},
$$

所以有以下推导

$$
\begin{aligned}
&\tfrac{1}{n}H^{\mathrm{T}}(I-S)^{\mathrm{T}}(I-S)Z_i=\tfrac{1}{n}H^{\mathrm{T}}(I-S)^{\mathrm{T}}(I-S)(H+o_p(1))\\
=&\tfrac{1}{n}\sum_{i=1}^{n}\left\{H_i-\Phi^{\mathrm{T}}(U_i)\,\Gamma^{-1}(U_i)\,X_i\left[1+O_p(c_n)\right]\right\}^2\\
=&\tfrac{1}{n}\sum_{i=1}^{n}\left\{\left[H_i-\Phi^{\mathrm{T}}(U_i)\,\Gamma^{-1}(U_i)\,X_i\right]+O_p(c_n)\,\Phi^{\mathrm{T}}(U_i)\,\Gamma^{-1}(U_i)\,X_i\right\}^2\\
=&\tfrac{1}{n}\sum_{i=1}^{n}\left\{\left[H_i-\Phi^{\mathrm{T}}(U_i)\,\Gamma^{-1}(U_i)\,X_i\right]\right\}^2\\
&+O_p(c_n)\,\tfrac{1}{n}\sum_{i=1}^{n}\left\{\left[H_i-\Phi^{\mathrm{T}}(U_i)\,\Gamma^{-1}(U_i)\,X_i\right]\Phi^{\mathrm{T}}(U_i)\,\Gamma^{-1}(U_i)\,X_i\right\}\\
&+O_p\left(c_n{}^2\right)\tfrac{1}{n}\sum_{i=1}^{n}\left\{\Phi^{\mathrm{T}}(U_i)\,\Gamma^{-1}(U_i)\,X_iX_i{}^{\mathrm{T}}\Gamma^{-1}(U_i)\,\Phi(U_i)\right\}=\Sigma_1+o_p(1),
\end{aligned}
$$

$$
\frac{1}{n}G^{\mathrm{T}}(\beta)(I-S)(Y-Z\beta)=\frac{1}{n}G^{\mathrm{T}}(\beta)(I-S)(M+m),
$$

由引理 1 得到 $\left(X_i{}^{\mathrm{T}}, 0_{1\times q}\right)\left\{X^{\mathrm{T}}\left(U_1\right)W_{hu_1}X\left(U_1\right)\right\}^{-1}X^{\mathrm{T}}\left(U_1\right)W_{hu_1}M = X_i{}^{\mathrm{T}}\alpha\left(U_i\right)\left(1+O_p\left(c_n\right)\right)$, 由此可以得到下面式子

$$
\begin{aligned}
&\tfrac{1}{n}H^{\mathrm{T}}(I-S)^{\mathrm{T}}\left(I-S\right)M\\
&=\tfrac{1}{n}\sum_{i=1}^{n}\left[H_i-\Phi^{\mathrm{T}}\left(U_i\right)\Gamma^{-1}\left(U_i\right)X_i\left(1+O_p\left(c_n\right)\right)\right]\left[X_i{}^{\mathrm{T}}\alpha\left(U_i\right)-X_i{}^{\mathrm{T}}\alpha\left(U_i\right)\left(1+O_p\left(c_n\right)\right)\right]\\
&=\tfrac{1}{n}\sum_{i=1}^{n}\left(H_i-\Phi^{\mathrm{T}}\left(U_i\right)\Gamma^{-1}\left(U_i\right)X_i\right)X_i{}^{\mathrm{T}}\alpha\left(U_i\right)\left(1+O_p\left(c_n\right)\right)O_p\left(c_n\right)\\
&=O_p\left(c_n^2\right),
\end{aligned}
$$

同理 $\left(X_i{}^{\mathrm{T}}, 0_{1\times q}\right)\left\{X^{\mathrm{T}}\left(U_1\right)W_{hu_1}X\left(U_1\right)\right\}^{-1}X^{\mathrm{T}}\left(U_1\right)W_{hu_1}m = O_p\left(c_n\right)$, 所以有下式

$$
\frac{1}{n}H^{\mathrm{T}}(I-S)^{\mathrm{T}}\left(I-S\right)m=\frac{1}{n}\sum_{i=1}^{n}\left[H_i-\Phi^{\mathrm{T}}\left(U_i\right)\Gamma^{-1}\left(U_i\right)X_i\right]m_i\left(1+O_p\left(c_n\right)\right),
$$

此处定义 $\psi=\frac{1}{n}\sum_{i=1}^{n}\left[H_i-\Phi^{\mathrm{T}}\left(U_i\right)\Gamma^{-1}\left(U_i\right)X_i\right]\left(1+O_p\left(c_n\right)\right)\left(I-\rho W\right)^{-1}\varepsilon_i$, 所以有 $\frac{1}{n}G^{\mathrm{T}}\left(\beta\right)\left(I-S\right)\left(Y-Z\beta\right)=\psi_n+O_p\left(c_n{}^2\right)$, 那么 (15) 式成立.

综合以上式子引理 2 证明完成.

**引理 3** 假设条件 C1-C8 成立, 那么

$$
1/\sqrt{n}\sum_{i=1}^{n}\xi\left(\beta\right)\to N\left(0,\Sigma_2\right), \tag{16}
$$

$$
\max_{1\le i\le n}\left\|\xi\left(\beta\right)\right\|=o_p\left(n^{1/2}\right), \tag{17}
$$

**证**

$$
\begin{aligned}
1/\sqrt{n}\sum_{i=1}^{n}\xi\left(\beta\right)&=1/\sqrt{n}\sum_{i=1}^{n}G_i\left(\beta\right)\left(\tilde{Y}_i-\tilde{Z}_i^{\mathrm{T}}\beta\right)=1/\sqrt{n}\sum_{i=1}^{n}G_i\left(\beta\right)\left(I-S\right)\left(Y-Z^{\mathrm{T}}\beta\right)\\
&=\sqrt{n}\left(\psi_n+O_p\left(c_n{}^2\right)\right),
\end{aligned}
$$

因为 $E\left(\psi_n\right)=0$, $Cov\left(\psi_n\right)=\frac{1}{n}\sigma^2\Sigma_1(I-\rho W)^{-2}$, 其中 $\Sigma_2=\sigma^2\Sigma_1(I-\rho W)^{-2}$ 则 (16) 式成立.

$$
\max_{1\le i\le n}\left\|\xi\left(\beta\right)\right\|\le\max_{1\le i\le n}\left\|G_i\left(\beta\right)\right\|\cdot\max_{1\le i\le n}\left\|\tilde{Y}_i-\tilde{Z}_i\beta\right\|,
$$

由假设 $E\|X\|^{2s}\le\infty, E\|Z\|^{2s}\le\infty, E\|H\|^{2s}\le\infty, E\|U\|^{2s}\le\infty, E\|\varepsilon\|^{2s}\le\infty$, 得

$$
\max_{1\le i\le n}\left\|G_i\left(\beta\right)\right\|=o\left(n^{1/2s}\right),\max_{1\le i\le n}\left\|\tilde{Y}_i-\tilde{Z}_i\beta\right\|=o\left(n^{1/2s}\right)
$$

(17) 式成立.

**引理 4** 假设条件 C1-C7 成立, 则 $\|\lambda\|=O_p\left(a_n\right)$.

**证** 设 $\lambda=\alpha\delta$, 其中 $\alpha\geq 0\delta\in R^p$ 且 $\|\delta\|=1$, 并且令 $S(\beta)=\frac{1}{n}\sum\limits_{i=1}^{n}\xi_i(\beta)\xi_i{}^{\mathrm{T}}(\beta)$, $\bar{\xi}_i(\beta)=\frac{1}{n}\sum\limits_{i=1}^{n}\xi_i(\beta)$, $\xi_i{}^*(\beta)=\max\limits_{1\leq i\leq n}\|\xi_i(\beta)\|$, 由 (17) 式 $\xi_i{}^*(\beta)=O_p(a_n)$.

$$
\begin{aligned}
0&=\frac{1}{n}\sum_{i=1}^{n}\frac{\xi_i(\beta)}{1+\lambda^{\mathrm{T}}\xi_i(\beta)}=\frac{1}{n}\sum_{i=1}^{n}\frac{\xi_i(\beta)}{1+\alpha\delta^{\mathrm{T}}\xi_i(\beta)}=\frac{1}{n}\sum_{i=1}^{n}\delta^{\mathrm{T}}\xi_i(\beta)-\alpha\frac{1}{n}\sum_{i=1}^{n}\frac{\left[\delta^{\mathrm{T}}\xi_i(\beta)\right]^2}{1+\alpha\delta^{\mathrm{T}}\xi_i(\beta)}\\
&\leq\delta^{\mathrm{T}}\bar{\xi}(\beta)-\frac{\alpha}{1+\alpha\xi^*(\beta)}\delta^{\mathrm{T}}S(\beta)\delta,
\end{aligned}
$$

因此 $\alpha\left[\delta^{\mathrm{T}}S(\beta)\delta-\xi^*(\beta)\delta^{\mathrm{T}}\bar{\xi}(\beta)\right]\leq\delta^{\mathrm{T}}\bar{\xi}(\beta)$, 由 Owen[8] 的 (3.36) 和引理 3 得到 $\|\lambda\|=O_p(a_n)$.

**引理 5** 假设条件 C1-C9 成立, 那么在 $D_n=\{\beta:\|\beta-\beta_0\|\leq ca_n\}$ 这个领域内, $L_p(\beta)$ 在 $D_n$ 内有最小值.

**证** 当 $\beta\in D_n$, 有 $\max\limits_{1\leq i\leq n}\left\|\lambda^{\mathrm{T}}\xi_i(\beta)\right\|=o_p(1)$, 将 $Q_{1n}(\beta,\lambda)=\frac{1}{n}\sum\limits_{i=1}^{n}\frac{\xi_i(\beta)}{1+\lambda^{\mathrm{T}}\xi_i(\beta)}=0$ 泰勒展开得到

$$
0=\bar{\xi}(\beta)-S(\beta)\lambda+\gamma_n,
$$

其中 $\gamma_n=\frac{1}{n}\sum\limits_{i=1}^{n}\xi_i(\beta)\frac{\left[\lambda^{\mathrm{T}}\xi_i(\beta)\right]^2}{1+\lambda^{\mathrm{T}}\xi_i(\beta)}$ 且 $|\varsigma_i|\leq\left|\lambda^{\mathrm{T}}\xi_i(\beta)\right|$, 所以 $\lambda=S^{-1}(\beta)\bar{\xi}(\beta)+S^{-1}(\beta)\gamma_n$ 代入 $L(\beta)$ 得到

$$
\begin{aligned}
2L(\beta)&=2\sum_{i=1}^{n}\ln\left(1+\lambda^{\mathrm{T}}\xi_i(\beta)\right)\\
&=n\bar{\xi}^{\mathrm{T}}(\beta)S^{-1}(\beta)\bar{\xi}(\beta)-n\gamma_n{}^{\mathrm{T}}S^{-1}(\beta)\gamma_n+\frac{2}{3}\sum_{i=1}^{n}\left[1+\lambda^{\mathrm{T}}\xi_i(\beta)\right]^3(1+\varsigma_i)^{-4},
\end{aligned}
$$

取 $\varsigma_i$ 并且满足 $|\varsigma_i|\leq\left|\lambda^{\mathrm{T}}\xi_i(\beta)\right|$, $\beta\in\partial D_n$ 这里 $\partial D_n$ 代表 $D_n$ 得边界, $\beta=\beta_0+ca_n\theta_\beta$, $\theta_\beta$ 是单位向量

$$
\begin{aligned}
2L(\beta)=&n\bar{\xi}^{\mathrm{T}}(\beta)S^{-1}(\beta)\bar{\xi}(\beta)+n\left[\bar{\xi}(\beta)-\bar{\xi}(\beta_0)\right]S^{-1}(\beta)\left[\bar{\xi}(\beta)-\bar{\xi}(\beta_0)\right]\\
&+n\bar{\xi}(\beta_0)\left[S^{-1}(\beta)-S^{-1}(\beta_0)\right]\bar{\xi}(\beta_0)+2n\bar{\xi}(\beta_0)S^{-1}(\beta)\left[\bar{\xi}(\beta)-\bar{\xi}(\beta_0)\right]\\
&-n\gamma_n{}^{\mathrm{T}}S^{-1}(\beta)\gamma_n+\frac{2}{3}\sum_{i=1}^{n}\left[1+\lambda^{\mathrm{T}}\xi_i(\beta)\right]^3(1+\varsigma_i)^{-4}\\
=&T_0+T_1+T_2,
\end{aligned}
$$

当 $n\to\infty$ 时, 由引理 3

$$
\begin{aligned}
\bar{\xi}(\beta)-\bar{\xi}(\beta_0)&=\frac{1}{n}\sum_{i=1}^{n}\xi_i(\beta)-\frac{1}{n}\sum_{i=1}^{n}\xi_i(\beta_0)=\frac{1}{n}\sum_{i=1}^{n}\left[\xi_i(\beta)-\xi_i(\beta_0)\right]\\
&=-\frac{1}{n}G(I-S)Z(\beta-\beta_0)+o_p(1),
\end{aligned}
$$

因此 $T_1 = n{\Sigma_1}^2{a_n}^2\Sigma_2$, $T_2/T_1 \xrightarrow{p} 0$, 且 $2L(\beta_0) - T_1 = o_p(1)$, 这就意味着, 当 $n \to \infty$, $P\{2L(\beta) - 2L(\beta_0) > c\} \to 1, c \in R$. 此外, 对于 $n$ 足够大, C(9) 和 SCAD 惩罚得无偏性意味着 $p_\tau(|\beta_i|) = p_\tau(|\beta_{i0}|), i \in \mathrm{A}$, 因此当 $n$ 足够大有

$$\begin{aligned} L_p(\beta) - L_p(\beta_0) &= L(\beta) - L(\beta_0) + n\sum_{i=1}^{p}\left[p_\tau(|\beta_i|) - p_\tau(|\beta_{i0}|)\right] \\ &\geq L(\beta) - L(\beta_0) + n\sum_{i\in \mathrm{A}}^{p}\left[p_\tau(|\beta_i|) - p_\tau(|\beta_{i0}|)\right] \geq L(\beta) - L(\beta_0), \end{aligned}$$

因此 $P\{L_p(\beta) - L_p(\beta_0)\} \to 1$ 对于 $\beta \in \partial D_n$ 成立有个最小值.

**定理 1** 在 C1-C9 的假设条件下, 有 $\sqrt{n}\left(\hat{\beta} - \beta\right) \xrightarrow{D} N\left(0, \Sigma_1\sigma^2(I - \rho W)^{-2}\right)$ 其中 $\xrightarrow{D}$ 表示依分布收敛.

**证** 将 $Q_{1n}(\beta, \lambda) = \frac{1}{n}\sum\limits_{i=1}^{n}\frac{\xi_i(\beta)}{1+\lambda^{\mathrm{T}}\xi_i(\beta)}$, $Q_{2n}(\beta, \lambda) = \frac{\lambda}{1+\lambda^{\mathrm{T}}\xi_i(\beta)}\left(\frac{\partial\xi_i(\beta)}{\partial\beta}\right)^{\mathrm{T}}$ 在 $(\beta, 0)$ 处泰勒展开得到下式:

$$Q_{in}\left(\hat{\beta}, \hat{\lambda}\right) = Q_{in}(\beta, 0) + \frac{\partial Q_{in}(\beta, 0)}{\partial\beta}\left(\hat{\beta} - \beta\right) + \frac{\partial Q_{in}(\beta, 0)}{\partial\lambda^{\mathrm{T}}}\left(\hat{\lambda} - 0\right) + o_p\left(n^{-1/2}\right),$$

写成矩阵形式

$$\begin{pmatrix} \hat{\lambda} \\ \hat{\beta} - \beta \end{pmatrix} = {K_n}^{-1}\begin{pmatrix} Q_{1n}(\beta, 0) + o_p\left(n^{-1/2}\right) \\ o_p\left(n^{-1/2}\right) \end{pmatrix},$$

$$K_n = \begin{pmatrix} \frac{\partial Q_{1n}(\beta,0)}{\partial\lambda^{\mathrm{T}}} & \frac{\partial Q_{1n}(\beta,0)}{\partial\beta} \\ \frac{\partial Q_{2n}(\beta,0)}{\partial\lambda^{\mathrm{T}}} & \frac{\partial Q_{2n}(\beta,0)}{\partial\beta} \end{pmatrix} \to K = \begin{pmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{pmatrix},$$

$$K^{-1} = \begin{pmatrix} {K_{11}}^{-1} + {K_{11}}^{-1}K_{21}{K_{22.1}}^{-1}{K_{11}}^{-1}{K_{12}}^{-1} & -{K_{11}}^{-1}K_{12}{K_{22.1}}^{-1} \\ -{K_{11}}^{-1}K_{12}{K_{22.1}}^{-1} & {K_{22.1}}^{-1} \end{pmatrix},$$

其中 $K_{22.1} = K_{22} - K_{21}{K_{11}}^{-1}K_{12}$, 那么他的逆矩阵 ${K_{22.1}}^{-1} = -\left(K_{21}{K_{11}}^{-1}K_{12}\right)^{-1}$. 由 Owen[8] 得

$$\frac{\partial Q_{1n}(\beta, 0)}{\partial\beta} = \frac{1}{n}\sum_{i=1}^{n}\frac{\partial\xi_i(\beta)}{\partial\beta}, \frac{\partial Q_{1n}(\beta, 0)}{\partial\lambda^{\mathrm{T}}} = -\frac{1}{n}\xi_i(\beta)\xi_i(\beta)^{\mathrm{T}},$$

$$\frac{\partial Q_{2n}(\beta, 0)}{\partial\beta} = 0, \frac{\partial Q_{2n}(\beta, 0)}{\partial\lambda^{\mathrm{T}}} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{\partial\xi_i(\beta)}{\partial\beta}\right)^{\mathrm{T}},$$

所以 $\sqrt{n}\left(\hat{\beta} - \beta\right) = {K_{22.1}}^{-1}K_{21}{K_{11}}^{-1}\sqrt{n}Q_{1n}(\beta, 0) + o_p(1) \xrightarrow{D} N\left(0, \Sigma_1\sigma^2(I - \rho W)^{-2}\right)$, 定理 1 证明完成.

**定理 2** 假设条件 C1-C9 成立, 那么

$$-2L(\beta) = 2\sum_{i=1}^{n}\log\left\{1 + \lambda^{\mathrm{T}}\xi_i(\beta)\right\} \xrightarrow{D} \chi^2_{p+1},$$

其中 $\chi^2_{p+1}$ 表示具有 $p+1$ 个自由度的卡方分布.

**证** 假设 $\frac{1}{n}\sum\limits_{i=1}^{n}\xi_i{}^{\mathrm{T}}(\beta)\xi_i(\beta)$ 是正定的, 且 $\frac{1}{n}\sum\limits_{i=1}^{n}\frac{\xi_i(\beta)}{1+\lambda^{\mathrm{T}}\xi_i(\beta)}=0$, 使用 Owen[8] 的 (2.14) 同样的方法结合引理 4, 得

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}\frac{\xi_i(\beta)}{1+\lambda^{\mathrm{T}}\xi_i(\beta)}&=\frac{1}{n}\sum_{i=1}^{n}\xi_i(\beta)\left[1-\lambda^{\mathrm{T}}\xi_i(\beta)+\frac{\left[\lambda^{\mathrm{T}}\xi_i(\beta)\right]^2}{1+\lambda^{\mathrm{T}}\xi_i(\beta)}\right]\\&=\frac{1}{n}\sum_{i=1}^{n}\xi_i(\beta)-\frac{1}{n}\sum_{i=1}^{n}\xi_i{}^{\mathrm{T}}(\beta)\xi_i(\beta)\lambda^{\mathrm{T}}+\frac{1}{n}\sum_{i=1}^{n}\frac{\xi_i(\beta)\left[\lambda^{\mathrm{T}}\xi_i(\beta)\right]^2}{1+\lambda^{\mathrm{T}}\xi_i(\beta)}=0,\end{aligned}$$

因此 $\lambda=S^{-1}(\beta)\bar{\xi}(\beta)+o_p(a_n)$ 通过泰勒展开得:

$$2L(\beta)=2n\lambda^{\mathrm{T}}\bar{\xi}(\beta)-n\lambda^{\mathrm{T}}S(\beta)\lambda+o_p(1)=n\bar{\xi}(\beta)S^{-1}(\beta)\bar{\xi}(\beta)+o_p(1),$$

结合引理 4, 我们可以证明当 $n\to\infty$ 时,

$$\left\{\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\xi_i(\beta)\right\}^{\mathrm{T}}\left\{\frac{1}{n}\sum_{i=1}^{n}\xi_i{}^{\mathrm{T}}(\beta)\xi_i(\beta)\right\}^{-1}\left\{\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\xi_i(\beta)\right\}\xrightarrow{D}\chi^2_{p+1},$$

定理 2 证明完成.

**定理 3** 假设条件 C1-C9 成立, 且 $p^5/n\to0$, 当 $n\to\infty$ 有
1. 稀疏性: $\hat{\beta}_2=0$.
2. 渐近正态性: $\sqrt{n}B_n\Sigma_p{}^{-\frac{1}{2}}\left(\hat{\beta}_1-\hat{\beta}_{10}\right)\xrightarrow{D}N(0,g)$, 其中 $B_n$ 是 $s\times d$ 矩阵且 $B_nB_n{}^{\mathrm{T}}\to g$, $\Sigma_n=\Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$. 根据引理 5, 当 $\beta\in D_n$ 时, $j=1,2,\cdots,p$ 有

$$\frac{1}{n}\frac{\partial L_p(\beta)}{\partial\beta_j}=\frac{1}{n}\sum_{i=1}^{n}\lambda^{\mathrm{T}}\frac{\partial\xi_i(\beta)}{\partial\beta_j}\{1+op(1)\}+p'_\tau(|\beta_j|)\,sign(\beta_j)=\mathrm{I}+\mathrm{II},$$

结合假设条件条件 C1-C9 有

$$\max_{j\in\mathrm{A}}|\mathrm{I}|\leq\max_{j\in\mathrm{A}}\left|\lambda^{\mathrm{T}}\right|\left|\frac{1}{n}\sum_{i=1}^{n}\frac{\partial\xi_i(\beta)}{\partial\beta_j}\right|=O_p(a_n),$$

当 $\tau(n/p)^{\frac{1}{2}}\to\infty$ 时, 对于任意, $j\in\mathrm{A}$,$p'_\tau(|\beta_j|)\,sign(\beta_j)$ 渐进的主导 $\frac{\partial\xi_i(\beta)}{\partial\beta_j}$ 的符号, 因此对于任意 $j\notin\mathrm{A}$, 当 $n\to\infty$ 时, 依概率趋近于 1, 则

$$\frac{\partial L_p(\beta)}{\partial\beta_j}>0,\beta_j\in(0,ca_n),\frac{\partial L_p(\beta)}{\partial\beta_j}<0,\beta_j\in(-ca_n,0),$$

故对于任意 $j\in\mathrm{A}$, 依概率趋近于 1 有 $\hat{\beta}_j=0$, 则稀疏性: $\hat{\beta}_2=0$ 得证.

取 $\psi_1$ 和 $\psi_2$ 满足 $\psi_1\beta=\beta_1$ 和 $\psi_2\beta=\beta_2$, 由拉格朗日乘法得到 $L_p(\beta)$ 的极小值,

$$\frac{1}{n}\sum_{i=1}^{n}\log\left\{1+\lambda^{\mathrm{T}}\xi_i(\beta)\right\}\sum_{j=1}^{p}p'_\tau(|\beta_j|)+\nu^{\mathrm{T}}\psi_2\beta,$$

其中, $\nu$ 表示 $(p-s)$ 维拉格朗日算子, 定义

$$\tilde{Q}_{1n}(\beta,\lambda,\nu)=\frac{1}{n}\sum_{i=1}^{n}\frac{\xi_i(\beta)}{1+\lambda^{\mathrm{T}}\xi_i(\beta)},$$
$$\tilde{Q}_{2n}(\beta,\lambda,\nu)=\frac{1}{n}\sum_{i=1}^{n}\frac{1}{1+\lambda^{\mathrm{T}}\xi_i(\beta)}\left(\frac{\partial\xi_i(\beta)}{\partial\beta^{\mathrm{T}}}\right)^{\mathrm{T}}\lambda+b(\beta)+{\psi_2}^{\mathrm{T}}\nu,$$
$$\tilde{Q}_{3n}(\beta,\lambda,\nu)=\psi_2\beta,$$

其中 $b(\beta)=\left\{{p'}_\tau(|\beta_1|)\,sign(\beta_1),{p'}_\tau(|\beta_2|)\,sign(\beta_2),\cdots,{p'}_\tau(|\beta_p|)\,sign(\beta_p),0\cdots,0\right\}^{\mathrm{T}}$, 关于 $\left(\hat{\beta},\hat{\lambda},\hat{\nu}\right)$ 的最小值满足 $0=\tilde{Q}_{jn}\left(\hat{\beta},\hat{\lambda},\nu\right)$, $(j=1,2,3)$. 因为 $\|\lambda\|=O_p(a_n)$ 对于 $\beta\in D_n$, 可以从 $0=\tilde{Q}_{2n}\left(\hat{\beta},\hat{\lambda},\hat{\nu}\right)$ 知 $\|\nu\|=O_p(a_n)$, 将 $\tilde{Q}_{jn}\left(\hat{\beta},\hat{\lambda},\nu\right)$ 在 $(\beta_0,0,0)$ 处泰勒展开, 给出以下偏导

$$\frac{\partial\tilde{Q}_{1n}(\beta,0,0)}{\partial\lambda}=-S(\beta_0),\frac{\partial\tilde{Q}_{1n}(\beta,0,0)}{\partial\beta}=\frac{1}{n}\sum_{i=1}^{n}\frac{\partial\xi_i(\beta_0)}{\partial\beta},\frac{\partial\tilde{Q}_{1n}(\beta,0,0)}{\partial\hat{\nu}}=0,$$
$$\frac{\partial\tilde{Q}_{2n}(\beta,0,0)}{\partial\lambda}=\frac{1}{n}\sum_{i=1}^{n}\frac{\partial\xi_i(\beta_0)}{\partial\beta},\frac{\partial\tilde{Q}_{2n}(\beta,0,0)}{\partial\beta}=b'(\beta),\frac{\partial\tilde{Q}_{2n}(\beta,0,0)}{\partial\hat{\nu}}=\psi_2,$$
$$\frac{\partial\tilde{Q}_{3n}(\beta,0,0)}{\partial\lambda}=0,\frac{\partial\tilde{Q}_{3n}(\beta,0,0)}{\partial\beta}=\psi_2,\frac{\partial\tilde{Q}_{3n}(\beta,0,0)}{\partial\hat{\nu}}=0,$$

通过泰勒展开写成矩阵形式

$$\begin{pmatrix}\tilde{Q}_{1n}(\beta_0,0,0)\\0\\0\end{pmatrix}=\begin{pmatrix}-\Sigma_2&\Sigma_1&0\\\Sigma_1&0&{\psi_2}^{\mathrm{T}}\\0&\psi_2&0\end{pmatrix}\begin{pmatrix}\hat{\lambda}\\\hat{\beta}-\beta_0\\\hat{\nu}\end{pmatrix}+R_n,$$

其中 $R_n=\sum\limits_{l=1}^{5}R_n^{(l)},R_n^{(l)}=\left(R_{1n}^{(1)\mathrm{T}},R_{2n}^{(1)\mathrm{T}},0\right)^{\mathrm{T}},R_{jn}^{(1)\mathrm{T}}\in R^p\,(j=1,2)\,,k=1,2,\cdots,p$ 且第 $k$ 个分量为 $R_{jn,k}^{(1)\mathrm{T}}=\frac{1}{2}\left(\hat{\phi}-\phi_0\right)^{\mathrm{T}}\frac{\partial^2\tilde{Q}_{jn,k}(\tilde{\phi})}{\partial\phi\partial\phi^{\mathrm{T}}}\left(\hat{\phi}-\phi_0\right)$, 其中 $\phi=\left(\beta^{\mathrm{T}},\lambda^{\mathrm{T}}\right)^{\mathrm{T}},\tilde{\phi}=\left(\tilde{\beta}^{\mathrm{T}},\tilde{\lambda}^{\mathrm{T}}\right)^{\mathrm{T}}$, 满足 $\left\|\tilde{\phi}-\phi_0\right\|\leq\left\|\hat{\phi}-\phi_0\right\|$.

$${R_n}^{(2)}=(0,b(\beta_0),0)^{\mathrm{T}},$$
$${R_n}^{(3)}=\left[0,\left[b'\left(\tilde{\beta}\right)\left(\hat{\beta}-\beta_0\right)\right],0\right]^{\mathrm{T}},$$
$${R_n}^{(4)}=\left\{\left[\left(\Sigma_2-S(\beta_0)\right)\hat{\lambda}\right]^{\mathrm{T}}+\left[\left(\frac{1}{n}\sum_{i=1}^{n}\frac{\partial\xi_i(\beta)}{\partial\beta}-\Sigma_1\right)\left(\hat{\beta}-\beta_0\right)\right]^{\mathrm{T}},0,0\right\}^{\mathrm{T}},$$
$${R_n}^{(5)}=\left\{0,\left[\left(\frac{1}{n}\sum_{i=1}^{n}\frac{\partial\xi_i(\beta)}{\partial\beta}-\Sigma_1\right)\hat{\lambda}\right]^{\mathrm{T}},0\right\}^{\mathrm{T}},$$

结合条件得, 对 $l=1,\cdots,5.{R_n}^{(l)}=o_p\left(n^{-1/2}\right)$, 记 $\Pi_{11}=-\Sigma_2,\Pi_{12}=(\Sigma_1,0),\Pi_{21}={\Pi_{12}}^{\mathrm{T}}$,

$$\Pi_{22}=\begin{pmatrix}0&{\psi_2}^{\mathrm{T}}\\\psi_2&0\end{pmatrix},\Pi=\begin{pmatrix}\Pi_{11}&\Pi_{12}\\\Pi_{21}&\Pi_{22}\end{pmatrix},$$

设 $\kappa=\left(\beta^{\mathrm{T}},\nu^{\mathrm{T}}\right)$

$$\begin{pmatrix}\hat{\lambda}\\\hat{\kappa}-\kappa_0\end{pmatrix}=\Pi^{-1}\left\{\begin{pmatrix}-\tilde{Q}_{1n}(\beta,0,0)\\0\end{pmatrix}+R_n\right\},$$

且 $\Pi^{-1}=\begin{pmatrix} {\Pi_{11}}^{-1}+{\Pi_{11}}^{-1}\Pi_{21}{\Pi_{22.1}}^{-1}{\Pi_{11}}^{-1}{\Pi_{12}}^{-1} & -{\Pi_{11}}^{-1}\Pi_{12}{\Pi_{22.1}}^{-1} \\ -{\Pi_{11}}^{-1}\Pi_{12}{\Pi_{22.1}}^{-1} & {\Pi_{22.1}}^{-1} \end{pmatrix}$, 其中 $\Pi_{22.1}=\Pi_{22}-\Pi_{21}{\Pi_{11}}^{-1}\Pi_{12}=\begin{pmatrix} \Sigma^{-1} & {\psi_2}^{\mathrm{T}} \\ \psi_2 & 0 \end{pmatrix}$, 因此 $\hat{\kappa}-\kappa_0={\Pi_{22.1}}^{-1}\Pi_{21}{\Pi_{11}}^{-1}\tilde{Q}_{1n}\left(\beta_0,0,0\right)+o_p\left(n^{-1/2}\right)$.

$${\Pi_{22.1}}^{-1}=\begin{pmatrix} \Sigma-\Sigma{\psi_2}^{\mathrm{T}}\left(\psi_2\Sigma{\psi_2}^{\mathrm{T}}\right)\psi_2\Sigma & -\Sigma{\psi_2}^{\mathrm{T}}\left(\psi_2\Sigma{\psi_2}^{\mathrm{T}}\right)^{-1} \\ -\left(\psi_2\Sigma{\psi_2}^{\mathrm{T}}\right)^{-1}\psi_2\Sigma & \left(\psi_2\Sigma{\psi_2}^{\mathrm{T}}\right)^{-1} \end{pmatrix},$$

因此 $\hat{\beta}-\beta_0=\left\{\Sigma-\Sigma{\psi_2}^{\mathrm{T}}\left(\psi_2\Sigma{\psi_2}^{\mathrm{T}}\right)^{-1}\psi_2\Sigma\right\}\left\{{\Sigma_1}^{\mathrm{T}}{\Sigma_2}^{-1}\bar{\xi}\left(\beta_0\right)+o_p\left(n^{-1/2}\right)\right\}$, 对 $\hat{\beta}_1$ 展开得

$$\hat{\beta}_1-\beta_{10}=\left\{\psi_1\Sigma-\psi_1\Sigma{\psi_2}^{\mathrm{T}}\left(\psi_2\Sigma{\psi_2}^{\mathrm{T}}\right)^{-1}\psi_2\Sigma\right\}\left\{{\Sigma_1}^{\mathrm{T}}{\Sigma_2}^{-1}\bar{\xi}\left(\beta_0\right)+o_p\left(n^{-1/2}\right)\right\},$$

其中 $\psi_1\Sigma-\psi_1\Sigma{\psi_2}^{\mathrm{T}}\left(\psi_2\Sigma{\psi_2}^{\mathrm{T}}\right)^{-1}\psi_2\Sigma=\left(\Sigma_n,0\right)$ 则

$$\sqrt{n}B_n{\Sigma_n}^{-1/2}\left(\hat{\beta}_1-\beta_{10}\right)\xrightarrow{D}N\left(0,g\right)$$

定理 3 证明完成.

## 参 考 文 献


[1] Wei Chanhua, Guo Shuang, Zhai Shufen. Statistical inference of partially linear varying coefficient spatial autoregressive models[J]. Economic Modelling, 2017, 64: 553–559.

[2] Luo Guowang, Wu Mixia, Pang Zhen. Empirical likelihood inference for the semiparametric varying-coefficient spatial autoregressive model[J]. Journal of Systems Science and Com-plexity, 2022, 34(6): 2310–2333.

[3] Liang Xuan, Gao Jiti, Gong Xiaodong. Semiparametric spatial autoregressive panel data model with fixed effects and time-varying coefficients[J]. Journal of Business Economic Statistics, 2022, 40(4): 1784–1802.

[4] Su Liangjun. Semiparametric GMM estimation of spatial autoregressive models[J]. Journal of Econometrics, 2012, 167(2): 543–560.

[5] Zhang Yuanqing, Li Hong, Feng Yaqin. Inference for partially linear additive higher-order spatial autoregressive model with spatial autoregressive error and unknown heteroskedas-ticity[J]. Communications in Statistics - Simulation and Computation, 2023, 52(3): 898–924.

[6] Su Liangjun, Jin Saina. Profile quasi-maximum likelihood estimation of partially linear spatial autoregressive models[J]. Journal of Econometrics, 2009, 157(1): 18–33.

[7] 陈建宝, 乔宁宁. 半参数变系数空间误差回归模型的估计 [J]. 数量经济技术经济研究, 2017, 34(04): 129–146.

[8] Owen A. Empirical likelihood ratio confidence regions[J]. The Annals of Statistics. 1990, 18(1): 90–120.

[9] Tang Chengyong, Leng Chenlei. Penalized high-dimensional empirical likelihood[J]. Bio-metrika, 2010, 97(4): 905–920.

[10] Chen Xia, Tan Xiaoyan, Yan Li. Penalized empirical likelihood for high-dimensional generalized linear models with longitudinal data[J]. Journal of Statistical Computation and Simulation, 2023, 93(10): 1515–1531.

[11] He Bangqiang, Hong Xingjian, Fan Guoliang. Penalized empirical likelihood for partially linear errors-in-variables panel data models with fixed effects[J]. Statistical Papers, 2020, 61: 2351–2381.

[12] Wang Xinyang, Wang Dehui. Penalized empirical likelihood inference for the GINAR(p) model[J]. Statistics, 2022, 56(4): 785–804.

[13] Fan Jianqing, Li Runze. Variable selection via nonconcave penalized likelihood and its oracle properties[J]. Journal of the American Statistical Association, 2001, 96(456): 1348–1360.

[14] Fan Guoliang, Xu Hongxia, Liang Huaying. Empirical likelihood inference for partially time-varying coefficient errors-in-variables models[J]. Electronic Journal of Statistics, 2012, 6: 1040–1058.

[15] Fan Guoliang, Liang Huaying. Empirical likelihood for longitudinal partially linear model with α -mixing errors[J]. Journal of Systems Science and Complexity, 2013, 26(2): 232–248.

[16] You Jinhong, Chen Gemai. Estimation of a semiparametric varying-coefficient partially linear errors-in-variables model[J]. Journal of Multivariate Analysis, 2005, 97(2): 324–341.

[17] Li Yinghua, Qin Yongsong, Li Yuan. Empirical likelihood for nonparametric regression models with spatial autoregressive errors[J]. Journal of the Korean Statistical Society, 2021, 50: 1–32.

[18] Li Gaorong, Lin Lu, Zhu Lixing. Empirical likelihood for a varying coefficient partially linear model with diverging number of parameters[J]. Journal of Multivariate Analysis, 2012, 105(1): 85–111.

[19] 周婷, 秦永松. 带固定效应空间误差面板数据模型的经验似然推断 [J]. 数学杂志, 2020, 40(5): 551–564.

[20] 程素丽. 部分线性可加空间误差回归模型的 GMM 估计 [D]. 福州: 福建师范大学, 2019.

# PENALIZED EMPIRICAL LIKELIHOOD FOR SEMIPARAMETRIC VARYING COEFFICIENT SPATIAL ERROR REGRESSION MODEL

WANG Yi-heng, HE Bang-qiang

(*School of Mathematics-Physics and Finance, Anhui Polytechnic University, Wuhu 241000, China*)

**Abstract:** The article considers the problem of parameter estimation and variable selection in semi-parametric variable coefficient spatial error regression model. The local linear estimation method was used to estimate the variable coefficient function, then the maximum empirical log-likelihood ratio estimation of parametric and non-parametric components is constructed, and PEL was suggested to select the variables, and the square and sum method was used to estimate the variance of the spatial coefficient and error term. The estimated values of pa-rameters and non-parameters and the superiority of PEL method in selecting variables are obtained. Under suitable conditions, the Penalized empirical likelihood has oracle charac-teristics and obeys the asymptotic chi-square distribution under the null hypothesis.

**Keywords:** partially linear model; penalized empirical likelihood; SAR; Errors-in-variables

**2010 MR Subject Classification:** 62G05; 62E20