第 32 卷第 1 期
2015 年 1 月
统计研究
Statistical Research
Vol. 32, No. 1
Jan. 2015

随机效应空间滞后单指数面板模型*

陈建宝　孙 林

内容提要: 本文构建了随机效应空间滞后单指数面板模型的截面极大似然估计方法,从理论证明和数值模拟两方面分别考察了其估计量的大样本性质和小样本表现。研究结果表明:在大样本条件下,估计量均具有一致性,并且参数估计量具有渐近正态性;在小样本条件下,各估计量依然具有良好的表现,其精度随着样本容量的增加而提高。空间权重矩阵结构的复杂性对空间相关系数的估计量影响较大,但对其他估计量的影响较小。

关键词: 随机效应;空间滞后单指数面板模型;截面极大似然估计;数值模拟

中图分类号: F222.3　　**文献标识码:** A　　**文章编号:** 1002-4565(2015)01-0095-07

Spatial Lag Single-index Panel Model with Random Effects

Chen Jianbao & Sun Lin

Abstract: In this paper, a profile maximum likelihood estimation method is constructed for a spatial lag single-index panel model with random effects and the large sample properties and finite sample performances of its estimators are studied by theoretical proofs and numerical simulation respectively. The research results are summarized as follows: (1) All estimators have consistency and parametric estimators have asymptotic normality under large sample. (2) All estimators still have good performance under finite sample and their accuracy is improved with the increasing of sample size; the complexity of spatial weighted matrix has remarkable effect on the estimator of spatial correlation coefficient and little effects on the other estimators.

Key words: Random Effects; Spatial Lag Single-index Panel Model; Profile Maximum Likelihood Estimation; Numerical Simulation

一、引言

在实际问题研究中,变量之间的相互关系通常较为复杂,通过预先设定形式的参数模型往往很难准确描述这种关系。近年来,伴随着计算机技术的快速发展,从数据出发,让数据“说话”的非参数模型引起了众多研究者的关注并取得了丰富的研究成果(Fan 和 Gijbels,1996)。当非参数模型中的解释变量较多时,往往会出现所谓的“维数灾难”问题,导致估计的可靠性降低。为了有效地克服“维数灾难”问题,人们提出了一些新模型,其中一种行之有效的方法就是建立单指数模型。

单指数模型的基本数学形式为 $Y = g(x'\beta) + \varepsilon$ (Friedman 和 Stuetzle,1981),其中,β 为未知参数,$g(\cdot)$ 为未知连接函数。其优点在于通过连接函数 $g(\cdot)$ 实现了降维效果,有效地避免了“维数灾难”问题,同时还可以更好地反映变量间的相互关系。在此基础上,人们发展和研究了各种拓展的单指数模型,主要包括截面数据单指数模型和面板数据单指数模型,研究的重点在于如何寻求 β 和 $g(\cdot)$ 的有效估计。

对于截面数据单指数模型,人们提出了各种估计方法。借鉴投影寻踪回归的思想,Ichimura(1993)和 Carroll 等(1997)基于非参连接函数的最小二乘准则,分别采用不同函数估计形式替代未知函数进行估计;Ai 和 Chen(2003)则利用半参数极大似然估计研究了单指数模型的估计问题;Wang 等(2010)和 Cui 等(2011)基于参数约束假设分别给出关于单指数模型未知参数和未知函数估计的非迭代方法,提高了运算速度;其他估计方法还有半参数秩估计法、平均导数法、半参数 M 估计方法等。

* 本文为教育部人文社会科学项目“空间自回归单指数模型的理论和实践”(13YJA9100002)的部分研究成果。

与截面数据单指数模型相比，面板数据单指数模型更为复杂，目前研究成果相对较少。基于截面最小二乘估计的思想，Pang 和 Xue(2012) 给出了随机效应单指数面板模型的估计方法，并证明了估计量的渐近正态性；Chen 等(2013) 则对固定效应部分线性单指数面板模型构建了半参数最小平均方差估计方法，并证明了估计的渐近性质。

随着研究的深入，空间计量模型理论发展迅速，应用日趋广泛。其中，参数空间计量模型的理论研究已趋于成熟，对于截面数据空间计量模型研究和面板数据空间计量模型研究均得到了丰富的、较完整的理论成果(Kelejian 和 Prucha，2010；Lee 和 Yu，2010)。应用研究成果表明，对一些具有空间相关性的实际问题，由于参数空间回归模型预先设定了模型形式，并不能很好地解释空间变量间的非线性关系(Gress，2004)。近年来，对模型形式限制较少的非参数空间计量模型开始引起学者们的关注(Robinson，2010；Su，2012)，其理论研究才刚刚起步，成果较少。为了克服非参数空间滞后模型中存在的“维数灾难”问题，本文提出了一种现有研究文献还未涉及的具有随机效应的空间滞后单指数面板模型。针对该模型，构建了截面极大似然估计方法，从理论上证明了其估计量的优良性，通过数据模拟考察了其估计量的小样本表现，并考察了空间权重矩阵结构复杂性对估计的影响。

二、理论模型

参考 Su(2012) 的模型设定，随机效应空间滞后单指数面板模型的数学表达式为：

$$y_{ij} = \rho(W_0Y_j)_i + g(x'_{ij}\beta) + b_i + \varepsilon_{ij} \quad 1 \leqslant i \leqslant N, 1 \leqslant j \leqslant T \tag{1}$$

其中，y_{ij} 为被解释变量在第 i 个截面第 j 个时刻的观测值，$Y_j = (y_{1j},\cdots,y_{Nj})'$，$\rho$ 为待估空间相关系数，$W_0 = (w_{ij})_{N\times N}$ 为预先设定的空间权重矩阵，$(W_0Y_j)_i$ 代表 W_0Y_j 的第 i 个分量，$g(\cdot)$ 是未知连接函数，$x_{ij} = (x_{ij1},\cdots,x_{ijp})'$ 为 p 个解释变量在第 i 个截面第 j 个时刻的观测值向量，$\beta \in R^p$ 为待估参数向量，b_i 为第 i 个截面的个体效应，ε_{ij} 为第 i 个截面第 j 个时刻观测值的随机误差项，$b_i \sim i.i.d.N(0,\sigma_b^2)$，$\varepsilon_{ij} \sim i.i.d.N(0,\sigma_\varepsilon^2)$，且 b_i 和 ε_{ij} 相互独立。

记 $Y \hat{=} (y_{11},\cdots,y_{1T},\cdots,y_{N1},\cdots,y_{NT})'$，则模型(1) 的矩阵形式可等价地表示为：

$$Y = \rho WY + G + Ub + \varepsilon \tag{2}$$

其中，$W = W_0 \otimes I_T$，$G = (g(x'_{11}\beta),\cdots,g(x'_{1T}\beta),\cdots,g(x'_{N1}\beta),\cdots,g(x'_{NT}\beta))'$，$U = I_N \otimes 1_T$，$b = (b_1,\cdots,b_N)'$，$\varepsilon = (\varepsilon_{11},\cdots,\varepsilon_{1T},\cdots,\varepsilon_{N1},\cdots,\varepsilon_{NT})'$，$I_N$ 为 N 阶单位矩阵，“$\otimes$”为 Kronecker 乘积。

记 $\theta = (\lambda',\sigma')'$，其中，$\lambda = (\beta',\rho)'$，$\sigma = (\sigma_\varepsilon^2,\sigma_b^2)'$。对于模型(2) 而言，统计推断最重要的第一步就是找到合适的估计方法给出未知参数向量 θ 和函数 $g(\cdot)$ 的估计。通常的估计方法是极大似然估计方法，下面考察该方法是否适用于我们提出的模型(2)。

令 $A(\rho) = I - \rho W$①，则模型(2) 可写为：

$$A(\rho)Y = G + Ub + \varepsilon \tag{3}$$

即：$Ub + \varepsilon = A(\rho)Y - G$。易知，$\partial(Ub+\varepsilon)/\partial Y = A(\rho)$。因 $b_i \sim i.i.d.N(0,\sigma_b^2)$，$\varepsilon_{ij} \sim i.i.d.N(0,\sigma_\varepsilon^2)$，所以，$\Sigma(\sigma) \hat{=} E(Ub+\varepsilon)(Ub+\varepsilon)' = \sigma_\varepsilon^2 I + \sigma_b^2 I_N \otimes (1_T1'_T)$，于是

$$|\Sigma(\sigma)| = \sigma_\varepsilon^{2(NT-N)}(\sigma_b^2 + T\sigma_\varepsilon^2)^N,\quad [\Sigma(\sigma)]^{-1} = \frac{1}{\sigma_\varepsilon^2}I + \left(\frac{1}{\sigma_\varepsilon^2 + T\sigma_b^2} - \frac{1}{\sigma_\varepsilon^2}\right)I_N \otimes \left(\frac{1}{T}1_T1'_T\right)$$

因此，模型(2) 的似然函数可以写为：

$$L(\theta) = (2\pi)^{-NT/2}|\Sigma(\sigma)|^{-1/2}\exp\left\{-\frac{1}{2}[A(\rho)Y - G]'[\Sigma(\sigma)]^{-1}[A(\rho)Y - G]\right\}|A(\rho)| \tag{4}$$

将式(4) 代入相应的参数表达式可得其对数似然函数为：

$$\begin{aligned}\ln L(\theta) = &-\frac{N(T-1)}{2}\ln\sigma_\varepsilon^2 - \frac{N}{2}\ln(\sigma_\varepsilon^2 + T\sigma_b^2) \\ &- \frac{1}{2(\sigma_\varepsilon^2 + T\sigma_b^2)}[A(\rho)Y - G]'H[A(\rho)Y - G] \\ &- \frac{1}{2\sigma_\varepsilon^2}[A(\rho)Y - G]'(I - H)[A(\rho)Y - G] \\ &+ \ln|A(\rho)| + const\end{aligned} \tag{5}$$

其中，$H = I_N \otimes \left(\frac{1}{T}1_T1'_T\right)$。为便于计算，去掉常数项，对上式关于 σ_b^2 和 σ_ε^2 分别求偏导并令其为 0，可得 σ_ε^2 和 σ_b^2 的极大似然估计分别为：

$$\hat{\sigma}_{\varepsilon T}^2 = [N(T-1)]^{-1}[A(\rho)Y - G]'(I - H)[A(\rho)Y - G] \tag{6}$$

$$\hat{\sigma}_{bT}^2 = (NT)^{-1}[A(\rho)Y - G]'H[A(\rho)Y - G] - T^{-1}\hat{\sigma}_{\varepsilon T}^2 \tag{7}$$

---

① 在本文中未指明单位矩阵具体形式时，$I = I_{NT}$。

将式(6)和式(7)代入式(5)得到关于 λ 的集中对数似然函数:

$$\ln L(\lambda) = \frac{-N(T-1)}{2}\ln\hat{\sigma}_{\varepsilon T}^2 - \frac{N}{2}\ln(\hat{\sigma}_{\varepsilon T}^2 + T\hat{\sigma}_{bT}^2) - \frac{NT}{2} + \ln|A(\rho)| \quad (8)$$

对于式(8),因为 $g(x'_{ij}\beta)$ 未知,$i=1,\cdots,N,j=1,\cdots,T$,无法直接通过关于 λ 求最大化得到 λ 的估计值,因此,传统的极大似然估计方法在此并不适用。

三、截面极大似然估计方法

由于一般极大似然估计方法不适用于模型(2),本文尝试使用截面极大似然估计方法寻求模型(2)中未知参数和函数的估计,具体估计方法的实施步骤如下:

步骤 1　假定 θ 已知,用局部线性法得到 $g(u)$ 的可行初始估计 $\hat{g}_{IN}(u)$,即 $\hat{g}_{IN}(u)=\hat{a}$,而关于$(\hat{a},\hat{b})$的定义为:

$$(\hat{a},\hat{b}) = \arg\min_{a,b}\frac{1}{NT}\sum_{i=1}^{N}\sum_{j=1}^{T}[\tilde{y}_{ij} - a - b(u_{ij} - u)]^2 k_h(u_{ij}-u) \quad (9)$$

其中,$\dot{\hat{g}}_{IN}(u)=\hat{b}$,$\tilde{y}_{ij}=y_{ij}-\rho(W_0Y_j)_i$,$k_h(u_{ij}-u)=h^{-1}k((u_{ij}-u)/h)$,$u_{ij}=x'_{ij}\beta$,$i=1,\cdots,N$,$j=1,\cdots,T$,$k(\cdot)$ 为一元核函数,h 为窗宽。

令 $Z(u,\beta)=\begin{pmatrix}1 & \cdots & 1\\ \frac{u_{11}-u}{h} & \cdots & \frac{u_{NT}-u}{h}\end{pmatrix}'$,$K(u,\beta)=diag(k_h(u_{11}-u),\cdots,k_h(u_{NT}-u))$,$\delta=(a,hb)'$,$\hat{\delta}=(\hat{a},h\hat{b})'$,则式(9)可写成如下形式:

$$\hat{\delta} = \arg\min_{a,b}\frac{1}{NT}[A(\rho)Y - Z(u,\beta)\delta]'K(u,\beta)[A(\rho)Y - Z(u,\beta)\delta]$$

即有:

$$\hat{\delta} = [Z'(u,\beta)K(u,\beta)Z(u,\beta)]^{-1}Z'(u,\beta)K(u,\beta)A(\rho)Y \quad (10)$$

令 $[Z'(u,\beta)K(u,\beta)Z(u,\beta)]^{-1}Z'(u,\beta)K(u,\beta)=S(u,\beta)$,则 $\delta=S(u,\beta)A(\rho)Y$。易知,

$$\hat{g}_{IN}(u) = \hat{a} = e'_1\hat{\delta} = e'_1S(u,\beta)A(\rho)Y = s(u,\beta)A(\rho)Y \quad (11)$$

其中,$e'_1S(u,\beta)=s(u,\beta)$,$e_1=(1,0)'$。记 $S(\beta)=(s(u_1,\beta)',\cdots,s(u_{NT},\beta)')'$,则 G 的初始估计为:

$$\hat{G}_{IN} = (\hat{g}_{IN}(u_1),\cdots,\hat{g}_{IN}(u_{NT}))' = (s(u_1,\beta)A(\rho)Y,\cdots,s(u_{NT},\beta)A(\rho)Y)' = (s(u_1,\beta)',\cdots,s(u_{NT},\beta)')'A(\rho)Y = S(\beta)A(\rho)Y \quad (12)$$

步骤 2　用 $\hat{G}_{IN}$ 替代式(5)中的 G,得到对数似然函数的近似值为:

$$\ln\tilde{L}(\theta) = -\frac{N(T-1)}{2}\ln\sigma_\varepsilon^2 - \frac{N}{2}\ln(\sigma_\varepsilon^2 + T\sigma_b^2) - \frac{1}{2(\sigma_\varepsilon^2+T\sigma_b^2)}[A(\rho)Y-\hat{G}_{IN}]'H[A(\rho)Y-\hat{G}_{IN}] - \frac{1}{2\sigma_\varepsilon^2}[A(\rho)Y-\hat{G}_{IN}]'(I-H)[A(\rho)Y-\hat{G}_{IN}] + \ln|A(\rho)| \quad (13)$$

上式为未知参数 θ 的函数,其极大似然估计为 $\hat{\theta}=\arg\max\limits_{\theta}\frac{1}{NT}\ln\tilde{L}(\theta)$。

在实际应用中,往往分两步来完成:

首先,假定 λ 已知,对式(13)求解关于$(\sigma_\varepsilon^2,\sigma_b^2)'$的最大化问题,得到 σ_ε^2 和 σ_b^2 的初始估计分别为:

$$\hat{\sigma}_{\varepsilon IN}^2 = [N(T-1)]^{-1}[A(\rho)Y-\hat{G}_{IN}]'(I-H)[A(\rho)Y-\hat{G}_{IN}] \quad (14)$$

$$\hat{\sigma}_{bIN}^2 = (NT)^{-1}[A(\rho)Y-\hat{G}_{IN}]'H[A(\rho)Y-\hat{G}_{IN}] - T^{-1}\hat{\sigma}_{\varepsilon IN}^2 \quad (15)$$

其次,将 $\hat{\sigma}_{\varepsilon IN}^2$、$\hat{\sigma}_{bIN}^2$ 分别替代式(13)中的 σ_ε^2 和 σ_b^2,得到关于 λ 的集中对数似然函数:

$$\ln\tilde{L}(\lambda) = -\frac{N(T-1)}{2}\ln\hat{\sigma}_{\varepsilon IN}^2 - \frac{N}{2}\ln(\hat{\sigma}_{\varepsilon IN}^2 + T\hat{\sigma}_{bIN}^2) - \frac{NT}{2} + \ln|A(\rho)| \quad (16)$$

从而 λ 的估计为:

$$\hat{\lambda} = \arg\max_{\lambda}\frac{1}{NT}\ln\tilde{L}(\lambda) \quad (17)$$

上式为非线性最优化求解问题,可用迭代方法来实现。

得到 $\hat{\lambda}=(\hat{\beta},\hat{\rho})'$后,用 $\hat{\lambda}$ 替代式(14)和式(15)中的 λ,可得 σ_ε^2 和 σ_b^2 的最终估计分别为:

$$\hat{\sigma}_\varepsilon^2 = \frac{1}{N(T-1)}[A(\hat{\rho})Y - S(\hat{\beta})A(\hat{\rho})Y]'(I-H)[A(\hat{\rho})Y - S(\hat{\beta})A(\hat{\rho})Y] \quad (18)$$

$$\hat{\sigma}_b^2 = \frac{1}{NT}[A(\hat{\rho})Y - S(\hat{\beta})A(\hat{\rho})Y]'H[A(\hat{\rho})Y - S(\hat{\beta})A(\hat{\rho})Y] - \frac{1}{T}\hat{\sigma}_\varepsilon^2 \tag{19}$$

步骤 3　用步骤 2 得到的 $\hat{\lambda}$ 替代式(11)中的 λ，则非参数函数 $g(u)$ 的最终估计为：

$$\hat{g}(u) = s(u,\hat{\beta})A(\hat{\rho})Y \tag{20}$$

四、估计的大样本性质

(一)假设条件

记 $\theta_0 = (\lambda'_0, \sigma'_0)'$ 为待估参数的真实值，其中，$\lambda_0 = (\beta'_0, \rho_0)'$，$\sigma_0 = (\sigma^2_{\varepsilon,0}, \sigma^2_{b,0})'$。为了推导 $\hat{\theta} = (\hat{\beta}', \hat{\rho}, \hat{\sigma}^2_\varepsilon, \hat{\sigma}^2_b)'$ 和 $\hat{g}(u)$ 的一致性和渐近正态性，我们需要建立适当的正则假设条件。

条件 1，关于模型中变量的假设条件：

(1) $\{x_{ij}, y_{ij}, \varepsilon_{ij}\}_{i=1,j=1}^{N,T}$ 为 $i.i.d.$ 随机序列，$\{x_{ij}, y_{ij}, b_i, \varepsilon_{ij}\}_{i=1}^{N}$ 在时刻 j 固定时为 $i.i.d.$ 随机序列，$j = 1,\cdots,T$，$x'_{ij}\beta$ 的边际密度函数 $f_j(u)$ 在 $u_{ij,0} = x'_{ij}\beta_0 \in U$ 处连续可微，同时 $f_j(u)$ 一致有界且不为零，其中，U 为 $k(u)$ 的支撑集；ε_{ij} 和 b_i 相互独立，b_i 满足 $E(b_i \mid x_1,\cdots,x_{NT}) = 0$，$Var(b_i \mid x_{11},\cdots,x_{NT}) = \sigma_b^2 < \infty$ 及 $E(\|b_i x'_{ij}\|) < \infty$；$\varepsilon_{ij}$ 满足 $E(\varepsilon_{ij} \mid x_1,\cdots,x_{NT}) = 0$，$Var(\varepsilon_{ij} \mid x_{11},\cdots,x_{NT}) = \sigma_\varepsilon^2 < \infty$ 及 $E(\|\varepsilon_{ij} x'_{ij}\|) < \infty$；其中，$i = 1,\cdots,N$，$j = 1,\cdots,T$。

(2) 实值函数 $g(\cdot)$ 为二阶连续可微的有界函数，在 $u \in U$ 上满足一阶 Lipschitz 条件，对于任意支撑集上的点 u 都有 $|g(u)| \leq m_g$，其中，m_g 为正常数。

条件 2，于模型中常量的假设条件：

(1) W_0 的对角元为零，非对角元 $w_{0,ij}$ 一致小于 $O(1/l_N)$，并且 $\lim_{N\to\infty} l_N/N = 0$。对任意 $\rho \in \Theta$，$I_N - \rho W_0$ 非奇异，其中，Θ 为凸紧集，ρ_0 为 Θ 的内点。

(2) W_0 和 $(I_N - \rho W_0)^{-1}$ 在 $\rho \in \Theta$ 上满足绝对行和与绝对列和一致有界①。

条件 3，关于核函数的假设条件：

$k(\cdot)$ 的支撑集为有界闭集，且在其支撑集上为连续非负的偶函数。即存在常数 $m_k > 0$，使得支撑集 $[-m_k, m_k] \subset R$，当 $|v| \leq m_k$ 时，$k(v) \geq 0$；令 $\mu_l = \int k(v)v^l dv$，$\nu_l = \int k^2(v)v^l dv$，则对于任意的正奇数 l，$\mu_l = v_l = 0$，同时 $\mu_0 = 1$，$\mu_2 \neq 0$。

条件 4，关于窗宽的假设条件：在 $N \to \infty$ 及 $h \to 0$ 时，$Nh \to \infty$。

条件 5，参数估计唯一性条件：存在唯一的 $\theta = \theta_0$，使得模型(2)成立。

条件 6，模型(2)中单指数部分可识别的条件：β_0 为 B 的内点，其中，$B \in R^p$ 为凸紧集，且 $\|\beta_0\| = 1$ 而且向量 β_0 的第一个分量为正，其中，$\|\cdot\|$ 表示欧几里得范数②。

条件 7，$\lim\limits_{N\to\infty} \frac{1}{N}(\dot{G}(X\beta_0)X \quad B_0G(X \quad \beta_0))' \left(\frac{H}{\sigma^2_{\varepsilon,0} + T\sigma^2_{b,0}} + \frac{I-H}{\sigma^2_{\varepsilon,0}}\right)(\dot{G}(X\beta_0)X \quad B_0G(X\beta_0))$ 存在并且非奇异，其中，$B_0 = WA^{-1}(\rho_0)$。

条件 8，$\lim\limits_{N\to\infty} \frac{1}{N}\{tr[(B'_0)^2] + tr(B_0 \quad B'_0) - \frac{2}{N(T-1)}[tr((I-H) \quad B_0)]^2 - \frac{2}{N}[tr(HB_0)]^2\} > 0$。

以上条件中，条件 1 和条件 2 描述了本文引入的模型及其空间权重矩阵的特征；条件 3 和条件 4 给出了核函数和窗宽条件；条件 5 和条件 6 为唯一性识别条件；条件 7 和条件 8 是参数部分渐近正态性条件。

(二)主要结论

在给出估计量的大样本性质之前，先列出几个有用的引理。

引理 1：在假设条件 1 ~4 下，$N^{-1}Z'(u,\beta)K(u,\beta)Z'(u,\beta) \xrightarrow{p} \sum_{j=1}^{T} f_j(u)\,diag\{1,\mu_2\}$。

引理 2：在假设条件 1 ~4 下，$S(\beta)(Ub+\varepsilon) = op(1)$ ③，$S(\beta)B(Ub+\varepsilon) = op(1)$，并且 $(I - S(\beta))G = op(1)$，$(I - S(\beta))BG = Op(c)$，其中，$c > 0$ 为常数。

引理 3：设 A 为 $NT \times NT$ 对称矩阵，$A_{ij,st}$ 为其第 ij 行第 st 列元素，则二次型 $Q = (Ub+\varepsilon)'A(Ub+\varepsilon)$ 的期望和方差分别为：

---

① 矩阵 $D = (D_{ij})_{N\times N}$ 绝对行(或列)和一致有界是指存在非负常数 m_D 使得 $\sum_{i=1}^{N}|D_{ij}| \leq m_D$ 及 $\sum_{j=1}^{N}|D_{ij}| \leq m_D$，常数 m_D 在下文中会根据不同的矩阵进行不同定义。

② Carroll 等(1997)，Wang 等(2010)也采用此识别条件。

③ 对于矩阵向量，这里表示每一个分量均为 $op(1)$。

$$E(Q) = \sigma_b^2 tr(\tilde{A}) + \sigma_\varepsilon^2 tr(A)$$

$$Var(Q) = (\xi_4 - 3\sigma_b^4)\sum_{i=1}^{N}\tilde{A}_{ii}^2 + 2\sigma_b^4[tr(\tilde{A})^2]$$

$$+ 2\sigma_\varepsilon^4[tr(A)^2] + (\eta_4 - 3\sigma_\varepsilon^4)\sum_{i=1}^{N}\sum_{j=1}^{T}A_{ij,ij}^2$$

$$+ 4\sigma_b^2\sigma_\varepsilon^2[\sum_{s=1}^{N}\sum_{t_1=1}^{T}\sum_{i=1}^{N}\sum_{j=1}^{T}(A_{ij,st_1}\sum_{t_2=1}^{T}A_{ij,st_2})]$$

其中，$\tilde{A} = U'AU$，$\tilde{A}_{ij}$ 为其第 i 行第 j 列元素，$\xi_4 = Eb_i^4$，$\eta_4 = E\varepsilon_{ij}^4$。进一步可得:

$$(Q - E(Q))/\sqrt{Var(Q)} \xrightarrow{L} N(0,1)$$

引理 4: 若 A 为 N 阶方阵，有 $tr(A^2) \leq tr(A'A)$。进一步，若 A 的 N 个特征值均为实数并且其中 k 个非零，则有 $tr(A^2) > 0$，并且 $(trA)^2/tr(A^2) \leq k$。

在给出主要结论前，记:

$\Theta_0 = \{\rho \mid \rho \in \Theta, \|\rho - \rho_0\| \leq c_1 N^{-1/2}\}$，$B_0 = \{\beta \mid \beta \in B, \|\beta\| = 1, \|\beta - \beta_0\| \leq c_2 N^{-1/2}\}$

其中，c_1 和 c_2 为正常数，类似约束参见 Härdle 等(1993)。

定理 1: 在假设条件 1 ~ 6 下，$\hat{\rho} - \rho_0 = op(1)$，$\hat{\beta} - \beta_0 = op(1)$。

定理 2: 在假设条件 1 ~ 6 下，$\hat{\sigma}_\varepsilon{}^2 - \sigma_{\varepsilon,0}^2 = op(1)$，$\hat{\sigma}_b{}^2 - \sigma_{b,0}^2 = op(1)$，对 $\rho \in \Theta_0$ 与 $\beta \in B_0$ 成立。

定理 3: 在假设条件 1 ~ 6 下，$\hat{g}(u) - g(u) = op(1)$，对 $\rho \in \Theta_0$，$\beta \in B_0$ 及 $u \in U$ 成立。

定理 4: 在假设条件 1 ~ 7 下或假设条件 1 ~ 6 和条件 8 下，

$$\sqrt{N}(\hat{\theta} - \theta_0) \xrightarrow{L} N(0, \Sigma_{\theta_0}^{-1})$$

其中，$\Sigma_{\theta_0} = -\lim_{N\to\infty} E\left[\frac{1}{N}\frac{\partial^2 \ln L(\theta)}{\partial\theta\partial\theta'}\Big|_{\theta=\theta_0}\right]$ ①。

定理 1 和定理 2 分别给出了模型参数估计量 $(\hat{\beta}', \hat{\rho})'$ 和 $(\hat{\sigma}_\varepsilon^2, \hat{\sigma}_b^2)'$ 的一致性，定理 3 给出了模型非参数估计量 $\hat{g}(u)$ 的一致性，定理 4 则给出了所有参数估计量的联合渐近正态分布。

五、数值模拟

估计量的小样本表现如何，对于实际应用非常重要。下面我们将利用蒙特卡洛数值模拟方法评估上面构建的估计量小样本效果。对于模型中参数部分的估计，使用的评价标准为样本标准差以及两种均方根误 RMSE1 和 RMSE2:

$$RMSE1 = \left[\frac{1}{mcn}\sum_{i=1}^{mcn}(\hat{\theta}_i - \theta_0)^2\right]^{1/2}, RMSE2 = (\hat{\theta}_{0.5} - \theta_0)^2 + \frac{\hat{\theta}_{0.75} - \hat{\theta}_{0.25}}{1.35}$$

其中，mcn 为模拟次数，$\hat{\theta}_i$ 为每次模拟所得到的参数估计值，$i = 1, \cdots, mcn$，θ_0 为参数真实值，$\hat{\theta}_{0.25}$、$\hat{\theta}_{0.5}$、$\hat{\theta}_{0.75}$ 分别为参数估计的上四分位数、中位数和下四分位数; 对于模型中的非参数部分估计，采用平均绝对误差作为评价标准，其计算公式为:

$$MADE_j = Q^{-1}\sum_{q=1}^{Q}|\hat{g}_j(u_q) - g_j(u_q)|, j = 1, \cdots, mcn$$

其中，$\{u_q\}_{q=1}^{Q}$ 为 u 的支撑集内所选取的 Q 个固定网格点。

由于难以选择最优窗宽(Su，2012)，我们利用拇指准则在非参数估计时选择窗宽，这里使用的核函数为常用的 Epanechnikov 核函数。

(一) 数据生成过程

我们考虑以下数据生成过程，对于模型(2)，通过代入已知参数数据生成样本，假设模拟样本为两期，即 $T = 2$。对于数据生成过程，我们做如下具体设计:

1. 令 x_{ij}，$i = 1, \cdots, N$，$j = 1, \cdots, T$ 为二维随机变量，其每个分量服从均匀分布 $U = (-3, 3)$，$\beta_0 = (\beta_{10}, \beta_{20})' = (\sqrt{1/3}, \sqrt{2/3})' \approx (0.5774, 0.8165)'$，$g(u_{ij}) = 2e^{-u_{ij}^2}$，$u_{ij} = x'_{ij}\beta_0$。随机误差项 ε 的各分量 $\varepsilon_{ij} \sim i.i.d. N(0, \sigma_{\varepsilon,0}^2)$，随机效应 b 的各分量 $b_i \sim i.i.d. N(0, \sigma_{b,0}^2)$，其中，$\sigma_0 = (\sigma_{\varepsilon,0}^2, \sigma_{b,0}^2)' = (0.5, 1)'$。

2. 取 $\rho_0 = 0.5$，为了考察空间权重矩阵的影响，使用 Case(1991) 中一类权重矩阵②，这里 $N = R \times M$，分别取 $R = 20$、30、40 和 $M = 2$、4 进行模拟得到结果。另外，为了比较不同空间权重矩阵的影响，同时采用实际中常用的 Rook 权重矩阵③，分别取 $N = 49$、64、81、100 进行模拟以对比最终结果。

---

① 限于篇幅，有关引理和定理的证明过程均未列出，有兴趣的读者可向作者索要。

② Case(1991) 权重矩阵假设存在 R 个区域，每个地区有 M 个成员，同一地区内的成员互为"邻居"并且具有相同的权重，满足 $W_0 = I_R \otimes B_M$，其中，$B_M = (1_M 1'_M - I_M)/(M-1)$。

③ 参考 Su(2012)，Rook 相邻型矩阵指每一个空间单元和其成直角关系的空间单元连接，因此居中的单元具有四个相邻空间，边界的单元具有三个相邻空间，角落的单元仅具有两个相邻空间。

（二）数据模拟结果

利用 Matlab 进行 300 次模拟，记录每次模拟结果，计算各项评级指标，参数结果见表 1。

表 1 参数模拟结果（Case 权重矩阵）

R	参数	$M=2$ Mean	Std. dev	RMSE1	RMSE2	$M=4$ Mean	Std. dev	RMSE1	RMSE2
20	ρ	0.4956	0.0562	0.0562	0.0558	0.4935	0.0764	0.0765	0.0542
	β_1	0.5693	0.0666	0.0670	0.0559	0.5730	0.0392	0.0394	0.0372
	β_2	0.8179	0.0502	0.0502	0.0386	0.8182	0.0273	0.0273	0.0261
	σ_ε^2	0.5654	0.1360	0.1507	0.1436	0.5556	0.0908	0.1063	0.0973
	σ_b^2	0.8613	0.2857	0.3172	0.1818	0.9321	0.1999	0.2108	0.0764
30	ρ	0.4957	0.0458	0.0459	0.0470	0.4937	0.0407	0.0411	0.0396
	β_1	0.5710	0.0437	0.0440	0.0426	0.5738	0.0292	0.0293	0.0275
	β_2	0.8192	0.0304	0.0305	0.0296	0.8182	0.0204	0.0204	0.0193
	σ_ε^2	0.5564	0.1190	0.1315	0.1366	0.5440	0.0680	0.0809	0.0736
	σ_b^2	0.9078	0.2452	0.2615	0.1402	0.9520	0.1694	0.1758	0.0318
40	ρ	0.4998	0.0366	0.0366	0.0367	0.4959	0.0373	0.0375	0.0368
	β_1	0.5713	0.0365	0.0369	0.0337	0.5762	0.0245	0.0245	0.0244
	β_2	0.8196	0.0255	0.0256	0.0235	0.8167	0.0173	0.0173	0.0172
	σ_ε^2	0.5512	0.0936	0.1065	0.1015	0.5404	0.0605	0.0727	0.0612
	σ_b^2	0.9232	0.1889	0.2036	0.0785	0.9583	0.1355	0.1415	0.0022

观察表 1，我们发现，参数 ρ、σ_ε^2、σ_b^2、β_1 和 β_2 的估计值与其真实值在各类情况下均较为接近，说明各估计量在小样本下均有较好的表现，表明本文所提出的估计方法具有优良的实用性，同时我们发现：

第一，当空间复杂度固定时（M 相同），所有参数的估计值与其真实值的偏误随地区数 R 的增加而减小，主要表现为 RMSE1 或 RMSE2 随 R 的增加而减少，说明在空间复杂度相同的情况下，样本容量越大，参数估计的偏误越小，同时可以看到参数估计值的标准差随着地区数 R 的增加而减小，说明参数的估计值会随着样本容量的增大而收敛。

第二，在地区数固定时（R 相同），参数 ρ 的估计偏误没有因空间复杂度 M 的增加而减少，与参数 ρ 不同，σ_ε^2、σ_b^2、β_1 和 β_2 在 M 为 4 时的估计偏误和标准差均小于 M 为 2 时的相应结果，说明在地区数 R 相同情况下，空间复杂度的增加会抵消由于样本容量增加所带来的参数 ρ 的估计偏误的减少。

第三，在样本容量固定时（N 相同），参数 ρ 的估计偏误会因空间复杂度的增加而增加，其他参数的估计偏误和标准差受空间复杂度的影响较小。为说明结论，考虑样本 $N = 80$，即 $R = 20$，$M = 4$ 与 $R = 40$，$M = 2$ 的情形，我们发现当 $R = 20$，$M = 4$ 时，ρ 的估计偏误和标准差均比 $R = 40$，$M = 2$ 时大，而其他参数的估计偏误和标准差在两种情形下相近。

表 2 给出了未知函数 $g(\cdot)$ 在 20 个固定格点处估计的 MADE 值的中位数和标准差。对比模拟结果发现，未知函数的估计效果与空间复杂度（M 的取值）关系不大，主要受样本容量（N 的取值）的影响，随着样本容量的增大，未知函数 $g(\cdot)$ 的 MADE 的中位数和标准差均趋于下降，说明未知函数的估计是收敛的。同时，样本容量的增加会改善未知函数的估计效果。

表 2 未知函数 $g(\cdot)$ 估计的 MADE 值的中位数和标准差（Case 权重矩阵）

统计量	$M=2$ $R=20$	$R=30$	$R=40$	$M=4$ $R=20$	$R=30$	$R=40$
Median	0.3238	0.2910	0.2678	0.2610	0.2349	0.2108
Std. dev	0.0959	0.0593	0.0536	0.0478	0.0452	0.0393

为了说明不同空间权重矩阵的影响，再利用 Rook 邻接矩阵作为空间权重矩阵，进行模拟以对比最终结果。由表 3 可知，与权重矩阵为 Case 时类似，参数 ρ、σ_ε^2、σ_b^2、β_1 和 β_2 的估计值与其真实值在各类情况下均较为接近，说明各估计量在小样本下均有较好的表现，表明本文所提出的估计方法具有优良的实用性，同时我们发现：所有参数的估计值与真实值的偏误随样本容量的增加而减小，主要表现为 RMSE1 或 RMSE2 随 N 的增加而减少，说明样本容量越大，参数估计的偏误越小；从表 3 中还可以看到，参数估计值的标准差随样本数的增加而减小，这说明参数的估计值会随着样本容量的增大而收敛。结合上述两结果知参数估计值会随着样本容量的增大而收敛到参数的真值。

表 3 参数模拟结果（Rook 权重矩阵）

参数	$N=49$ Mean	Std. dev	RMSE1	RMSE2	$N=64$ Mean	Std. dev	RMSE1	RMSE2
ρ	0.4577	0.1041	0.1122	0.1069	0.4785	0.0920	0.0944	0.0958
β_1	0.5711	0.0541	0.0544	0.0542	0.5706	0.0442	0.0447	0.0441
β_2	0.8181	0.0390	0.0389	0.0377	0.8195	0.0309	0.0310	0.0308
σ_ε^2	0.5644	0.1193	0.1354	0.1140	0.5585	0.0989	0.1147	0.1067
σ_b^2	0.9024	0.2419	0.2604	0.1345	0.9219	0.2198	0.2329	0.0826
	$N=81$ Mean	Std. dev	RMSE1	RMSE2	$N=100$ Mean	Std. dev	RMSE1	RMSE2
ρ	0.4751	0.0760	0.0799	0.0798	0.4861	0.0663	0.0676	0.0701
β_1	0.5741	0.0355	0.0356	0.0356	0.5716	0.0320	0.0325	0.0312
β_2	0.8177	0.0249	0.0249	0.0250	0.8196	0.0223	0.0225	0.0218
σ_ε^2	0.5533	0.0862	0.1012	0.0859	0.5495	0.0774	0.0918	0.0818
σ_b^2	0.9319	0.2001	0.2111	0.0726	0.9342	0.1779	0.1894	0.0425

表 4 给出了未知函数 $g(\cdot)$ 在 20 个固定格点处的估计值的 300 个 MADE 值的中位数和标准差。对

比表 4 中的 4 次模拟结果发现,随着样本容量的增大,未知函数 $g(\cdot)$ 估计的 MADE 值的中位数和标准差均趋于下降,说明未知函数的估计是收敛的。同样地,未知函数的估计效果会随样本容量的增加而改善。

表 4　未知函数 $g(\cdot)$ 估计的 MADE 值的中位数和标准差(Rook 权重矩阵)

统计量	$N=49$	$N=64$	$N=81$	$N=100$
Median	0. 3156	0. 2935	0. 2661	0. 2495
Std. dev	0. 0992	0. 0718	0. 0565	0. 0543

六、总结

非参数回归模型往往存在“维数灾难”问题,导致估计的精度会随着解释变量维数的增加而降低,而单指数回归模型既有非参数的自身优势同时又有规避“维数灾难”的降维功能。据此,针对空间面板数据,本文提出了一种新的具有随机效应的空间滞后单指数面板模型,针对该模型,我们在传统极大似然估计方法不适用的情况下,借助“截面极大似然估计”的思想,构建了该模型的截面极大似然估计方法。在此基础上,重点从理论证明和数值模拟两方面进行了研究,相关结果总结如下: 在一定的正则假设条件下,证明了我们所构建的有关未知参数和函数的估计量具有一致性,并进一步证明了有关未知参数估计量的渐近正态性,表明估计方法具有优良的渐近性质。在小样本的情况下,所有估计量的偏误(RMSE 或 MADE) 和标准差均较小,表明本文提出的估计方法依然具有优良的小样本表现; 估计量的偏误(RMSE 或 MADE) 和标准差均会随着样本容量的增加而下降,说明其表现与理论结果一致; 空间相关系数的估计量对于空间复杂性较为敏感,其精度在一定程度上会受空间复杂性的影响,对相同容量的样本,空间权重矩阵的结构越复杂,估计量的标准差和偏误越大,但样本容量的增加能够减少这种影响,使空间相关系数估计量的稳健性得到改善,其他估计量对于空间复杂性则具有良好的稳健性,即精度受空间权重矩阵结构复杂性的影响较小。

参考文献

[1] Ai C. , Chen X. . Efficient Estimation of Models with Conditional Moment Restrictions Containing Unknown Functions [J]. Econometrica, 2003, 71(6): 1795 -1843.

[2] Carroll R. J. , Fan J. , Gijbels I. , Wand M. P. . Generalized Partially Single-Index Models [J]. Journal of the American Statistical Association, 1997, 92(438): 477 -489.

[3] Case A. C. . Spatial Patterns in Household Demand [J]. Econometrica, 1991, 59(4): 953 -965.

[4] Chen J. , Gao J. T. , Li D. G. . Estimation in Partially Linear Single-index Panel Data Models with Fixed Effects [J]. Journal of Business & Economic Statistics, 2013, 31(3): 1 -42.

[5] Cui X. , Härdle W. , Zhu L. X. . The EFM Approach for Single-index Models [J]. Annals of Statistics, 2011, 39(3): 1658 -1688.

[6] Fan J. Q. , Gijberts I. . Local Polynomial Modelling and Its Applications [M]. London: Chapman & Hall, 1996.

[7] Friedman J. H. , Stuetzle W. . Projection Pursuit Regression [J]. Journal of the American Statistical Association, 1981, 76(376): 817 -823.

[8] Gress B. . Using Semi-parametric Spatial Autocorrelation Models to Improve Hedonic Housing Price Prediction [R]. Working Paper, UC Riverside Economics Department, 2004.

[9] Härdle W. , Hall P. , Ichimura H. . Optimal Smoothing in Single-index Models [J]. Annals of Statistics, 1993, 21(1): 157 -178.

[10] Ichimura H. . Semiparametric Least Squares (SLS) and Weighted SLS Estimation of Single-index Models [J]. Journal of Econometrics, 1993, 58(1 -2): 71 -120.

[11] Kelejian H. H. , Prucha I. R. . Specification and Estimation of Spatial Autoregressive Models with Autoregressive and Heteroskedastic Disturbances [J]. Journal of Econometrics, 2010, 157(1): 53 -67.

[12] Lee L. F. , Yu J. H. . Some Recent Developments in Spatial Panel Data Models [J]. Regional Science and Urban Economics, 2010, 40: 255 -271.

[13] Pang Z. , Xue L. G. . Estimation for the Single-index Models with Random Effects [J]. Computational Statistics and Data Analysis, 2012, 56(6): 1837 -1853.

[14] Robinson P. M. . Efficient Estimation of the Semiparametric Spatial Autoregressive Model [J]. Journal of Econometrics, 2010, 157(1): 6 -17.

[15] Su L. J. . Semi-parametric GMM Estimation of Semiparametric Spatial Autoregressive Models [J]. Journal of Econometrics, 2012, 167(2): 543 -560.

[16] Wang J. L. , Xue L. G. , Zhu L. X. , Chong Y. S. . Estimation for a Partial-linear Single-index Model [J]. Annals of Statistics, 2010, 38(1): 246 -274.

作者简介

陈建宝,男,1965 年生,云南曲靖人,2003 年毕业于澳大利亚科庭理工大学统计学专业,获哲学博士学位,现为厦门大学经济学院副院长,统计系教授、博士生导师。研究方向为统计理论与方法、经济计量。

孙林,男,1986 年生,山东菏泽人,现为厦门大学经济学院统计系博士研究生。研究方向为统计理论和方法、经济计量。

(责任编辑: 曹　麦)