随机效应广义空间滞后半参数变系数面板模型的估计*

唐礼智　刘 玉

内容提要: 本文通过构建同时包含因变量和误差项空间滞后的随机效应半参数变系数面板模型，拓展了现有模型的灵活性和适应性; 采用截面极大似然估计方法得出了参数和非参数的估计。理论证明发现: 在一定的正则条件下，所有估计量具有一致性和渐近正态性。数值模拟显示: 估计量具有良好的小样本性质，估计精度随着样本容量的增加而增加; 空间权重矩阵的选择对估计量的表现没有产生显著差异，但是在 Case 权重矩阵下，当样本量相同时，空间相关系数的估计偏差随着空间权重结构复杂度的增加而扩大。

关键词: 半参数; 变系数; 广义空间滞后; 面板模型

DOI: 10. 19343/j. cnki. 11 －1302/c. 2018. 02. 011

中图分类号: O212　　**文献标识码**: A　　**文章编号**: 1002 －4565(2018) 02 －0119 －10

Estimation of Generalized Spatial Lag Semi-parametric Varying-coefficient Panel Model with Random Effects

Tang Lizhi & Liu Yu

Abstract: By constructing a semi-parametric varying-coefficient random effect panel model with spatial lag and spatial error, the flexibility and adaptability of the existing models are extended. It uses profile maximum likelihood method to construct both parametric and nonparametric estimators. The theoretical proof shows that all estimators have consistency and asymptotic normality under certain regular conditions. The numerical simulation shows that the estimators have good small sample properties, and the estimation accuracy increases with the increase of the sample capacity. The selection of the spatial weight matrix does not produce significant difference in the performance of the estimator, but under the condition of Case matrix, when the sample quantities are the same, the estimated deviation of spatial correlation coefficients increase with the increase of spatial weight structure complexity.

Key words: Semi-parametric; Varying-coefficient; Generalized Spatial Lag; Panel Model

一、引言

空间计量模型关注个体的空间相关性和异质性，极大地拓展了传统计量经济模型的研究范畴。早期的空间计量模型主要采用参数(线性) 模型设定形式，Kelejian 和 Prucha(1998) [1] 通过广义两阶段最小二乘法构造了同时包含误差空间滞后和因变量空间滞后的截面模型的估计量，并且验证其具有一致性和渐近正态性; Lee(2004) [2] 构造了截面空间滞后(SAR) 模型的拟极大

* 本文为教育部基地重大项目“矫正要素配置扭曲与促进经济有效增长”(16JJD790031) 的阶段性成果。

似然估计，通过证明得出估计量满足 $\sqrt{n}$ 一致收敛性和渐近正态性; Elhorst(2003) [3] 建立了固定效应、随机效应面板 SAR 和空间误差模型(SEM) 的识别和估计方法; Kapoor 等(2007) [4] 构建了面板空间误差分量模型(KKP 模型) 的广义矩估计(GMM) ，并证明其具有良好的大样本性; Lee 和 Yu(2010) [5] 研究了同时含有空间滞后项与空间误差项的固定效应面板模型，通过正交变换克服了“伴生参数”问题，得到了一致收敛的参数估计; Moscone 和 Tosetti(2011) [6] 在固定效应面板 SEM 的基础上引入扰动项的异方差性，采用 GMM 对参数进行估计并得出了估计量的渐近性质。

但是，单纯的参数(线性) 空间计量模型并不能很好地解释复杂的经济现象。为了增强现有参数空间计量模型的灵活性和适用性，非参数(半参数) 空间计量模型的研究逐步兴起。Basile (2008) [7] 将半参数估计方法与参数空间杜宾模型相结合对欧洲 155 个地区的经济增长进行了实证研究，结论表明初始人均收入和人力资本投资确实存在非线性效应; Su 和 Jin(2010) [8] 构建了部分线性空间自回归模型的截面拟极大似然估计; Su(2012) [9] 在非参数的框架下同时引入因变量和误差项的空间滞后效应，采用 GMM 得出了渐近一致估计量; 李坤明和陈建宝(2013) 构建了变系数 SAR 模型的截面极大似然估计量并证明了估计量的渐近性质; 陈建宝和孙林(2015) [10] 创新性地构建了随机效应单指数空间滞后面板模型，采用截面极大似然估计得到参数和非参数的一致估计量，并证明估计量满足渐近正态性; Zhang 和 Shen(2015) [11] 构建了包含内生解释变量的随机效应半参数变系数 SAR 面板模型，通过工具变量法得出了渐近一致估计量; Bai 等(2015) [12] 则在半参数变系数的框架下，设定误差项同时存在空间效应和时序相关性，分别构建了参数的加权最小二乘估计量和非参数的加权多项式样条估计量，并证明了估计量的大样本性质; 陈建宝和乔宁宁(2017) [13] 采用截面极大似然估计得出半参数变系数 SEM 的渐近一致估计量; 陈建宝和孙林(2017) [14] 对随机效应变系数 SAR 面板模型构建了截面极大似然估计，并且证明了估计量的一致性和渐近正态性。

综上所述，现有非参数(半参数) 空间计量模型存在以下不足: 一是空间效应的存在方式单一化，空间滞后和空间误差未同时兼顾，不能全面考察空间效应的影响; 二是非参数模型形式易产生“维数灾难”，如 Su 和 Jin(2010) [8] 、Su(2012) [9] 。虽然陈建宝和孙林(2017) [14] 采用了变系数形式，但是假设所有系数均是变系数可能与实际相悖; 三是部分研究采用截面数据，无法刻画个体效应、控制异质性，并且可能存在多重共线性问题。因此，本文创新性地设定了随机效应广义空间滞后半参数变系数面板模型，不仅同时包含可观测的被解释变量和不可观测的误差项的空间滞后效应，而且假设外生解释变量系数一部分是常数，另一部分是随着某个外生变量变化的未知函数，显著提高了模型的灵活性和适应性。

二、模型设定

本文假设空间相关性同时存在于扰动项和个体效应中(即遵从 KKP 模型的设定) ，则随机效应广义空间滞后半参数变系数面板模型如下所示:

$$\begin{cases} y_{it} = \rho \sum_{j=1}^{N} w_{ij}^{1} y_{jt} + x_{it}' \beta + z_{it}' \gamma(u_{it}) + v_{it} \\ v_{it} = \lambda \sum_{j=1}^{N} w_{ij}^{2} v_{it} + \mu_{it} \\ \mu_{it} = \alpha_i + \varepsilon_{it} \end{cases} \tag{1}$$

其中，y_{it} 是被解释变量，$x_{it}=(x_{it}^1,\cdots,x_{it}^p)'$ 是 $p\times1$ 维解释变量，$z_{it}=(z_{it}^1,\cdots,z_{it}^q)'$ 是 $q\times1$ 维变系数解释变量并且与 x_{it} 不相关。$\beta=(\beta_1,\cdots,\beta_p)'$ 是一个 $p\times1$ 维常系数向量，$\gamma(u_{it})=(\gamma_1(u_{it}),\cdots,\gamma_q(u_{it}))'$ 是 $q\times1$ 维未知变系数函数，为了简化推导过程，此处只考察 u_{it} 是一元变量的情形。ρ 和 λ 是空间相关系数，反映了空间相关性的强弱；w_{ij}^1 和 w_{ij}^2 是外生给定的常数，反映空间中单元 i 对单元 j 的影响大小，在现实中两者可能相等也可能不相等。$\alpha_i\sim_{iid}N(0,\sigma_\alpha^2)$，$\varepsilon_{it}\sim_{iid}N(0,\sigma_\varepsilon^2)$ 分别表示空间特定效应和扰动项并且满足 α_i 与 ε_{it} 相互独立。

为了分析方便，本文对面板数据先按时间再按个体排列，则有：

$Y=(y_{11},\cdots,y_{N1}\cdots,y_{1T},\cdots,y_{NT})'$，$X_t=(x_{1t},\cdots,x_{Nt})'$，$X=(X_1',\cdots,X_T')'$，

$M=(z_{11}'\gamma(u_{11}),\cdots,z_{N1}'\gamma(u_{N1}),\cdots,z_{1T}'\gamma(u_{1T}),\cdots,z_{NT}'\gamma(u_{NT}))'$，

$\alpha=(\alpha_1,\cdots,\alpha_N)'$，$v=(v_{11},\cdots,v_{N1},\cdots,v_{1T},\cdots,v_{NT})'$，$W_1=(w_{ij}^1)_{N\times N}$，$W_2=(w_{ij}^2)_{N\times N}$

$\mu=(\mu_{11},\cdots,\mu_{N1},\cdots,\mu_{1T},\cdots,\mu_{NT})'$，$\varepsilon=(\varepsilon_{11},\cdots,\varepsilon_{N1},\cdots,\varepsilon_{1T},\cdots,\varepsilon_{NT})'$，$e_T=(1,\cdots,1)'$

$\widetilde{W}_1=I_T\otimes W_1$，$\widetilde{W}_2=I_T\otimes W_2$，$B(\rho)=I_{NT}-\rho\widetilde{W}_1$，$B(\lambda)=I_{NT}-\lambda\widetilde{W}_2$

因而式(1)的矩阵表示形式如下：

$$\begin{cases}Y=\rho\widetilde{W}_1Y+X\beta+M+v\\ v=\lambda\widetilde{W}_2v+\mu\\ \mu=e_T\otimes\alpha+\varepsilon\end{cases}\tag{2}$$

进一步化简可得其简约形式：

$$B(\lambda)[B(\rho)Y-X\beta-M]=\mu\tag{3}$$

记 $\theta=(\beta',\rho,\lambda,\sigma_\alpha^2,\sigma_\varepsilon^2)'\in\Theta$（$\Theta$ 是有限维的参数空间），则对于上述模型，首要任务是选择合适的方法估计参数 θ 和非参数函数 M。首先，本文尝试用极大似然估计的方法。

由 $\mu=e_T\otimes\alpha+\varepsilon$ 可知，$\Sigma_\mu=E(e_T\otimes\alpha+\varepsilon)(e_T\otimes\alpha+\varepsilon)'=\sigma_\varepsilon^2Q_{NT}+(\sigma_\varepsilon^2+T\sigma_\alpha^2)J_{NT}$，$|\Sigma_\mu|=\sigma_\varepsilon^{2N(T-1)}(\sigma_\varepsilon^2+T\sigma_\alpha^2)^{-N}$，$\Sigma_\mu^{-1}=\dfrac{Q_{NT}}{\sigma_\varepsilon^2}+\dfrac{J_{NT}}{\sigma_\varepsilon^2+T\sigma_\alpha^2}$。其中，$J_{NT}=\dfrac{e_Te_T'\otimes I_N}{T}$，$Q_{NT}=I_{NT}-J_{NT}$，$e_T$ 是元素全为 1 的 $T\times1$ 列向量。式(3)的对数似然函数如下：

$$L(\theta)=-\frac{NT}{2}\ln(2\pi)-\frac{N(T-1)}{2}\ln(\sigma_\varepsilon^2)-\frac{N}{2}\ln(\sigma_\varepsilon^2+T\sigma_\alpha^2)+\ln|B(\lambda)|+\ln|B(\rho)|-\frac{1}{2\sigma_\varepsilon^2}[B(\rho)Y-X\beta-M]'B'(\lambda)Q_{NT}B(\lambda)[B(\rho)Y-X\beta-M]-\frac{1}{2(\sigma_\varepsilon^2+T\sigma_\alpha^2)}[B(\rho)Y-X\beta-M]'B'(\lambda)J_{NT}B(\lambda)[B(\rho)Y-X\beta-M]\tag{4}$$

关于式(4)对参数 σ_ε^2 和 σ_α^2 求一阶导数可得极大似然估计：

$$\hat{\sigma}_\varepsilon^2(\beta',\rho,\lambda)=\frac{1}{N(T-1)}[B(\rho)Y-X\beta-M]'B'(\lambda)Q_{NT}B(\lambda)[B(\rho)Y-X\beta-M]\tag{5}$$

$$\hat{\sigma}_\alpha^2(\beta',\rho,\lambda)=\frac{1}{NT}[B(\rho)Y-X\beta-M]'B'(\lambda)J_{NT}B(\lambda)[B(\rho)Y-X\beta-M]-\frac{\hat{\sigma}_\varepsilon^2(\beta^T,\rho,\lambda)}{T}\tag{6}$$

将式(5)和式(6)代入式(4)可得 β，ρ 和 λ 的集中似然函数：

$$L(\beta',\rho,\lambda)=-\frac{NT}{2}[\ln(2\pi)+1]-\frac{N(T-1)}{2}\ln\hat{\sigma}_\varepsilon^2-\frac{N}{2}\ln(\hat{\sigma}_\varepsilon^2+T\hat{\sigma}_\alpha^2)+\ln|B(\lambda)|+\ln|B(\rho)|\tag{7}$$

式(7)中 M 未知，无法通过非线性迭代得到 β、ρ 和 λ 的估计值，故传统极大似然估计失效，需要寻找未知函数矩阵 M 的可行估计。

三、模型估计

借鉴 Fan 和 Huang(2005)[16] 的研究，本文采用截面极大似然估计法估计未知函数矩阵 M，通过局部线性化逼近未知函数 $\gamma(u)$。应注意的是，为了区分符号，本文用 $\dot{f}$ 表示函数 f 的一阶导数，用 A'、a' 表示矩阵 A 或者向量 a 的转置，具体步骤如下：

(一) 局部线性化

对 $\gamma(u_{it})$ 关于其很小的邻域中的一点 u 进行 Taylor 展开，则 $\gamma_j(u_{it}) \approx \gamma_j(u) + \dot{\gamma}_j(u)(u_{it}-u)$，$j=1,2,\cdots,q$。令：

$$S(u)=\begin{bmatrix} z'_{11} & h^{-1}(u_{11}-u)z'_{11} \\ \vdots & \vdots \\ z'_{N1} & h^{-1}(u_{N1}-u)z'_{N1} \\ \vdots & \vdots \\ z'_{1T} & h^{-1}(u_{1T}-u)z'_{1T} \\ \vdots & \vdots \\ z'_{NT} & h^{-1}(u_{NT}-u)z'_{NT} \end{bmatrix}$$

$$\delta(u)=(\gamma_1(u),\cdots,\gamma_q(u),h\dot{\gamma}_1(u),\cdots,h\dot{\gamma}_q(u))'$$

$$K(u)=\mathrm{diag}(k_h(u_{11}-u),\cdots,k_h(u_{N1}-u),\cdots,k_h(u_{1T}-u),\cdots,k_h(u_{NT}-u)),k_h(u_{it}-u)$$

$$=h^{-1}k\left(\frac{u_{it}-u}{h}\right)$$

其中 $k(\cdot)$ 是一元核函数，$K(u)$ 是权重矩阵，h 是窗宽，需要进行选择。给定 β' 和 ρ，则 $\delta(u)$ 满足：$\delta(u)=\mathrm{argmin}\dfrac{1}{NT}[B(\rho)Y-X\beta-S(u)\delta(u)]'K(u)[B(\rho)Y-X\beta-S(u)\delta(u)]$，因而：

$$\hat{\delta}(u)=[S(u)'K(u)S(u)]^{-1}S(u)'K(u)Y^*=R^{-1}(u)T(u) \tag{8}$$

其中：

$$Y^*=B(\rho)Y-X\beta=(y^*_{11}\ \cdots\ y^*_{NT})'$$

$$R(u)=\frac{1}{NT}S(u)'K(u)S(u)=\frac{1}{NT}\begin{bmatrix} \sum_{t=1}^{T}\sum_{i=1}^{N}z_{it}z'_{it}k_h(u_{it}-u) & \sum_{t=1}^{T}\sum_{i=1}^{N}z_{it}z'_{it}\left(\frac{u_{it}-u}{h}\right)k_h(u_{it}-u) \\ \sum_{t=1}^{T}\sum_{i=1}^{N}z_{it}z'_{it}\left(\frac{u_{it}-u}{h}\right)k_h(u_{it}-u) & \sum_{t=1}^{T}\sum_{i=1}^{N}z_{it}z'_{it}\left(\frac{u_{it}-u}{h}\right)^2k_h(u_{it}-u) \end{bmatrix}$$

$$T(u)=\frac{1}{NT}S(u)'K(u)Y^*=\frac{1}{NT}\begin{bmatrix} \sum_{t=1}^{T}\sum_{i=1}^{N}z_{it}y^*_{it}k_h(u_{it}-u) \\ \sum_{t=1}^{T}\sum_{i=1}^{N}z_{it}y^*_{it}\left(\frac{u_{it}-u}{h}\right)k_h(u_{it}-u) \end{bmatrix}$$

令 $e=(I_q,0_q)'$，则变系数函数的初始估计为：

$$\hat{\gamma}_{IN}(u)=(\hat{\gamma}_1(u),\cdots,\hat{\gamma}_q(u))'=e'\hat{\delta}(u)$$

令 $L = \begin{bmatrix} [e'_{11} \quad 0_q][S(u_{11})'K(u_{11})S(u_{11})]^{-1}S(u_{11})'K(u_{11}) \\ \vdots \\ [e'_{N1} \quad 0_q][S(u_{N1})'K(u_{N1})S(u_{N1})]^{-1}S(u_{N1})'K(u_{N1}) \\ \vdots \\ [e'_{1T} \quad 0_q][S(u_{1T})'K(u_{1T})S(u_{1T})]^{-1}S(u_{1T})'K(u_{1T}) \\ \vdots \\ [e'_{NT} \quad 0_q][S(u_{NT})'K(u_{NT})S(u_{NT})]^{-1}S(u_{NT})'K(u_{NT}) \end{bmatrix}$

进而非参数部分的初始估计为:

$$\hat{M}_{IN} = LY^* \tag{9}$$

(二) 计算对数截面似然函数

将式(9) 带入式(4) ,可得近似初始对数截面似然函数如下:

$$\begin{aligned}\tilde{L}(\theta) = & -\frac{NT}{2}\ln(2\pi) - \frac{N(T-1)}{2}\ln(\sigma_\varepsilon^2) - \frac{N}{2}\ln(\sigma_\varepsilon^2 + T\sigma_\alpha^2) + \ln|B(\rho)| + \\ & \ln|B(\lambda)| - \frac{1}{2\sigma_\varepsilon^2}Y^{*\prime}(I_{NT} - L)'B'(\lambda)Q_{NT}B(\lambda)(I_{NT} - L)Y^* - \\ & \frac{1}{2(\sigma_\varepsilon^2 + T\sigma_\alpha^2)}Y^{*\prime}(I_{NT} - L)'B'(\lambda)J_{NT}B(\lambda)(I_{NT} - L)Y^*\end{aligned} \tag{10}$$

对式(10) 关于 σ_ε^2 和 σ_α^2 求一阶导数,可得初始似然估计:

$$\hat{\sigma}_{\varepsilon IN}^2(\beta',\rho,\lambda) = \frac{1}{N(T-1)}Y^{*\prime}(I_{NT} - L)'B'(\lambda)Q_{NT}B(\lambda)(I_{NT} - L)Y^* \tag{11}$$

$$\hat{\sigma}_{\alpha IN}^2(\beta',\rho,\lambda) = \frac{1}{NT}Y^{*\prime}(I_{NT} - L)'B'(\lambda)J_{NT}B(\lambda)(I_{NT} - L)Y^* - \frac{\hat{\sigma}_{\varepsilon IN}^2}{T} \tag{12}$$

(三) 计算集中似然函数

将式(11) 和式(12) 代入式(10) 可得关于 β'、ρ、λ 的集中对数似然函数:

$$\begin{aligned}\tilde{L}(\beta',\rho,\lambda) = & -\frac{NT}{2}[\ln(2\pi) + 1] - \frac{N(T-1)}{2}\ln\hat{\sigma}_{\varepsilon IN}^2 - \frac{N}{2}\ln(\hat{\sigma}_{\varepsilon IN}^2 + T\hat{\sigma}_{\alpha IN}^2) + \ln|B(\rho)| \\ & + \ln|B(\lambda)|\end{aligned} \tag{13}$$

对式(13) 进行非线性迭代可以得出 ρ、λ、β' 的最优估计 $\hat{\rho}$、$\hat{\lambda}$、$\hat{\beta}'$,进而方差以及变系数的最终估计为:

$$\hat{\sigma}_\varepsilon^2 = \frac{1}{N(T-1)}\hat{Y}^{*\prime}(I_{NT} - L)'B'(\hat{\lambda})Q_{NT}B(\hat{\lambda})(I_{NT} - L)\hat{Y}^* \tag{14}$$

$$\hat{\sigma}_\alpha^2 = \frac{1}{NT}\hat{Y}^{*\prime}(I_{NT} - L)'B'(\hat{\lambda})J_{NT}B(\hat{\lambda})(I_{NT} - L)\hat{Y}^* - \hat{\sigma}_\varepsilon^2/T \tag{15}$$

$$\hat{\gamma}(u) = e'[S(u)'K(u)S(u)]^{-1}S(u)'K(u)(B(\hat{\rho})Y - X\hat{\beta}) \tag{16}$$

其中,$\hat{Y}^* = B(\hat{\rho})Y - X\hat{\beta}$。

四、估计的大样本性质

通常期望所构造的估计量具有良好的大样本性质,例如一致性和渐近正态性,这些性质有助于进行统计推断和检验。为了方便证明参数估计量、非参数估计的一致性和渐近正态性,本文将模型改为如下简化形式:

$$Y = X\beta + M + \rho G_1(X\beta + M) + B^{-1}(\rho) B^{-1}(\lambda)\mu$$

其中，$B^{-1}(\rho) = I_{NT} + \rho G_1$，$G_1 = \tilde{W}_1 B^{-1}(\rho)$，$B^{-1}(\lambda) = I_{NT} + \lambda G_2$，$G_2 = \tilde{W}_2 B^{-1}(\lambda)$。进一步定义 $H = G_1(X\beta + M)$，并且令 $\theta_0 = (\beta_0', \rho_0, \lambda_0, \sigma_{\alpha 0}^2, \sigma_{\varepsilon 0}^2)'$ 和 M_0 分别是待估参数 θ 和变系数部分 M 的真实值，$B_1 = B(\rho_0)$，$B_2 = B(\lambda_0)$，$G_{10} = \tilde{W}_1 B^{-1}(\rho_0) = \tilde{W}_1 B_1^{-1}$，$G_{20} = \tilde{W}_2 B^{-1}(\lambda_0) = \tilde{W}_2 B_2^{-1}$，$H_0 = G_{10}(X\beta_0 + M_0)$。

（一）假设条件

1. 关于模型中变量的假设。

（1）$\{x_{it}\}_{i=1,t=1}^{N,T}$、$\{z_{it}\}_{i=1,t=1}^{N,T}$ 是独立同分布随机序列，且具有有界支撑集；$\{\varepsilon_{it}\}_{i=1,t=1}^{N,T}(i = 1,2,\cdots,N; t = 1,2,\cdots,T)$ 是独立同分布于 $N(0,\sigma_\varepsilon^2)$ 的随机序列，并且与 $\{x_{it}\}_{i=1,t=1}^{N,T}$、$\{z_{it}\}_{i=1,t=1}^{N,T}$ 不相关，满足 $E(\varepsilon_{it} \mid x_{11}\cdots,x_{NT},z_{11},\cdots,z_{NT}) = 0$，$Var(\varepsilon_{it} \mid x_{11}\cdots,x_{NT},z_{11},\cdots,z_{NT}) = \sigma_\varepsilon^2 < \infty$，$E(\|\varepsilon_{it}x_{it}'\|) < \infty$，$E(\|\varepsilon_{it}z_{it}'\|) < \infty$。

（2）$\{u_{it}\}_{i=1,t=1}^{N,T}$ 是独立同分布随机序列，边际密度函数 $f_t(u)(t = 1,\cdots,T)$ 是 Lipschitz 连续的且对支撑集上任意 u 均有 $0 < f_t(u) < \infty$；$\{u_{it}\}_{i=1,t=1}^{N,T}$ 与 $\{\varepsilon_{it}\}_{i=1,t=1}^{N,T}$ 无关，且有 $E(x_{it}\varepsilon_{it} \mid u_{it}) = 0(i = 1,2,\cdots,N; t = 1,2,\cdots,T)$；此外，$\Omega_{11}(u) = E(z_{it}z_{it}' \mid u_{it} = u)$ 存在且非奇异，$\Omega_{11}(u) = E(z_{it}z_{it}' \mid u_{it} = u)$，$\Omega_{12}(u) = E(z_{it}x_{it}' \mid u_{it} = u)$ 和 $\Omega_{22}(u) = E(x_{it}x_{it}' \mid u_{it} = u)$ 中每一个元素都是二阶连续可微的，$E(x_{it}x_{it}')$ 是非奇异的常数矩阵。

（3）$\{\alpha_i\}_{i=1}^N$ 是个体随机效应，满足独立同分布性质，并且 $\{\alpha_i\}_{i=1}^N$ 与 $\{\varepsilon_{it}\}_{i=1,t=1}^{N,T}$ 相互独立，$E(\alpha_i \mid x_{11}\cdots,x_{NT},z_{11},\cdots,z_{NT}) = 0$，$Var(\alpha_i \mid x_{11}\cdots,x_{NT},z_{11},\cdots,z_{NT}) = \sigma_\alpha^2 < \infty$，$E(\|\alpha_i x_{it}'\|) < \infty$，$E(\|\alpha_i z_{it}'\|) < \infty(i = 1,2,\cdots,N; t = 1,2,\cdots,T)$。

（4）实值函数 $\gamma_i(u)(i = 1,2,\cdots,q)$ 是二阶连续可导的有界函数，在 $u \in U$ 上满足一阶 Lipschitz 条件，即对于任意支撑集上的点 u 都有 $|\gamma_i(u)| < m_\gamma$，其中 m_γ 是一个正常数。

（5）存在 $r = max\{4,s\}$ 使得 $E\|X\|^r < \infty$，$E\|Z\|^r < \infty$，$E\|\alpha\|^r < \infty$，$E\|\varepsilon\|^r < \infty$，并且存在 $\xi < 2 - s^{-1}$ 使得 $(NT)^{2\xi-1}h \to \infty$。

2. 关于模型中常量的假设。

（1）空间权重矩阵 $W_1 = (w_{ij}^1)$、$W_2 = (w_{ij}^2)(i,j = 1,2,\cdots,N)$ 是外生给定的矩阵，所有元素非随机，并且 $w_{ij}^r = O_p(1/l_N)$，$\lim_{N\to\infty} l_N/N = 0$，作为一种标准性约定，所有对角线元素 $w_{ii}^r = 0$，即空间单元自身与自身不存在空间相关性；W_r 和 $I_N - \rho W_r$ 的绝对行和与绝对列和一致有界（$r = 1,2$），即：对任意矩阵 $A = (a_{ij})(i,j = 1,2,\cdots,N)$ 存在非负常数 m_a，使得 $\sum_{i=1}^N |a_{ij}| \leqslant m_a$ 和 $\sum_{j=1}^N |a_{ij}| \leqslant m_a$ 成立。

（2）空间相关系数 $|\rho| < 1$，空间误差系数 $|\lambda| < 1$，且 $I_N - \rho W_1$、$I_N - \lambda W_2$ 是非奇异矩阵，对任意的 $\rho,\lambda \in \Theta$ 可逆，Θ 是紧参数空间，$(I_N - \rho W_1)^{-1}$、$(I_N - \lambda W_2)^{-1}$ 同样满足绝对行和与绝对列和一致有界。

3. 关于核函数的假设。

（1）$k(\cdot)$ 的支撑集是有界闭集，在其支撑集上是连续的非负偶函数。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，$\mu_l = \nu_l = 0$，并且 $\mu_0 = 1$，$\mu_2 \neq 0$。

（2）关于窗宽的假设：在 $N \to \infty$、$T \to \infty$ 及 $h \to 0$ 时，$NTh \to \infty$。

4. 关于模型参数估计唯一可识别条价的假设。

参数 θ 存在唯一真实值 $\theta_0 \in \Theta$，使得原模型成立。

5. 关于模型中参数估计的渐近正态性假设条件。

(1) 对于截面似然函数 $L(\theta)$,满足 $\Sigma_{\theta_0} = -\lim\limits_{N,T\to\infty} E\left(\frac{1}{NT}\frac{\partial^2 L(\theta)}{\partial\theta\partial\theta'}\bigg|_{\theta_0}\right)$ 存在且非奇异。

(2) $\phi_{H_0H_0} - \phi'_{XH_0}\phi_{XX}^{-1}\phi_{XH_0} > 0$,其中:$\Lambda = (I_{NT}-L)'B_2'\Sigma_{\mu_0}^{-1}B_2(I_{NT}-L)$,$\phi_{H_0H_0} = \lim\limits_{N,T\to\infty}\frac{1}{NT}H_0'\Lambda H_0$,$\phi_{XH_0} = \lim\limits_{N,T\to\infty}\frac{1}{NT}X'\Lambda H_0$,$\phi_{XX} = \lim\limits_{N,T\to\infty}\frac{1}{NT}X'\Lambda X$。

(3) $\lim\limits_{N,T\to\infty}\frac{1}{NT}\left\{\mathrm{tr}(G_{20}^2) + \mathrm{tr}(G'_{20}G_{20}) - \frac{2}{N}\mathrm{tr}^2(G'_{20}J_{NT}) - \frac{2}{N(T-1)}\mathrm{tr}^2(G'_{20}Q_{NT})\right\} > 0$,

$\lim\limits_{N,T\to\infty}\frac{1}{NT}\left\{\mathrm{tr}(G_{10}^2) + \mathrm{tr}(G'_{10}G_{10}) - \frac{2}{N}\mathrm{tr}^2(J_{NT}G'_{10}) - \frac{2}{N(T-1)}\mathrm{tr}^2(G'_{10}Q_{NT})\right\} > 0$,

$\lim\limits_{N,T\to\infty}\frac{1}{NT}\left\{\mathrm{tr}(G'_{10}G'_{20}) + \mathrm{tr}(G'_{10}G_{20}) - \frac{2}{N}\mathrm{tr}(G'_{20}J_{NT})\mathrm{tr}(G'_{10}J_{NT}) - \frac{2}{N(T-1)}\mathrm{tr}(G'_{20}Q_{NT})\mathrm{tr}(G'_{10}Q_{NT})\right\} > 0$。

假设 1 和假设 3 对应于 Fan 和 Huang(2005)[15] 变系数部分线性模型中的一些基本假设; 假设 2 是空间计量经济模型中常见的假设; 假设 4 是参数唯一性识别条件; 假设 5 是证明估计量的渐近正态分布的充分条件。总之,上述所有假设都是为了得出估计量的大样本性质而设定的。

(二) 主要结论

1. 相关引理。

在得出主要结论之前,本文首先给出以下重要引理。

引理 1: 在假设 1 ~3 满足时:

$$R(u) = \frac{1}{NT}S(u)'K(u)S(u) \xrightarrow{p} \frac{1}{T}\sum_{t=1}^{T}f_t(u)\begin{pmatrix}\Omega_{11}(u) & 0\\ 0 & \mu_2\Omega_{11}(u)\end{pmatrix}$$

$$T(u) = \frac{1}{NT}S(u)'K(u)Y^* \xrightarrow{p} \frac{1}{T}\sum_{t=1}^{T}f_t(u)\begin{pmatrix}\Omega_{11}(u)\gamma(u)\\ 0\end{pmatrix}$$

引理 2: 在假设 1 ~4 满足时: $L\mu = o_p(1)$;$LB^{-1}(\lambda)\mu = o_p(1)$;$(I_{NT}-L)M = o_p(1)$;$\frac{1}{NT}A'(I_{NT}-L)'B'(\lambda)\Sigma_\mu^{-1}B(\lambda)(I_{NT}-L)M = o_p(1)$,其中 $A = X、M、H、\mu$。$\frac{1}{NT}A'(I_{NT}-L)'B'(\lambda)\Sigma_\mu^{-1}B(\lambda)(I_{NT}-L)G_1B^{-1}(\lambda)\mu = o_p(1)$,其中 $A = X、M、H$。

引理 3: 设 A 是一个 $NT\times NT$ 方阵,$\mu = D\alpha+\varepsilon$,其中:$D = e_T\otimes I_N$,$\alpha \sim N(0,\sigma_\alpha^2 I_N)$,$\varepsilon \sim N(0,\sigma_\varepsilon^2 I_{NT})$,则二次型 $\mu'A\mu$ 的期望和方差如下:

$$E(\mu'A\mu) = \sigma_\alpha^2\mathrm{tr}(\tilde{A}) + \sigma_\varepsilon^2\mathrm{tr}(A)$$

$$Var(\mu'A\mu) = \sigma_\alpha^4[\mathrm{tr}(\tilde{A}^2) + \mathrm{tr}(\tilde{A}\tilde{A}')] + \sigma_\varepsilon^4[\mathrm{tr}(A^2) + \mathrm{tr}(AA')] + \sigma_\alpha^2\sigma_\varepsilon^2[\mathrm{tr}(D'A'AD) + \mathrm{tr}(D'AA'D) + 2\mathrm{tr}(D'A^2D)]$$

其中: $\tilde{A} = D'AD$。

引理 4: 设 A 是一个 N 阶方阵,则: ①$\mathrm{tr}(A^2) \leqslant \mathrm{tr}(AA')$;② 若 A 的 N 个特征根均为实数,其中存在 k 个为非 0,那么有 $\mathrm{tr}(A^2) > 0$,并且满足$\frac{(\mathrm{tr}A)^2}{\mathrm{tr}(A^2)} \leqslant k$。

2. 主要结论。

定理 1: 在假设 1 ~4 成立时,$\hat{\theta} \xrightarrow{p} \theta_0$。

定理 2: 在假设 1 ~5 成立时,参数估计 $\hat{\theta}$ 具有渐近正态性,即 $\sqrt{NT}(\hat{\theta}-\theta_0)\xrightarrow{L}N(0,\Sigma_{\theta_0}^{-1})$,其中 $\Sigma_{\theta_0}=-\lim\limits_{N,T\to\infty}E\left(\frac{1}{NT}\frac{\partial^2 L(\theta)}{\partial\theta\partial\theta'}\bigg|_{\theta_0}\right)$。

此时,由于 $L(\theta)$ 中非参数矩阵 M 未知,因而本文由 $\tilde{L}(\theta)$ 出发计算 Σ_{θ_0} 的渐近估计量 $\hat{\Sigma}_{\theta_0}$。

定理 3: 在假设 1 ~5 成立时有 $\hat{\Sigma}_{\theta_0}\xrightarrow{p}\Sigma_{\theta_0}$,其中 $\hat{\Sigma}_{\theta_0}=-\lim\limits_{N,T\to\infty}\frac{1}{NT}\frac{\partial^2\tilde{L}(\theta)}{\partial\theta\partial\theta'}|_{\theta=\theta_0}$ 且 $\hat{\Sigma}_{\theta_0}$ 非奇异。

定理 4: 在假设 1 ~5 成立时,非参数估计满足 $\hat{\gamma}(u)\xrightarrow{p}\gamma(u)$。

定理 5: 在假设 1 ~5 成立时,$\sqrt{NTh}(\hat{\gamma}(u)-\gamma(u)-\frac{h^2}{2}\ddot{\gamma}(u)\mu_2)\xrightarrow{L}N(0,\varphi^2(u))$,其中 $\varphi^2(u)=v_0\left[\frac{1}{T}\sum_{t=1}^{T}f_t(u)\right]^{-2}F(u)\Omega_{11}^{-1}(u)$,$F(u)=\frac{1}{NT}\sum_{n=1}^{T}\sum_{m=1}^{N}f_n(u)\tilde{B}_{(mn,\cdot)}\Sigma_{\mu_0}^{-1}\tilde{B}'_{(mn,\cdot)}$,$\tilde{B}_{(mn,\cdot)}$ 是 $B_2^{-1}(\lambda)$ 的第 mn 行。特别的,当 $NTh^5\xrightarrow{p}0$ 时,$\sqrt{NTh}(\hat{\gamma}(u)-\gamma(u))\xrightarrow{L}N(0,\varphi^2(u))$。

五、MonteCarlo 数值模拟

为了考察所构建估计量的有限样本性质,本文进行如下 Monte Carlo 模拟试验: 核函数选取常用的 Epanechnikov 函数; 对于参数部分,本文用 $m=300$ 次模拟的均值作为估计值,用样本标准差以及两种均方根误 $RMSE_1$ 和 $RMSE_2$ 衡量估计的优劣,$RMSE_1=\sqrt{\frac{1}{m}\sum_{i=1}^{m}(\hat{\theta}_i-\theta_0)^2}$,$RMSE_2=(\hat{\theta}_{0.5}-\theta_0)^2+\frac{(\hat{\theta}_{0.75}-\hat{\theta}_{0.25})}{1.35}$,其中 $\hat{\theta}_{0.25}$、$\hat{\theta}_{0.5}$ 和 $\hat{\theta}_{0.75}$ 分别是估计值的下四分位数、中位数和上四分位数; 对于非参数部分,窗宽的选择借鉴 Su(2012) 的经验法则进行选取,选取支撑集内部 $n=100$ 个检验点计算绝对均方根误 $RASE=\sqrt{\frac{1}{n}\sum_{i=1}^{n}\|\hat{\gamma}(u_i)-\gamma(u_i)\|^2}$,并以 300 次模拟所得 $RASE$ 的中位数和标准差来衡量估计的好坏。

(一) 数据生成过程

1. 检验模型。

令 $Y=\rho\tilde{W}_1Y+\gamma_1{}'(u)z_1+\gamma_2{}'(u)z_2+X\beta+(I_{NT}-\lambda\tilde{W}_2)^{-1}(\varepsilon+e_T\otimes\alpha)$,其中 $\rho=0.5$、$\lambda=0.3$,$\gamma_1(u)=\sin(2\pi u)+2u$,$\gamma_2(u)=1.5e^{-u^2}+\cos(2\pi u)$,$\beta=1$,$X$ 是一元随机变量、服从 $N(0,1)$,$Z=(z_1,z_2)$ 是二元正态随机变量,其各个分量独立同分布于 $N(0,1)$,个体随机效应 α_i 独立同分布于 $N(0,1)$,扰动项 ε_{it} 独立同分布于 $N(0,0.5)$,变系数函数的协变量 u 独立同分布于 $U[0,1]$,同时本文假设 $\tilde{W}_1=\tilde{W}_2$。

2. 空间矩阵选择。

为了比较不同空间权重矩阵对上述模型估计的影响,本文设定 $T=2$,分别选取 Rook($N=49$,64,81,100) 和 Case($M=2$,4; $R=20$,30,40) 空间权重矩阵共 10 种情形进行模拟。

(二) 数值模拟结果

在 Matlab 软件中进行 300 次模拟,结果如下所示:

表 1 是空间权重矩阵为 Rook 矩阵下的模拟结果,可以发现在时间长度 T 确定时,随着空间单元数 N 的增加,真实值与估计值的偏误越来越小,主要体现在参数估计的标准差和两种均方根误

越来越小。这表明参数估计值会随着样本容量的增大而收敛，这与本文所证明的大样本性质相一致。

表 1　　Rook 权重矩阵下参数估计的均值和方差

$N = 49$					$N = 81$				
参数	均值	标准差	$RMSE_1$	$RMSE_2$	参数	均值	标准差	$RMSE_1$	$RMSE_2$
ρ	0.5232	0.0770	0.0803	0.0752	ρ	0.5101	0.0617	0.0624	0.0610
λ	0.2214	0.1449	0.1646	0.1688	λ	0.2429	0.1143	0.1276	0.1038
β	1.0108	0.1077	0.1080	0.1022	β	0.9940	0.0792	0.0793	0.0822
σ_α^2	0.8373	0.2587	0.3053	0.2839	σ_α^2	0.8650	0.1759	0.2215	0.1816
σ_ε^2	0.5343	0.1200	0.1246	0.1198	σ_ε^2	0.5290	0.0899	0.0943	0.0876
$N = 64$					$N = 100$				
参数	均值	标准差	$RMSE_1$	$RMSE_2$	参数	均值	标准差	$RMSE_1$	$RMSE_2$
ρ	0.5187	0.0712	0.0735	0.0640	ρ	0.5121	0.0566	0.0578	0.0537
λ	0.2357	0.1389	0.1529	0.1426	λ	0.2558	0.1093	0.1177	0.1150
β	1.0032	0.0959	0.0958	0.0951	β	0.9969	0.0663	0.0662	0.0668
σ_α^2	0.8403	0.2000	0.2557	0.2393	σ_α^2	0.8846	0.1790	0.2128	0.2047
σ_ε^2	0.5403	0.0977	0.1056	0.1010	σ_ε^2	0.5245	0.0769	0.0805	0.0730

由 Rook 空间权重矩阵下非参数函数的估计情况可以发现随着样本量的增加，$RASE$ 的中位数和标准差趋于下降，未知函数的估计是收敛的，同样印证了非参数估计量的大样本性质。

表 2 是 Case 空间权重矩阵下的参数估计表。虽然参数估计值与真实值很接近，估计精度随着样本总量的增加而提高，但是空间结构复杂程度对空间相关系数 ρ 和 λ 的精度影响较大。可以发现：在样本容量相同时，参数 ρ 和 λ 的估计偏误会随着空间复杂程度 M 的增加而增加，例如：$R = 20$、$M = 4$ 时，ρ 和 λ 估计的标准差和均方根误显著大于 $R = 40$、$M = 2$ 时的情形。

表 2　　Case 权重矩阵下参数估计表

R	参数	$M = 2$				$M = 4$			
		均值	标准差	$RMSE_1$	$RMSE_2$	均值	标准差	$RMSE_1$	$RMSE_2$
20	ρ	0.5278	0.0479	0.0553	0.0507	0.5160	0.0474	0.0499	0.0475
	λ	0.2190	0.1009	0.1293	0.1110	0.2321	0.0937	0.1156	0.0867
	β	0.9911	0.1203	0.1204	0.1241	0.9981	0.0803	0.0802	0.0798
	σ_α^2	0.7850	0.2711	0.3456	0.3208	0.8752	0.1787	0.2177	0.2169
	σ_ε^2	0.5584	0.1398	0.1513	0.1442	0.5309	0.0896	0.0947	0.0891
30	ρ	0.5201	0.0388	0.0437	0.0382	0.5082	0.0416	0.0423	0.0372
	λ	0.2392	0.0835	0.1032	0.0871	0.2635	0.0817	0.0893	0.0777
	β	0.9949	0.1045	0.1044	0.0959	0.9940	0.0598	0.0600	0.0610
	σ_α^2	0.8673	0.2425	0.2761	0.2796	0.9257	0.1578	0.1742	0.1638
	σ_ε^2	0.5616	0.1133	0.1288	0.1182	0.5274	0.0755	0.0802	0.0779
40	ρ	0.5142	0.0346	0.0374	0.0383	0.5075	0.0364	0.0371	0.0318
	λ	0.2492	0.0693	0.0858	0.0673	0.2674	0.0725	0.0794	0.0745
	β	0.9958	0.0786	0.0786	0.0795	0.9987	0.0494	0.0494	0.0505
	σ_α^2	0.8761	0.1825	0.2203	0.2202	0.9482	0.1347	0.1441	0.1444
	σ_ε^2	0.5502	0.0978	0.1097	0.1027	0.530	0.0596	0.0666	0.0575

Case 矩阵下未知函数估计值的 $RASE$ 的中位数和标准差，与 Rook 权重相似，随着样本容量的增大，中位数和标准差逐渐下降，趋于收敛。

六、结论

本文构建的随机效应广义空间滞后半参数变系数面板模型有效解决了非参数估计中的“维数灾难”问题，全面刻画了空间滞后效应和空间误差效应，并且半参数变系数的设定拓展了现有参数

模型的形式，具有更强的适用性和解释力。所构建的截面似然估计量满足一定条件下的一致性和渐近正态性，并且 Monte Carlo 数值模拟表明估计量具有稳健的小样本性质: ①空间权重矩阵的选择对估计量的性质影响不大，即 Rook 矩阵和 Case 矩阵都能得到性质较优的估计值; ②Rook 矩阵下，参数和非参数的估计精度随着样本总量的增加而提高; ③Case 矩阵下，空间相关系数 ρ 和 λ 的估计偏差随着空间权重结构复杂程度 M 的提升而增大，而其余参数的估计精度则随着样本总量的增大而提高。

参考文献

[1] H HKelejian, I R Prucha. A Generalized Spatial Two-stage Least Squares Procedure for Estimating a Spatial Autoregressive Model with Autoregressive Disturbances [J]. Journal of Real Estate Finance and Economics, 1998, 17(1): 99 - 121.

[2] L F Lee. Asymptotic Distributions of Quasi-maximum Likelihood Estimators for Spatial Autoregressive Models [J]. Econometrica, 2004, 72(6): 1899 - 1925.

[3] J P Elhorst. Specification and Estimation of Spatial Panel Data Models [J]. International Regional Science Review, 2003, 26(3): 244 - 268.

[4] M Kapoor, H HKelejian, I R Prucha. Panel Data Models with Spatially Correlated Error Components [J]. Journal of Econometrics, 2007, 140(1): 97 - 130.

[5] L F Lee, J H Yu. Estimation of Spatial Autoregressive Panel Data Models with Fixed Effects [J]. Journal of Econometrics, 2010, 154(2): 165 - 185.

[6] F Moscone, E Tosetti. GMM Estimation of Spatial Panels with Fixed Effects and Unknown Heteroskedasticity [J]. Regional Science and Urban Economics, 2011, 41(5): 487 - 497.

[7] R Basile. Regional Economic Growth in Europe: A Semiparametric Spatial Dependence Approach [J]. Papers in Regional Science, 2008, 87(4): 527 - 544.

[8] L J Su, S N Jin. Profile Quasi-maximum Likelihood Estimation of Partially Linear Spatial Autoregressive Models [J]. Journal of Econometrics, 2010, 157(1): 18 - 33.

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

[10] 陈建宝，孙林. 随机效应空间滞后单指数面板模型 [J]. 统计研究，2015(1): 95 - 101.

[11] 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(4): 129 - 146.

[12] Y Q Zhang, D M Shen. Estimation of Semi-parametric Varying Coefficient Spatial Panel Data Models with Random-effects [J]. Journal of Statistical Planning and Inference, 2015(159): 64 - 80.

[13] Y Bai, J H Hu, J H You. Panel Data Partially Linear Varying-coefficient Model with both Spatially and Time-wise Correlated Errors [J]. StatisticaSinica, 2015, 25(1): 275 - 294.

[14] 陈建宝，孙林. 随机效应变系数空间自回归面板模型的估计 [J]. 统计研究，2017(5): 118 - 128.

[15] J Q Fan, T Huang. Profile Likelihood Inferences on Semiparametric Varying-coefficient Partially Linear Models [J]. Bernoulli, 2005, 11(6): 1031 - 1057.

作者简介

唐礼智，男，2002 年毕业于华东师范大学获人文地理学专业理学博士学位，现为厦门大学经济学院统计系副主任、教授、博士生导师，中国宏观经济管理教育学会副秘书长。研究方向为计量经济学理论与方法、政策量化评估分析。

刘玉，女，现为厦门大学经济学院统计系在读博士研究生。研究方向为空间计量经济学、经济统计学。

(责任编辑: 郭明英)