第 30 卷第 4 期 统计研究 Vol. 30, No. 4
2013 年 4 月 Statistical Research Apr. 2013

复杂环境下空间滞后模型的稳健 LM 检验*

钱争鸣 刘立虎

内容提要: 针对误差项非正态分布和不同空间布局的复杂环境,本文提出一种新的空间滞后模型的稳健 LM 检验统计量及其检验方法,对比现有 LM 检验方法,它不仅消除了空间误差效应对检验的影响,而且克服了传统检验方法会受误差项非正态分布以及不同空间布局差异的影响。同时还具有在小样本情况下表现依然良好的优点。最后通过蒙特卡罗模拟的结果支持上述论断。

关键词: 非正态分布; 空间布局; 模型设定偏误; 矩估计; 稳健检验

中图分类号: F222.3　　**文献标识码:** A　　**文章编号:** 1002 – 4565(2013) 04 – 0099 – 07

The Robust LM Test of Spatial Lag Model with Non-normal Distribution and Spatial Layout

Qian Zhengming & Liu Lihu

Abstract: This paper proposes a robust LM statistic for diagnosing spatial lag effects, which is not sensitive to spatial error effect, spatial layout and distribution misspecification, while the existing test is not. Meanwhile, it has good performance under small sample. Monte Carlo results prove this.

Key words: Non-normal Distribution; Spatial Layout; Model Misspecification; GMM; Robust Test

一、引言

空间计量经济学理论认为,空间依赖性来源于非解释变量(扰动项)间的空间相关性以及解释变量间的空间溢出效应;而空间异质性来源于不同的空间分布,空间依赖性和空间异质性共同形成空间复杂的环境。它们会影响空间模型的稳健性。因此如何构建有效的空间模型稳健性检验统计量,探索有效的检验方法,具有理论与实证方面的重要意义。

国外学者在空间计量模型的模型设定与检验方法的研究已经做了一些较好的工作。相关的研究文献主要有: Cliff 和 Ord(1972) 扩展了 Moran(1950) 的统计量,有效检验了是否存在空间自相关效应; Burridge(1980) 建议采用拉格朗日乘子检验(LM 检验) 来识别空间误差模型和空间滞后模型; Anselin(2001) 推导了空间自回归移动平均(SARMA) 模型的 LM 检验,该检验是一阶 SAR 和 SEM 模型检验的一般化; Anselin(1988) 在考虑空间误差效应的情况下,建立了新的 LM 检验,但是该统计量要用极大似然法估计空间误差模型的参数,而且不具有较好的检验水平和检验功效; Anselin et al(1996) 巧妙地应用 Bera 和 Yoon(1993) 的修正 LM 检验方法,避免了对空间误差模型的极大似然估计,但是该检验统计量只有在空间误差项参数较小的情况下才是有效; Saavedra(2003) 提出应用 Kelejian 和 Prucha(1999) 的矩估计方法估计出空间误差效应的参数,然后应用 Newey 和 West(1987) 的理论构造检验统计量,这种检验方法依赖于空间误差参数的矩估计量,在小样本情况下,检验结果会受到一定的影响。国内的研究主要有张进峰(2011) 以 Bera 和 Yoon(1993) 的理论为基础构建了稳健 LM 统计量,该检验避免了 Anselin et al(1988, 1996) 方法计算上的繁琐和范围受限。

但是上述检验都是在假设模型误差项服从正态

* 本文为第四届统计学年会推荐论文。

分布及一般空间布局①的条件下构建和论证的，具有一定的局限性和缺陷，特别是在误差项服从非正态分布以及不同空间布局的条件下，LM 检验统计量的稳健性是未知的。基于此，本文考虑存在空间误差效应，以及在不同空间布局和误差项服从非正态分布的复杂环境下，通过构造空间滞后效应的稳健 LM 检验统计量，提出一种新的空间滞后模型的稳健 LM 检验方法。它不受空间误差参数大小的影响，也不受误差项非正态分布以及空间布局差异的复杂环境影响，而且还具有适用范围广和计算简捷的优点。

本文其余部分的安排如下，第二部分对现有的设定检验方法做简要回顾和评述，第三部分提出稳健检验方法，第四部分为蒙特卡罗模拟及其结果，最后是研究结论。

二、检验方法回顾与评述

对现有检验方法的简要回顾和评述可以从 Anselin(1988) 提出的空间联合模型(SAC) 开始，该模型的形式如下：

$$y = \rho W_1 y + X\beta + u, u = \lambda W_2 u + \varepsilon \tag{1}$$

其中，y 为 $N \times 1$ 维被解释变量向量，X 为 $N \times K$ 维解释变量矩阵，β 为 $K \times 1$ 维回归系数向量，W_1 和 W_2 是 $N \times N$ 维的空间权重矩阵，ρ 和 λ 分别为空间滞后效应系数和空间误差效应系数，u 为 $N \times 1$ 维误差项向量，且 $\varepsilon \sim N(0, \sigma_\varepsilon^2 I)$。

为便于表述，记 $A = I - \rho W_1$，$B = I - \lambda W_2$，$C = BW_1 B^{-1}$，$G_B = W_2 B^{-1}$，L 表示对数似然函数，tr 表示矩阵的迹运算。则上述空间联合模型的对数极大似然函数可表示为：

$$L = -\frac{N}{2}\ln 2\pi - \frac{N}{2}\ln\sigma^2 + \ln|A| + \ln|B| - \frac{1}{2\sigma^2}\varepsilon'\varepsilon \tag{2}$$

其中，$|A|$ 和 $|B|$ 为雅克比行列式，$\varepsilon = BAy - BX\beta$。

为了检验是否存在空间滞后效应，提出原假设为 $H_\rho^0: \rho = 0$，备择假设为 $H_\rho^1: \rho \neq 0$。定义 $\theta = [\beta', \sigma^2, \rho, \lambda]'$，$\gamma = [\beta', \sigma^2]'$。在原假设条件下，式(2) 的 Rao Score 向量 $d = \partial L/\partial\theta$ 以及对应的 Fisher 信息矩阵 $J_{\theta\theta} = -E[\partial^2 L/\partial\theta\partial\theta']$ 可表示为：

$$d(\theta) = \begin{bmatrix} \partial L/\partial\beta \\ \partial L/\partial\sigma^2 \\ \partial L/\partial\lambda \\ \partial L/\partial\rho \end{bmatrix} = \begin{bmatrix} \frac{1}{\sigma^2}(BX)'\varepsilon \\ \frac{\varepsilon'\varepsilon}{2\sigma^4} - \frac{N}{2\sigma^2} \\ \frac{1}{\sigma^2}\varepsilon' G_B \varepsilon - trG_B \\ \frac{1}{\sigma^2}\varepsilon' B W_1 y \end{bmatrix} \tag{3}$$

$$J(\theta) = \begin{bmatrix} \frac{1}{\sigma^2}(BX)'BX & 0 & 0 & \frac{1}{\sigma^2}(BX)'BW_1X\beta \\ 0 & \frac{N}{2\sigma^4} & \frac{1}{\sigma^2}trG_B & 0 \\ 0 & \frac{1}{\sigma^2}trG_B & T_{BB} & T_{BC} \\ \frac{1}{\sigma^2}(BW_1X\beta)'BX & 0 & T_{CB} & \frac{1}{\sigma^2}tr(BW_1X\beta)'(BW_1X\beta) + T_{CC} \end{bmatrix} \tag{4}$$

其中，$T_{BB} = tr[(G_B' + G_B)G_B]$，$T_{CC} = tr[(C' + C)C]$，$T_{BC} = T_{CB} = tr[(G_B' + G_B)C]$。

式(4) 表明，忽略空间误差效应系数 λ 的存在而检验空间滞后模型，极有可能得到错误的推断。定义 $H_\lambda^0: \lambda = 0$ 和 $H_\lambda^1: \lambda = \lambda_0 + \delta/N$，其中 δ 为常数。在不考虑空间误差效应的影响，即默认 $H_\lambda^0: \lambda = 0$ 成立时，Burridge(1980) 在误差项服从正态分布的条件下推导出 LM 检验统计量：

$$LM_\rho = \frac{[e'W_1 y/\hat{\sigma}^2]^2}{T_{11} + (W_1 X\beta)'M(W_1 X\beta)/\hat{\sigma}^2} \tag{5}$$

其中，$e = y - X\hat{\beta}$ 为 OLS 估计的残差，$\hat{\sigma}^2 = e'e/N$，$M = I - X(X'X)X'$，$T_{11} = tr[(W_1' + W_1)W_1]$。如果原假设成立，则 $LM_\rho \xrightarrow{D} \chi^2(1)$。在有限样本下，式(3) 对误差项分布具有稳健性，对不同的空间布局不具有稳健性，而且受到空间滞后效应的影响。

注意到式(5) 的检验只适用于真实模型不存在空间误差效应的情况，当存在空间误差项时，其均值和方差都会发生变化，从而不再服从中心化的卡方分布，导致检验水平和检验功效发生严重扭曲。为了消除空间误差参数的影响，Anselin(1988) 在误差项服从正态分布条件下构造了 LM_ρ^A 检验统计量：

$$LM_\rho^A = \frac{[e'B'BW_1 y/\hat{\sigma}^2]^2}{H_\rho - H_{\eta\rho}\mathrm{var}(\hat{\eta})H_{\rho\eta}} \tag{6}$$

① 最常使用的空是权重矩阵为 Rook 规则或者 Queen 规则。

其中,e 为 $y = X\beta + (I - \lambda W_2)^{-1}\varepsilon$ 极大似然法估计的残差向量,$\hat{\sigma}^2 = e'e/N$,$\eta = [\beta', \sigma^2, \lambda]'$,$B = I - \hat{\lambda} W_2$。$H_\rho = \frac{1}{\sigma^2} tr(BW_1X\beta)'(BW_1X\beta) + T_{CC}$,$H_{\eta\rho} = [\frac{1}{\sigma^2}(BX)'BW_1X\beta \quad 0 \quad T_{BC}]'$,$\text{var}(\hat{\eta})$ 为原假设下模型参数 η 的协方差矩阵的估计值。

虽然对式(6)的检验考虑了空间误差效应的影响,但是该方法仍然没有消除空间误差参数的影响,且该统计量需要用极大似然法估计空间滞后模型。如果样本量较大,那么对其估计将变得十分困难甚至难以实现。Anselin 等(1996)认为,如果允许空间误差参数在零附近的一个局部范围内存在,通过一定的变换就可以既避免式(5)计算复杂的问题又在一定范围内控制空间误差效应的影响,即默认 $H_\lambda^1: \lambda = \delta/\sqrt{n}$,但是不对 λ 进行估计,则在原假设 H_ρ^0 成立条件下,统计量形式如下:

$$LM_\rho^L = \frac{[e'W_1y/\hat{\sigma}^2 - T_{21}(T_{22})^{-1}e'W_2e/\hat{\sigma}^2]^2}{\hat{J}_{\rho\gamma} - (T_{21})^2(T_{22})^{-1}} \tag{7}$$

其中,e 为 $y = X\beta + \varepsilon$ 最小二乘法估计的残差向量,$\hat{\sigma}^2 = e'e/N$。$\hat{J}_{\rho\gamma} = \frac{1}{\hat{\sigma}^2}(W_1X\hat{\beta})'M(W_1X\hat{\beta}) + T_{11}$,$T_{22} = tr[(W_2 + W_2')W_2]$,$T_{21} = tr[(W_2 + W_2')W_1]$。该检验方法的优点是只需估计 $\lambda = \rho = 0$ 时的模型,从而使计算简便。但是只有当空间误差参数较小时,这种检验方法才是有效的。张进峰等(2011)应用 Bera 和 Yoon(1993)的理论,默认 $H_\lambda^1: \lambda = \lambda_0 + \delta/\sqrt{n}$ 成立,即通过 GMM 方法①事先获得 λ 的一致估计量 λ_0,允许 λ 在 λ_0 的局部范围内波动获得更加稳健的统计量,其形式如下:

$$LM_\rho^Z = \frac{[e'BW_1y/\tilde{\sigma}^2 - \tilde{T}_{BC}(J_{\lambda\gamma})^{-1}(e'G_Be/\tilde{\sigma}^2 - trG_B)]^2}{\tilde{J}_{\rho\gamma} - (\tilde{T}_{BC})^2(J_{\lambda\gamma})^{-1}} \tag{8}$$

其中,$\tilde{\gamma} = [\beta', \tilde{\sigma}^2]'$ 为 $\lambda = \lambda_0$ 和 $\rho = 0$ 时模型的极大似然估计量,$e = (I - \lambda_0 W_2)(y - X\tilde{\beta})$,$\tilde{\sigma}^2 = e'e/N$,$\tilde{J}_{\rho\gamma} = \frac{1}{\tilde{\sigma}^2}(B(\lambda_0)W_1X\tilde{\beta})'M_{BX}(B(\lambda_0)W_1X\tilde{\beta}) + T_{CC}(\lambda_0)$,$M_{BX} = I - (BX)[(BX)'(BX)]^{-1}(BX)'$,$\tilde{J}_{\lambda\gamma} = T_{BB}(\lambda_0) - \frac{2}{N}tr^2[G_B(\lambda_0)]$。

在 $\lambda_0 = 0$ 的假定下,式(8)简化为式(7)。如果 $\lambda_0 \neq 0$,式(7)会因为没有充分考虑空间误差效应的影响而得到错误的结论。显然,上述统计量都是在误差项服从正态分布的假设下获得的,如果误差项不服从正态分布或者空间权重矩阵 W 不再满足 Rook 规则(Queen 规则),那么上述统计量的稳健性就可能存在问题。

三、考虑不同空间布局和误差项非正态分布的稳健 LM 检验

在提出空间滞后模型的稳健 LM 检验之前,先对模型做如下假设:

假设 A1: 扰动项 $\{\varepsilon_i\} \sim IID(0, \sigma_\varepsilon^2)$,其过度(超额)峰度为 κ_ε。$\exists \alpha > 0$,使得 $E|\varepsilon_i|^{4+\alpha}$ 存在。

假设 A2: 对于所有的 i、j,空间权重矩阵 $W_{N\times N}$ 中的元素 w_{ij} 为 $O(h_N^{-1})$,其中序列 $\{h_N\}$ 是有界的或者发散的,且满足当 $N\to\infty$ 时,$h_N/N\to 0$,则矩阵 W 的行和与列和是一致有界的。例如矩阵行标准化后,对于所有的 i,对角元素 $w_{ii} = 0$ 并且 $\sum_j w_{ij} = 1$。

假设 A3: 对于所有的 N,矩阵 $\bar{X}_{N\times K}$ 中的元素是一致有界的,$\lim_{N\to\infty}\frac{1}{N}\bar{X}'\bar{X}$ 存在且非奇异。则 $\bar{X}(\bar{X}'\bar{X})^{-1}\bar{X}'$ 和 $I - \bar{X}(\bar{X}'\bar{X})^{-1}\bar{X}'$ 的行和与列和都是一致有界的。

假设 A4: 矩阵 W 与 $(I - \lambda_0 W)^{-1}$ 的范数是有界的②,则在 λ_0 的邻域内矩阵 $(I - \lambda W)^{-1}$ 是一致有界的。

上述假设设计的根据是: 假设 A1 为 Kelejian 和 Prucha(2001)关于线性二次型形式的中心极限定理的假设条件之一。假设 A2 为 Lee(2004a)为识别不同类型的空间依赖性所做的假设条件之一。特别是当每个空间单元有固定数量的邻单元,如果采用 Rook 规则或者 Queen 规则矩阵时,这种空间依赖性类型的 h_N 是有界的; 而当每个空间单元的邻单元数量随着 N 的增大而趋于无穷时,如果采用 Group 规则,h_N 是发散的。为了将空间依赖性限制在合理的范围,假设 A2 要求当 $N\to\infty$ 时,$h_N/N\to 0$。假设 A3 是 Lee(2004a)关于线性二次型形式的中心极限定理的假设条件之一。而假设 A4 表明,BX 满足假设 A3,则 $M_{BX} = I - (BX)[(BX)'(BX)]^{-1}(BX)'$ 亦满足假设 A3。

定义 $H_\lambda^0: \lambda = \lambda^*$ 和 $H_\lambda^1: \lambda = \lambda^* + \delta/\sqrt{N}$,其中 λ^*

---

① GMM 估计方法为 Kelejian 和 Prucha(1999)提出的方法。

② 即假定矩阵的行和与列和是有界的。

为已知常数，δ 为未知常数。Bera 和 Yoon(1993)表明，如果在真实的数据生成过程中，λ 为 H_{λ}^{1} 时被错误设定为 H_{λ}^{0}，那么在原假设 H_{ρ}^{0} 的条件下 $\frac{1}{\sqrt{N}}d_{\rho}(\tilde{\theta}) \xrightarrow{D} N(J_{\rho\lambda\gamma}\delta, J_{\rho\gamma})$，$\frac{1}{\sqrt{N}}d_{\lambda}(\tilde{\theta}) \xrightarrow{D} N(J_{\lambda\gamma}\delta, J_{\lambda\gamma})$。由于 δ 为未知常数且 $\delta=\sqrt{N}(\lambda-\lambda^{*})$，所以必须估计 λ。他们认为可以通过一定的变换消除未知常数 δ，得到 $\frac{1}{\sqrt{N}}d_{\rho}(\tilde{\theta})-J_{\rho\lambda\gamma}J_{\lambda\gamma}^{-1}\frac{1}{\sqrt{N}}d_{\lambda}(\tilde{\theta}) \xrightarrow{D} N(0, J_{\rho\gamma}-J_{\rho\lambda\gamma}J_{\lambda\rho\gamma}^{-1})$，从而构造稳健的 LM 检验。

为便于表述，记 $S_1=\frac{1}{N-K}tr(M_{BX}C)$，$P=M_{BX}(C-S_1 I)$，$S_{12}=\sum_i p_{ii}^2$，$S_{13}=tr(PP'+PP)$，$\{p_{ii}\}$ 为矩阵 P 的对角线元素；$\tilde{B}=I-\lambda_0 W_2$，$S_2=\frac{1}{N-K}tr(M_{BX}G_B)$，$Q=M_{BX}(G_B-S_2 I)_{BX}$，$S_{22}=\sum_i q_{ii}^2$，$\{q_{ii}\}$ 为矩阵 Q 的对角线元素，$S_{23}=tr(QQ'+QQ)$；$S_{32}=\sum_i p_{ii}q_{ii}$，$S_{33}=tr(PQ'+PQ)$。定义假设检验 H_{ρ}^{0}: $\rho=0$; H_{ρ}^{1}: $\rho\neq 0$，我们提出构造如下的稳健 LM 检验统计量：

$$LM_{\rho}^{R}=\{[e'BW_1 y/\tilde{\sigma}^2-NS_1)-(\tilde{\kappa}_{\varepsilon}S_{32}+S_{33})(\tilde{\kappa}_{\varepsilon}S_{22}+S_{23})^{-1}(e'G_B e/\tilde{\sigma}^2-NS_2)]^2\}/[(\tilde{\kappa}_{\varepsilon}S_{12}+S_{13}+\tilde{S}_{14})-(\tilde{\kappa}_{\varepsilon}S_{32}+S_{33})^2(\tilde{\kappa}_{\varepsilon}S_{22}+S_{23})^{-1}] \quad (9)$$

其中，$\tilde{\gamma}=[\tilde{\beta}',\tilde{\sigma}^2]'$ 为 $\lambda=\lambda^{*}$ 和 $\rho=0$ 时 SAC 模型的极大似然估计量，$e=(I-\lambda^{*}W_2)(y-X\tilde{\beta})$，$\tilde{\sigma}^2=e'e/N$，$\tilde{\kappa}_{\varepsilon}$ 为 e 的样本过度(超额)峰度。

定理 T1：如果式(9)中的 $W_i(i=1,2)$、$\{\varepsilon_i\}$、B 和 X 满足假设 A1～A4，则在原假设 H_{λ}^{1} 成立的条件下：①LM_{ρ}^{R} 渐进服从 $\chi^2(1)$；②当 $\kappa_{\varepsilon}=0$ 时，LM_{ρ}^{R} 渐进等价于 LM_{ρ}^{Z}。

因 $E(e'G_B e)=\sigma^2 tr(M_{BX}G_B)\neq 0$ 且 $E(e'W_1 y)=\sigma^2 tr(M_{BX}C)\neq 0$，所以 LM_{ρ}^{Z} 分子中的 $e'G_B e$ 项和 $e'W_1 y$ 项是非中心化的，导致 LM_{ρ}^{Z} 服从非中心化的卡方分布。考虑对 LM_{ρ}^{Z} 的均值和方差进行修正，即引入 $e'G_B e-\sigma^2 tr(M_{BX}G_B)$ 或者 $e'G_B e-\frac{e'e}{N-K}tr(M_{BX}G_B)=\varepsilon'Q\varepsilon$、$e'BW_1 y-\sigma^2 tr(M_{BX}C)$ 或者 $e'BW_1 y-\frac{e'e}{N-K}tr(M_{BX}C)$。只要 ε 存在四阶矩，可得到二者的均值和方差，将 σ^2 用其极大似然估计量代替可得 LM_{ρ}^{R}，因此 LM_{ρ}^{R} 不依赖于扰动项服从正态分布的假设。

尽管 LM_{ρ}^{Z} 检验统计量是在误差项服从正态分布的假设下推导的，但是由定理 T1 可知其与 LM_{ρ}^{R} 检验统计量渐进等价。也就是说，当样本量趋于无穷时，即使存在分布设定错误，两个统计量都是稳健的，但是统计量 LM_{ρ}^{Z} 和 LM_{ρ}^{R} 的有限样本性质不同，这两个统计量的主要差异在于修正 $e'BW_1 y/\sigma^2$ 的均值以及修正空间误差效应与滞后效应的交互影响。随着样本量的增加，这个修正量在一定的空间布局条件下可以忽略不计，但是在其他类型的空间布局条件下却是无法忽略的。LM_{ρ}^{Z} 和 LM_{ρ}^{R} 的关系可表示为：

$$\sqrt{LM_{\rho}^{R}}=\sqrt{\frac{T_0}{S_0}}\sqrt{LM_{\rho}^{Z}}-\frac{NS_1}{S_0^{1/2}}-[S_{03}S_{02}^{-1}(e'G_B e/\tilde{\sigma}^2-NS_2)-\tilde{T}_{BX}(J_{\lambda\gamma})^{-1}(e'G_B e/\tilde{\sigma}^2-trG_B)]/S_0^{1/2} \quad (10)$$

其中，$S_{01}\triangleq\tilde{\kappa}_{\varepsilon}S_{12}+S_{13}+\tilde{S}_{14}$，$S_{02}\triangleq\tilde{\kappa}_{\varepsilon}S_{22}+S_{23}$，$S_{03}\triangleq\tilde{\kappa}_{\varepsilon}S_{32}+S_{33}$，$S_0\triangleq S_{01}-S_{03}^2 S_{02}^{-1}$，$T_0\triangleq\tilde{J}_{\rho\gamma}-(\tilde{T}_{BC})^2(J_{\lambda\gamma})^{-1}$。$LM_{\rho}^{R}\sim LM_{\rho}^{Z}$，但是，$LM_{\rho}^{Z}$ 的均值修正项是否可以忽略，取决于式(10)中的第三项的比例大小。

四、蒙特卡罗模拟

本文通过设计一系列蒙特卡罗模拟，来验证这一新的空间滞后模型的稳健 LM 检验方法。为了获得检验统计量 LM_{ρ}^{R} 的有限样本性质，其中扰动项分布以及空间布局类型的设计可参考 Baltagi 和 Yang(2010)的蒙特卡罗试验。

(一)空间布局与误差项分布

考虑三种形式的空间权重矩阵：①Rook 规则；②Queen 规则；③Group 规则，Group 的数量 $G=N^{\psi}(0<\psi<1)$。在 Rook 或者 Queen 规则下的权重矩阵，每个空间单元的邻单元个数是固定的(Rook 规则为 2～4 个，Queen 规则为 3～8 个)，且不会随着 N 的增加而改变。在 Group 规则下的权重矩阵，每个空间单元的相邻单元个数会随着 N 的增加以一个比较小的速度增加，Group 的空间相邻单元之间也会发生变化。

考虑三种形式的误差项分布：①标准正态分布；②混合正态分布；③对数正态分布。后两种分布可标准化为均值为 0，方差为 1 的分布。与正态分布相比，混合正态分布呈对称、尖峰态分布；对数正态

分布呈偏态、尖峰态分布。标准化的混合正态分布生成方法如下：

$$\nu_i = [(1-\xi_i)Z_i + \xi_i \sigma Z_i]/(1-p+p\sigma^2)^{1/2} \quad (11)$$

其中，ξ 是概率为 p 的伯努利随机变量，Z 是依赖于 ξ 的标准正态随机变量，参数 p 也可作为两种正态分布混合的比例。在模拟中，选择 $p=0.05$，表示95%的随机变量来自标准正态分布，5%的随机变量来自标准差为 σ 的正态总体；选择 $\sigma=10$ 模拟误差项标准差较大的情况。标准化的对数正态分布生成方法如下：

$$\nu_i = [\exp(Z_i) - \exp(0.5)]/[\exp(2) - \exp(1)]^{1/2} \quad (12)$$

（二）检验的Size和Power分析

蒙特卡罗模拟的数据生成过程如下：

$$y = \rho W_1 y + X_1\beta_1 + X_2\beta_2 + X_3\beta_3 + \varepsilon$$
$$\varepsilon = \lambda W_2 \varepsilon + \nu \quad (13)$$

其中，X_1 为常数项，X_2 和 X_3 服从 $U(0,1)$ 的均匀分布。参数 $(\beta_1,\beta_2,\beta_3)=(1,1,1)$，$\rho=0,0.2,0.4,0.6,0.8$，$\lambda=0,0.2,0.4,0.6,0.8$。考虑3种样本量：$n=50,100,200$。每组模拟次数为1000次。

为验证本文统计量 LM_ρ^R 检验效果的改进，比较误差项分布以及空间布局对检验的影响，选择张进峰(2011)统计量 LM_ρ^Z 作为参照对象，分别利用蒙特卡罗模拟对统计量进行Size检验和Power检验。

下面对蒙特卡罗模拟的结果进行对比分析：

1\. LM_ρ^Z 和 LM_ρ^R 检验均不受误差项分布的影响，但是相对于 LM_ρ^Z，LM_ρ^R 表现略加稳健。空间权重矩阵为常用的Rook规则，当误差项服从正态分布且样本量为 $N=50$ 时，LM_ρ^Z 和 LM_ρ^R 检验表现无明显差异，拒绝原假设 $H_\rho^0: \rho=0$ 的概率均落在置信度为95%的置信区间①内；当样本量 $N=100$ 时，虽然二者的拒绝概率均落在了置信区间内，但是 LM_ρ^Z 检验拒绝概率明显偏小于0.05，而 LM_ρ^R 检验的拒绝概率在0.05附近波动；当误差项服从混合正态分布时，在小样本 $N=50$ 下 LM_ρ^Z 检验小于置信区间下限而 LM_ρ^R 检验均落在了置信区间内，但是当样本量 $N=100$ 时，LM_ρ^Z 检验和 LM_ρ^R 检验与拒绝概率5%无差异；当误差项服从对数正态分布时，两种样本量下 LM_ρ^Z 检验和 LM_ρ^R 检验的拒绝概率值均落在了置信区间内。空间权重矩阵为Queen规则，无论在何种误差项分布下，当样本量 $N=50$ 和 $N=100$ 时，LM_ρ^Z 和 LM_ρ^R 检验无差异且均落在置信区间内。这说明如果误差项服从非正态分布，LM_λ^Z 检验的小样本性质较差，随着样本量的增加，LM_ρ^Z 检验效果不断改善并与5%无显著差异，而 LM_ρ^R 检验表现良好。这也为定理T1的讨论提供了依据，即在一般空间权重矩阵下，随着样本量的增加，LM_ρ^Z 和 LM_ρ^R 渐进等价。

2\. LM_ρ^Z 检验对空间布局敏感，而 LM_ρ^R 检验表现稳健。如前所述，LM_ρ^Z 的均值修正项是否可以忽略取决于式(10)中第三项比例的大小。Size检验结果显示，在误差项服从正态分布下，如果空间权重矩阵为Group规则，当 $\phi=0.3$ 即 $h_N=N^{0.7}$ 时，LM_ρ^Z 检验的拒绝概率随着空间误差系数值的增大而增大，即使样本量 N 增加到100，检验结果也没有多少改进，但是 LM_ρ^R 检验除了 $N=100,\lambda=0.8$，其余情况均落在置信区间内；$\phi=0.5$ 即 $h_N=N^{0.5}$ 的情形与 $\phi=0.3$ 即 $h_N=N^{0.7}$ 的情形类似；当 $\phi=0.7$ 即 $h_N=N^{0.3}$ 且样本量 $N=50$ 时，LM_ρ^Z 检验的拒绝概率均落在置信区间内，LM_ρ^R 检验仅在 λ 为0.8时落在置信区间外(0.58)，当样本量 $N=100$ 时，二者均落在了置信区间内。这说明本文提出的 LM_ρ^R 检验对空间布局也具有稳健性。

3\. 当误差项分布以及空间布局均不服从常规性假设时，LM_λ^R 检验水平优于 LM_λ^Z。例如，在误差项分布为混合正态分布、空间权重矩阵为Group规则($\phi=0.3$ 即 $h_N=N^{0.7}$)的情况下，当样本量 $N=50$ 时，LM_λ^Z 检验拒绝原假设的概率呈现逐渐增大的趋势，且 $\lambda=0.8$ 时已经达到0.064，而 LM_ρ^R 检验约为0.05；当样本量 $N=100$ 时，LM_ρ^Z 检验拒绝原假设的概率没有多少改变。误差项分布为对数正态分布、空间权重矩阵为Group规则($\phi=0.3$ 即 $h_N=N^{0.7}$)的情况与之类似，除了 $N=100,\lambda=0.8$ 时为0.059。在空间权重矩阵为Group规则($\phi=0.5$ 即 $h_N=N^{0.5}$ 和 $\phi=0.7$ 即 $h_n=n^{0.3}$)的情况下，两种样本量的 LM_ρ^Z 检验均随着空间误差系数值的增大而增大，当 $\lambda=0.8$ 时落在了置信区间外，而 LM_ρ^R 检验则均落在了置信区间内。但是在空间权重矩阵为Group规则($\phi=0.7$ 即 $h_n=n^{0.3}$)的情况下，LM_ρ^Z 和 LM_ρ^R 检验

① 模拟1000次，则在5%的置信水平下，1倍标准差置信区间为[0.043，0.057]。

表 1 统计量的 Power 检验(Rook 规则)

	样本量		N = 50					N = 100				
误差项分布	ρ	λ	0.0	0.2	0.4	0.6	0.8	0.0	0.2	0.4	0.6	0.8
标准正态分布	0.2	LM_ρ^Z	0.753	0.745	0.762	0.801	0.779	0.980	0.981	0.979	0.983	0.987
		LM_ρ^R	0.780	0.772	0.795	0.826	0.803	0.984	0.983	0.982	0.988	0.991
	0.4	LM_ρ^Z	0.999	0.999	1.000	0.997	0.998	1.000	1.000	1.000	1.000	1.000
		LM_ρ^R	0.999	0.999	1.000	0.997	0.998	1.000	1.000	1.000	1.000	1.000
	0.6	LM_ρ^Z	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		LM_ρ^R	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.8	LM_ρ^Z	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		LM_ρ^R	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
混合正态分布	0.2	LM_ρ^Z	0.773	0.769	0.796	0.804	0.816	0.954	0.964	0.963	0.957	0.974
		LM_ρ^R	0.803	0.793	0.814	0.826	0.834	0.954	0.964	0.963	0.957	0.974
	0.4	LM_ρ^Z	0.991	0.989	0.992	0.992	0.994	1.000	1.000	1.000	1.000	1.000
		LM_ρ^R	0.993	0.996	0.995	0.994	0.995	1.000	1.000	1.000	1.000	1.000
	0.6	LM_ρ^Z	1.000	1.000	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		LM_ρ^R	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.8	LM_ρ^Z	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		LM_ρ^R	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
对数正态分布	0.2	LM_ρ^Z	0.847	0.854	0.860	0.857	0.893	0.976	0.979	0.980	0.971	0.990
		LM_ρ^R	0.862	0.871	0.877	0.879	0.908	0.977	0.979	0.983	0.975	0.991
	0.4	LM_ρ^Z	0.993	0.989	0.995	0.992	0.994	1.000	1.000	1.000	1.000	1.000
		LM_ρ^R	0.994	0.991	0.997	0.996	0.994	1.000	1.000	1.000	1.000	1.000
	0.6	LM_ρ^Z	1.000	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		LM_ρ^R	1.000	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.8	LM_ρ^Z	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		LM_ρ^R	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

表 2 统计量的 Power 检验(Group 规则:$\phi = 0.3$)

	样本量		N = 50					N = 100				
误差项分布	ρ	λ	0.0	0.2	0.4	0.6	0.8	0.0	0.2	0.4	0.6	0.8
标准正态分布	0.2	LM_ρ^Z	0.286	0.282	0.263	0.301	0.341	0.331	0.342	0.331	0.337	0.353
		LM_ρ^R	0.266	0.270	0.257	0.286	0.270	0.329	0.338	0.322	0.329	0.324
	0.4	LM_ρ^Z	0.694	0.696	0.667	0.668	0.722	0.784	0.772	0.766	0.764	0.796
		LM_ρ^R	0.693	0.693	0.664	0.661	0.676	0.770	0.766	0.755	0.760	0.767
	0.6	LM_ρ^Z	0.891	0.914	0.905	0.911	0.913	0.965	0.952	0.956	0.959	0.957
		LM_ρ^R	0.883	0.902	0.894	0.904	0.879	0.958	0.945	0.953	0.952	0.952
	0.8	LM_ρ^Z	0.981	0.995	0.990	0.977	0.982	0.997	0.997	0.996	0.997	0.995
		LM_ρ^R	0.963	0.985	0.974	0.961	0.956	0.988	0.989	0.987	0.990	0.989
混合正态分布	0.2	LM_ρ^Z	0.366	0.351	0.333	0.353	0.442	0.370	0.332	0.340	0.379	0.385
		LM_ρ^R	0.345	0.353	0.324	0.332	0.357	0.349	0.334	0.324	0.369	0.331
	0.4	LM_ρ^Z	0.737	0.747	0.702	0.724	0.743	0.797	0.771	0.757	0.750	0.797
		LM_ρ^R	0.719	0.732	0.702	0.690	0.685	0.774	0.759	0.746	0.738	0.752
	0.6	LM_ρ^Z	0.912	0.904	0.906	0.908	0.905	0.951	0.950	0.954	0.940	0.953
		LM_ρ^R	0.905	0.892	0.891	0.893	0.851	0.936	0.934	0.936	0.921	0.921
	0.8	LM_ρ^Z	0.983	0.982	0.979	0.982	0.990	0.995	0.994	0.992	0.991	0.996
		LM_ρ^R	0.972	0.961	0.959	0.956	0.952	0.975	0.977	0.977	0.971	0.972
对数正态分布	0.2	LM_ρ^Z	0.349	0.341	0.339	0.325	0.352	0.460	0.453	0.393	0.441	0.459
		LM_ρ^R	0.311	0.306	0.307	0.292	0.260	0.460	0.445	0.406	0.431	0.444
	0.4	LM_ρ^Z	0.714	0.745	0.712	0.701	0.730	0.869	0.860	0.848	0.803	0.843
		LM_ρ^R	0.675	0.701	0.669	0.652	0.628	0.864	0.857	0.849	0.807	0.842
	0.6	LM_ρ^Z	0.908	0.908	0.928	0.907	0.890	0.975	0.972	0.964	0.977	0.977
		LM_ρ^R	0.862	0.870	0.866	0.842	0.792	0.975	0.971	0.965	0.976	0.969
	0.8	LM_ρ^Z	0.985	0.981	0.984	0.984	0.973	0.999	0.993	0.997	0.998	0.998
		LM_ρ^R	0.925	0.924	0.927	0.913	0.903	0.996	0.989	0.992	0.993	0.994

结果在 $\lambda<0.8$ 时无差异,均落在置信区间内。

4. 在各种不同的模拟情况下,LM_ρ^R 的检验功效均优于 LM_ρ^Z。Power 检验结果如表 1 和表 2 所示。在空间权重矩阵为 Rook 规则(表 1)情况下,样本量 $N=50$ 时 LM_ρ^R 的检验功效明显优于 LM_ρ^Z,样本量 $N=100$ 时二者无明显差异,但是 LM_ρ^R 的检验功效仍然略优于 LM_ρ^Z。误差项服从非正态分布的情况与之类似。但是当空间权重矩阵为 Group 规则(表 2)时,无论误差项是否服从正态分布,LM_ρ^Z 的检验功效都略优于 LM_ρ^R,这是由于 LM_ρ^R 修正空间布局影响时造成的功效损失。无论何种空间布局,随着 λ 的增大,二者的检验功效都逐渐增大。

五、研究结论

本文构建了空间滞后模型的稳健检验统计量 LM_ρ^R,证明了 LM_ρ^R 不受误差项分布以及空间布局的影响。蒙特卡罗模拟结果说明,虽然在有限样本条件下 LM_ρ^Z 不受误差项分布的影响,但是对空间布局具有敏感性,这表明 LM_ρ^R 优于 LM_ρ^Z。例如,当空间权重矩阵为 Rook 或者 Queen 规则时,LM_ρ^Z 和 LM_ρ^R 检验随着样本量的增加渐进等价;而当空间权重矩阵为 Group 规则时,LM_ρ^Z 在三种误差项分布下各个样本量的检验水平均不平稳,与空间误差系数成正比,而 LM_ρ^R 则全部落在了置信区间内。

参考文献

[1] Anselin, L., 1988, Lagrange multiplier test diagnostics for spatial dependence and spatial heterogeneity [J]. Geographical Analysis, 20, 1 - 17.

[2] Anselin, L., 2001, Rao's score test in spatial econometrics [J]. Journal of Statistical Planning and Inference, 97, 113 - 139.

[3] Anselin, L., A. K. Bera, R. Florax, and M. J. Yoon, 1996, Simple diagnostic tests for spatial dependence [J]. Regional Science and Urban Economics, 26, 77 - 104.

[4] Bera, A. K., and M. J. Yoon, 1993, Specification testing with locally misspecified alternatives [J]. Econometric Theory, 9, 649 - 658.

[5] Baltagi, Badi and Yang, Zhenlin, (2010), Standardized LM Tests for Spatial Error Dependence in Linear or Panel Regressions [W]. No 11 - 2010, Working Papers, Singapore Management University, School of Economics.

[6] Burridge, P., 1980, On the Cliff-Ord test for spatial correlation [J]. Journal of the Royal Statistical Society. Series B, 42, 107 - 108.

[7] Kelejian, H. H. and Robinson, D. P. (1995). Spatial correlation: a suggested alternative to the autoregressive models. In: New Directions in Spatial Econometrics [C]. Edited by L. Anselin and R. J. G. M. Florax. Berlin: Springer-Verlag.

[8] Kelejian H. H. and Prucha, I. R. (2001). On the asymptotic distribution of the Moran I test statistic with applications [J]. Journal of Econometrics 104, 219 - 257.

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

[10] Lee, L. F. (2004b). A supplement to' asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models' [W]. Working paper, Department of Economics, Ohio State University.

[11] Lee, Lung-Fei, (2002), Consistency and Efficiency of Least Squares Estimation for Mixed Regressive, Spatial Autoregressive Models [J]. Econometric Theory, 18, issue 02, p. 252 - 277.

[12] Saavedra, Luz A., (2003), Tests for spatial lag dependence based on method of moments estimation, Regional Science and Urban Economics, 33, issue 1, p. 27 - 58.

[13] 张进峰,方颖. 空间误差模型的稳健检验 [J]. 数量经济技术经济研究,2011,(1).

[14] 钱争鸣,艾舍莱福,郭鹏辉. 非线性时序模型 LM 检验的两类临界值检验统计功效比较 [J]. 数量经济技术经济研究,2006(1):99 - 106.

作者简介

钱争鸣,男,江苏泰兴人,厦门大学经济学院统计系教授、博士生导师。研究方向为统计学与数量经济学。

刘立虎,男,河北唐山人,厦门大学经济学院统计系研究生。研究方向为空间统计与空间计量经济方法及其应用。

(责任编辑:禾　边)