

一类变系数空间滞后的混合地理加权回归模型*

唐志鹏^{1†}, 吴颖^{2,3}, 熊世峰³, 黄寰⁴

(1 中国科学院地理科学与资源研究所 区域可持续发展分析与模拟院重点实验室, 北京 100101; 2 中国科学院大学数学科学学院, 北京 100049; 3 中国科学院数学与系统科学研究院, 北京 100190; 4 成都理工大学商学院, 成都 610059)

(2022年7月22日收稿; 2022年10月10日收修改稿)

Tang Z P, Wu Y, Xiong S F, et al. A mixed geographically weighted regression model with varying-coefficient spatial lag[J]. Journal of University of Chinese Academy of Sciences, 2024, 41(3):345-356. DOI:10.7523/j.ucas.2022.081.

摘要 为解决因变量空间滞后存在的局域性问题,对现有常系数空间滞后的混合地理加权回归模型作了具有更广泛形式的拓展,提出一类变系数空间滞后的混合地理加权回归(MGWR-VSLR)模型。MGWR-VSLR模型实现了空间相关性与空间异质性融合,涵盖了绝大多数地理加权回归的模型形式,基于重构参数化方法和似然比检验分别给出模型的系数估计方法与显著性检验以及选取变系数的判别检验。在蒙特卡罗模拟与实际应用中, MGWR-VSLR模型均表现出优异的因变量拟合与预测能力。MGWR-VSLR模型的提出为量化研究空间效应问题设定适宜的模型形式提供了支撑依据。

关键词 空间异质性;混合地理加权回归;显著性检验;变系数

中图分类号:O212.7 文献标志码:A DOI:10.7523/j.ucas.2022.081

A mixed geographically weighted regression model with varying-coefficient spatial lag

TANG Zhipeng¹, WU Ying^{2,3}, XIONG Shifeng³, HUANG Huan⁴

(1 CAS Key Laboratory of Regional Sustainable Development Modeling, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China; 2 School of Mathematical Sciences,

University of Chinese Academy of Sciences, Beijing 100049, China; 3 Academy of Mathematics

and Systems Science, Chinese Academy of Sciences, Beijing 100190, China;

4 Business School, Chengdu University of Technology, Chengdu 610059, China)

Abstract Spatial correlation and spatial heterogeneity are the theoretical basis of spatial econometrics. In order to solve the local problem of spatial lag of dependent variables, this study extended the existing mixed geographically weighted regression model with constant-coefficient spatial lag, and proposed a mixed geographically weighted regression model with varying-coefficient spatial lag. The mixed geographically weighted regression model with varying-coefficient spatial lag

* 国家自然科学基金(42171177, 12171462)、成都市政府系统重大课题(B35360110202100097)和成都理工大学社科规划重大培育项目(YJ2021-XP001)资助

† 通信作者, E-mail: tangzp@igsrr.ac.cn

combines spatial correlation with spatial heterogeneity, and covers most of the model forms of geographically weighted regression. Based on the parameterization reconstruction method and likelihood ratio test, the coefficient estimation method, significance test of this model and the discriminant test of varying-coefficient are given respectively. Both in Monte Carlo simulation and practical application, the results show that the mixed geographically weighted regression model with varying-coefficient spatial lag renders itself well for the fitting and forecasting effect on dependent variable. The mixed geographically weighted regression model with varying-coefficient spatial lag provides a support for setting up a suitable model form for quantitative research on spatial effects.

Keywords spatial heterogeneity; mixed geographically weighted regression; significance test; varying-coefficient

空间计量经济学主要研究的是经济学中的空间效应问题,包括空间相关性和空间异质性。空间相关性表现为空间单元变量在一定程度上受到其他位置的空间单元影响,并且这种影响往往与距离存在一定相关性,而这种距离既可以是地理空间意义的距离,也可以是经济和社会的距离。空间异质性主要指地理空间上的区域缺乏均质性,从不同的空间观测点来观测会产生不同的观测结果。空间异质性反映了空间单元之间经济行为关系普遍存在的非均质性。应该说,空间相关性和空间异质性是空间计量经济学模型构建的理论基石。

在实际经济社会发展中,由于空间相关性的客观存在,空间自相关成为空间计量模型建立的依据。变量的空间自相关性表现为变量的数值不仅与所在空间位置点有关,还会受到其他空间点上变量取值的影响。自相关性的判断主要用于进行探索性空间数据分析,包括全局相关性和局部相关性,涉及常用的 2 种指数主要是 Moran's I 指数和 Geary G 指数^[1-2]。Fisher^[3]、Kelejian 和 Prucha^[4-7]、Lee^[8-12]对考虑空间相关性和空间异质性的截面线性回归模型进行研究,认为当空间系数在各空间点上不再恒定,此时采取变系数形式更合适。空间变系数的提出反映了空间局域性特征,但是增加了模型估计中待求解未知系数的个数,一定程度上增加了估计求解的难度。可以通过引入空间权重矩阵采用加权最小二乘法加以估计,而最合适的空间权重设置则是通过交叉验证所有空间点因变量的观测值与估计值最小平方和来实现。Brunsdon 等^[13-16]和 Fotheringham 等^[17]正是基于上述思路提出地理加权回归 (geographically weighted regression, GWR) 模型来揭示空间非均质性。GWR 模型强调空间变量系

数有别于传统的回归模型,它是变化的,并且与所在空间点和其他空间点的距离相关。这样,空间权重矩阵的引入也使得传统假设相互独立的空间样本点变得相互依赖起来。

近年来,GWR 模型在处理非均质空间数据得到了广泛应用。从空间变量来看,空间相关性是变量在空间分布特征最直接的体现,空间异质性则是空间结构的不稳定性表现。对于 GWR 而言,自变量并非全部是不稳定的空间结构,因此 Brunsdon 等^[16]提出混合地理加权回归 (mixed geographically weighted regression, MGWR),考虑了仅有部分自变量是稳定的空间结构。除自变量的空间异质性外,因变量本身还存在着空间滞后效应,至少从模型等式关系来看,自变量的空间效应也会影响到因变量。乔宁宁^[18]提出常系数空间滞后的混合地理加权回归 (mixed geographically weighted regression with constant-coefficient spatial lag regression, MGWR-CSLR) 模型。事实上,空间异质性结构不仅影响自变量,也影响着因变量, Brunsdon 等^[15]、Paze 等^[19]对于传统 GWR 进一步研究了因变量的空间滞后项回归,将滞后项的常系数替代为变系数,并且 Brunsdon 等^[15]认为各区域点因变量空间滞后的系数即使没有显著差异,也并不意味着常系数适用于每个地区。综上,从已有的 GWR 模型理论成果总结来看,不论是因变量还是自变量,所对应的变系数应该看作是常系数的一般性推广。常系数体现为变系数的一种特殊情形,譬如当表征变量的空间点由显著的空间集聚(扩散)分布变为随机分布,表征变量所对应每个空间点的变系数也就退化为对应所有点均值的常系数。由于模型变量需要考虑的空间特征越多,系数估计及检验也就越复杂,至今关于 GWR 考虑各种空间特征的复杂形式的理论与实

践研究均有待于解决。

为此,在前人研究基础上,本文试图对 GWR 模型做一个更具广泛意义的推广形式,将可能同时存在的因变量空间滞后性与自变量空间异质性 2 种效应融合并进行拓展研究,主要工作分为 3 方面:1)将变系数空间滞后回归与 MGWR 进行融合,放到一个全模型下进行考虑,因变量的空间滞后性与自变量空间异质性随着空间点变化能同时兼顾到并能充分展现,与乔宁宁^[18]提出的 MGWR-CSLR 模型相对应,将这一类新的模型称为变系数空间滞后的混合地理加权回归(mixed geographically weighted regression with varying-coefficient spatial Lag regression, MGWR-VSLR)模型。2)采用 Xiong^[20]提出的重构参数化方法对 MGWR-VSLR 模型进行系数估计,实现同步异带宽估计并给出模型与系数估计值的显著性检验,以及变系数选取的判别检验。3)基于蒙特卡罗模拟和实际应用检验了不同 GWR 形式的模型在处理空间相关性和空间异质性时所反映出的优良性质。基于本文的分析,有助于处理空间问题时选取适宜的 GWR 模型。

1 MGWR-VSLR 模型及其系数估计

1.1 MGWR-VSLR 模型

对于横截面空间数据来说,采用普通最小二乘回归(ordinary least square regression, OLSR)模型通常假定空间样本点相互独立,空间结构是稳定的,因此回归系数实质是样本均值回归的结果,该系数是不变的且适用于所有空间样本点,自变量对因变量的影响是全局性的。对所有空间点 i 而言,OLSR 模型自变量所对应的系数是不变的:

$$y_i = \beta_0 + \sum_{j=1}^p \alpha_j x_{ij} + \varepsilon_i, \quad i = 1, 2, \dots, n. \quad (1)$$

式中: y_i 和 $x_{i1}, x_{i2}, \dots, x_{ip}$ 都是对应空间点可观测的一般变量,其中 $\beta_0, \alpha_1, \alpha_2, \dots, \alpha_p$ 为回归系数, ε_i 为相应的随机误差项, $\varepsilon_i \sim N(0, \sigma^2)$ 。对于任意第 j ($1 \leq j \leq p$) 个自变量 X_j 中所有 n 个独立空间点观测值 $x_{1j}, x_{2j}, \dots, x_{nj}$, 所对应回归系数 α_j 是一样的,此时自变量也被看成是全局变量,对应系数即为常系数。当考虑到空间结构并不稳定,存在空间异质性时,则每个空间点的回归系数都是变化的,此时系数即为变系数,而自变量也被看成是局域变量。在 GWR 中每个空间样本点对应的回归系数是变化的,自变量对因变量的影响是局

域性的,Brundon 等^[15]引入空间点距离的高斯权重后,通过采用局部加权的最小二乘估计来求解变系数。

$$y_i = \beta_0(u_i, v_i) + \sum_{k=1}^q \beta_k(u_i, v_i) z_{ik} + \varepsilon_i, \quad i = 1, 2, \dots, n. \quad (2)$$

式(2)即为 GWR 模型,第 k 个自变量每个空间点 i 的系数均与所在空间位置的坐标 (u_i, v_i) 相关,不同空间点 i 所对应的回归系数也有所差异。随后,当考虑到仅有部分自变量在空间点上存在空间异质性,而其余变量空间异质性并不明显时,也即全局变量与局域变量同时存在于一个地理问题中,Brundson 等^[16]提出 MGWR 模型表示如下

$$y_i = \sum_{j=1}^p \alpha_j x_{ij} + \beta_0(u_i, v_i) + \sum_{k=1}^q \beta_k(u_i, v_i) z_{ik} + \varepsilon_i, \quad i = 1, 2, \dots, n. \quad (3)$$

式中:有 p 个自变量对应的系数 α_j ($j = 1, 2, \dots, p$) 被认为是常数,不会随空间位置变化而变化,其余剩下的自变量对应系数 β_k 以及截距项 β_0 则随着空间位置变化而变化,可以采用两步估计法^[21]求解。

尽管上述 MGWR 较为完整地反映了全局变量与局域变量空间结构的稳定性和不稳定性,却没有考虑到因变量存在的空间滞后性。乔宁宁^[18]在 MGWR 模型的基础上引入因变量的空间滞后性,建立了一种具有因变量空间滞后项的 MGWR 模型,由于假设因变量空间滞后项对不同空间点因变量的影响是一样的,也可以看成是 MGWR-CSLR 模型,如

$$y_i = \rho \sum_{j=1}^n w_{ij} y_j + \sum_{j=1}^p \alpha_j x_{ij} + \beta_0(u_i, v_i) + \sum_{k=1}^q \beta_k(u_i, v_i) z_{ik} + \varepsilon_i, \quad i = 1, \dots, n. \quad (4)$$

$$\text{或 } y_i = \rho \sum_{j=1}^n w_{ij} y_j + \alpha_0 + \sum_{j=1}^p \alpha_j x_{ij} + \sum_{k=1}^q \beta_k(u_i, v_i) z_{ik} + \varepsilon_i, \quad i = 1, 2, \dots, n. \quad (5)$$

在式(4)和式(5)中增加了因变量的空间滞后性, w_{ij} 为预先给定的 $n \times n$ 阶空间邻接矩阵元素,模型假设因变量空间滞后性对每个空间点因变量的影响均是相同的,因此,这里因变量的空间滞后项对应的系数 ρ 为待估计的常系数。

其中,式(4)的截距项 $\beta_0(u_i, v_i)$ 为一个与空间位置 (u_i, v_i) 有关的变参数,式(5)的截距项 α_0 为一个与空间位置无关的常参数。乔宁宁^[18]主

要针对式(5) 的求解过程做了详细介绍。在式(4) 和式(5) 中由于因变量的空间滞后项为常数 ρ , 表示因变量 y_i 所有空间点的滞后项对因变量影响均是一样的。Brunsdon^[15] 曾提出变系数的空间滞后回归 (varying-coefficient spatial lag regression, VSLR) 模型, 认为即使 $\rho(u_i, v_i)$ 在各个空间点差异不显著也并不意味着各个空间点适用于固定的常数 ρ 。从空间滞后的 MGWR 模型的形式来看, 自变量对应的变系数同样可能会影响因变量, 促使因变量表现出局域性特征, 将空间滞后项对应的常系数设为变系数, 更能反映因变量局部情况。因此, 与式(4) 和式(5) 的滞后项所对应的常系数 ρ 类似, 当滞后项前系数为变系数 $\rho(u_i, v_i)$, 分别得到:

$$y_i = \rho(u_i, v_i) \sum_{j=1}^n w_{ij} y_j + \sum_{j=1}^p \alpha_j x_{ij} + \beta_0(u_i, v_i) + \sum_{k=1}^q \beta_k(u_i, v_i) z_{ik} + \varepsilon_i, i = 1, 2, \cdots, n. \quad (6)$$

或

$$y_i = \rho(u_i, v_i) \sum_{j=1}^n w_{ij} y_j + \alpha_0 + \sum_{j=1}^p \alpha_j x_{ij} + \sum_{k=1}^q \beta_k(u_i, v_i) z_{ik} + \varepsilon_i, i = 1, 2, \cdots, n. \quad (7)$$

式(6) 和式(7) 的差异主要体现在截距项的区别, 在一般情况下, 式(6) 可以看作是式(7) 的更一般形式。因此为使本文研究的模型更具一般性, 后面的讨论均以式(6) 为例。MGWR-VSLR 模型的矩阵形式可写为

$$Y = \rho_n WY + X\alpha + M + \varepsilon. \quad (8)$$

其中:

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad X = [X_1, X_2, \cdots, X_p],$$
$$\rho_n = \begin{bmatrix} \rho(u_1, v_1) & 0 & \cdots & 0 \\ 0 & \rho(u_2, v_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \rho(u_n, v_n) \end{bmatrix},$$
$$X_i = \begin{bmatrix} x_{1i} \\ x_{2i} \\ \vdots \\ x_{ni} \end{bmatrix}, \alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_p \end{bmatrix},$$
$$\beta = \begin{bmatrix} \beta_0(u_1, v_1) & \beta_0(u_2, v_2) & \cdots & \beta_0(u_n, v_n) \\ \beta_1(u_1, v_1) & \beta_1(u_2, v_2) & \cdots & \beta_1(u_n, v_n) \\ \vdots & \vdots & \ddots & \vdots \\ \beta_q(u_1, v_1) & \beta_q(u_2, v_2) & \cdots & \beta_q(u_n, v_n) \end{bmatrix} =$$

$$\begin{bmatrix} \beta_0(u, v)^T \\ \beta_1(u, v)^T \\ \vdots \\ \beta_q(u, v)^T \end{bmatrix}, \quad Z = [I_n, Z_1, \cdots, Z_q],$$
$$Z_i = \begin{bmatrix} z_{1i} \\ z_{2i} \\ \vdots \\ z_{ni} \end{bmatrix}, \quad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix},$$
$$W = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1} & w_{n2} & \cdots & w_{nn} \end{bmatrix}.$$

设定 $\{d_1, \cdots, d_n\}$ 为 $Z\beta$ 的对角线元素, 则 $M = \text{diag}\{d_1, \cdots, d_n\} I_n, \rho_n WY$ 表示变系数空间滞后项。通过式(8), 可以看出把 MGWR-VSLR 模型作为全模型时: 1) 当不同变量对应的系数不再显著, 全模型分别退化为: $Y = \rho_n WY + \varepsilon$ 为变系数空间自回归 (varying-coefficient spatial auto-regression, VSAR) 模型; $Y = \rho_n WY + X\alpha + \varepsilon$ 为变系数空间滞后回归 VSLR 模型; $Y = \rho_n WY + M + \varepsilon$ 为变系数空间滞后的地理加权回归 (geographically weighted regression with varying-coefficient spatial lag regression, GWR-VSLR) 模型; 2) 而当 $\rho(u_i, v_i)$ 恒等于同一常数 ρ 时, 可以分别相应地得到传统常系数的空间自回归 (spatial auto-regression, SAR) 模型 $Y = \rho WY + \varepsilon$; 常系数的空间滞后回归 (spatial lag-regression, SLR) 模型 $Y = \rho WY + X\alpha + \varepsilon$; 常系数空间滞后的地理加权回归 (geographically weighted regression with constant-coefficient spatial lag regression, GWR-CSLR) 模型 $Y = \rho WY + M + \varepsilon$; 3) 而当 $\rho(u_i, v_i) \equiv 0 (i = 1, \cdots, n)$ 时, 此时 $Y = X\alpha + M + \varepsilon$ 为 MGWR 模型。因此, 从上述不同模型的形式总结可以看出, MGWR-VSLR 模型具有空间回归模型更一般的形式, 研究其变量对应系数的估计与统计量检验, 有助于更深刻地理解空间回归模型。

1.2 MGWR-VSLR 模型系数估计

在前人的研究中, 各类 GWR 都是利用空间局部加权最小二乘法来估计变系数部分, 然而应用这个方法难以实现一步估计含有变系数空间滞后项的模型系数。另一方面, 传统的 GWR 方法在估计系数时对所有系数使用同一个带宽, 这对不同空间点对应的变系数施加了同样的空间比

例,而在现实情况下,不同的变系数可能会在不同的空间尺度上产生影响。因此,本文采用 Xiong^[20]提出的重构参数化方法,将未知变系数关于空间位置的非参数回归转化为插值形式后,整个模型的极大似然估计便可以给出,估计过程可以对不同自变量项对应的系数以及因变量滞后项对应的系数同时分配不同的带宽并进行优化。另外,重构参数化方法将需要估计全部样本空间点的系数值转化为对少量节点对应的系数值进行估计,既减少了计算复杂度,又可以得到变系数关于空间位置的插值函数表达,有助于进一步分析变系数在空间上的变化。

参照 Xiong^[20]所给出的重构参数化方法求解过程,假设函数 $f(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^d$ 为待估函数,取线性插值算子 $\boldsymbol{\gamma}^T \mathbf{b}(\mathbf{x})$, 其中 $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_m)^T \in \mathbb{R}^m$ 为 m 个实数域节点 $\mathcal{A} = \{\mathbf{a}_1, \dots, \mathbf{a}_m\}$ 对应的真实值,即 $\boldsymbol{\gamma}^T \mathbf{b}(\mathbf{a}_i) = \gamma_i, i = 1, \dots, m$ 。因此 f 可以用 $\boldsymbol{\gamma}^T \mathbf{b}(\mathbf{x})$ 近似重构, 其中 $\boldsymbol{\gamma} = \mathbf{f}_{\mathcal{A}} = (f(\mathbf{a}_1), \dots, f(\mathbf{a}_m))^T$, 由此函数 f 的估计为

$$\tilde{f}(\mathbf{x}) = \tilde{\boldsymbol{\gamma}}^T \mathbf{b}(\mathbf{x}).$$

其中 $\tilde{\boldsymbol{\gamma}} = (\tilde{\gamma}_1, \dots, \tilde{\gamma}_m)^T$ 为下面目标函数的最优解:

$$\min_{\boldsymbol{\gamma} \in \mathbb{R}^m} \frac{1}{n} \sum_{i=1}^n [f(\mathbf{x}_i) - \boldsymbol{\gamma}^T \mathbf{b}(\mathbf{x}_i)]^2.$$

在本文中采用核插值,其中 $\mathbf{b}(\mathbf{x})$ 的计算方法如下

$$\mathbf{b}(\mathbf{x}) = P\mathbf{g}(\mathbf{x}) + Q\mathbf{r}_{\mathcal{A}}(\mathbf{x}). \quad (9)$$

其中:

$$\mathbf{P} = \mathbf{R}_{\mathcal{A}}^{-1} \mathbf{G}_{\mathcal{A}} (\mathbf{G}_{\mathcal{A}}^T \mathbf{R}_{\mathcal{A}}^{-1} \mathbf{G}_{\mathcal{A}})^{-1},$$

$$\mathbf{Q} = [\mathbf{I}_m - \mathbf{R}_{\mathcal{A}}^{-1} \mathbf{G}_{\mathcal{A}} (\mathbf{G}_{\mathcal{A}}^T \mathbf{R}_{\mathcal{A}}^{-1} \mathbf{G}_{\mathcal{A}})^{-1} \mathbf{G}_{\mathcal{A}}^T] \mathbf{R}_{\mathcal{A}}^{-1},$$

$$\mathbf{r}_{\mathcal{A}}(\mathbf{x}) = (R(\mathbf{x} - \mathbf{a}_1), \dots, R(\mathbf{x} - \mathbf{a}_m))^T,$$

$$\mathbf{R}_{\mathcal{A}} = (R(\mathbf{a}_i - \mathbf{a}_j))_{i=1, \dots, m, j=1, \dots, m},$$

$$\mathbf{g}(\mathbf{x}) = (1, \mathbf{x}^T)^T, \mathbf{G}_{\mathcal{A}} = (\mathbf{g}(\mathbf{a}_1), \dots, \mathbf{g}(\mathbf{a}_m))^T.$$

为与传统 GWR 方法一致,本文中取高斯核函数 $R(\mathbf{h}) = \exp(-\sum_{j=1}^2 h_j^2 / \theta^2)$, $\mathbf{h} = (h_1, h_2)^T, \theta > 0$ 为核参数,这里也称为带宽,表示局部影响效力的空间范围。

运用这一方法,给定 n 个空间点并从中选取 m 个节点,变系数向量 $\boldsymbol{\beta}_k = (\beta_k(u_1, v_1), \dots, \beta_k(u_n, v_n))^T (k=0, 1, \dots, q)$ 和空间滞后系数向量 $\boldsymbol{\rho} = (\rho(u_1, v_1), \dots, \rho(u_n, v_n))^T$ 可以近似写作

$$\tilde{\beta}_k(u_i, v_i) = \boldsymbol{\gamma}_k^T \mathbf{b}(u_i, v_i), k = 0, 1, \dots, q,$$

$$\tilde{\rho}(u_i, v_i) = \boldsymbol{\phi}^T \mathbf{b}(u_i, v_i). \quad (10)$$

式中, $\boldsymbol{\gamma}_k = (\gamma_{k1}, \dots, \gamma_{km})^T, \boldsymbol{\phi} = (\phi_1, \dots, \phi_m)^T$, 并且在式(10)中进一步记 $\mathbf{A}(\boldsymbol{\phi}) = \text{diag}\{\tilde{\rho}(u_1, v_1), \dots, \tilde{\rho}(u_n, v_n)\}, \tilde{\mathbf{Z}} = (\tilde{\mathbf{Z}}_0, \tilde{\mathbf{Z}}_1, \dots, \tilde{\mathbf{Z}}_q), \tilde{\mathbf{Z}}_0 = (\mathbf{b}(u_1, v_1), \dots, \mathbf{b}(u_n, v_n))^T, \tilde{\mathbf{Z}}_k = (\mathbf{b}(u_1, v_1)z_{1k}, \dots, \mathbf{b}(u_n, v_n)z_{nk})^T, k = 1, \dots, q$ 。令 $\boldsymbol{\gamma} = (\boldsymbol{\gamma}_0^T, \boldsymbol{\gamma}_1^T, \dots, \boldsymbol{\gamma}_q^T)^T$, $\mathbf{T}(\boldsymbol{\phi}) = \mathbf{I}_n - \mathbf{A}(\boldsymbol{\phi})\mathbf{W}, \mathbf{B} = (\mathbf{X}, \tilde{\mathbf{Z}})$ 以及 $\boldsymbol{\eta} = (\boldsymbol{\alpha}^T, \boldsymbol{\gamma}^T)^T$, 则 $\mathbf{B}\boldsymbol{\eta} = \mathbf{X}\boldsymbol{\alpha} + \mathbf{M}$, 式(8)可以简化为

$$\mathbf{Y} = \mathbf{T}(\boldsymbol{\phi})^{-1}(\mathbf{B}\boldsymbol{\eta} + \boldsymbol{\varepsilon}). \quad (11)$$

由于 $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_n)$, 式(11)的对数似然函数为

$$\ln L(\boldsymbol{\phi}, \boldsymbol{\eta}, \sigma^2) = -\frac{n}{2} \ln 2\pi\sigma^2 + \ln |\mathbf{T}(\boldsymbol{\phi})| -$$

$$\frac{1}{2\sigma^2} \{[\mathbf{T}(\boldsymbol{\phi})\mathbf{Y} - \mathbf{B}\boldsymbol{\eta}]^T [\mathbf{T}(\boldsymbol{\phi})\mathbf{Y} - \mathbf{B}\boldsymbol{\eta}]\}. \quad (12)$$

假设 $\boldsymbol{\phi}$ 已知,由式(12)容易得到 $\boldsymbol{\eta}$ 和 σ^2 的极大似然函数估计分别为

$$\begin{cases} \tilde{\boldsymbol{\eta}}(\boldsymbol{\phi}) = (\mathbf{B}^T \mathbf{B})^{-1} \mathbf{B}^T \mathbf{T}(\boldsymbol{\phi}) \mathbf{Y}, \\ \tilde{\sigma}^2(\boldsymbol{\phi}) = \frac{1}{n} (\mathbf{T}(\boldsymbol{\phi}) \mathbf{Y})^T \mathbf{S}^T \mathbf{S} \mathbf{T}(\boldsymbol{\phi}) \mathbf{Y}. \end{cases} \quad (13)$$

其中 $\mathbf{S} = \mathbf{I}_n - \mathbf{B}(\mathbf{B}^T \mathbf{B})^{-1} \mathbf{B}^T$, 则 $\boldsymbol{\phi}$ 的集中对数似然函数如下所示

$$\ln L(\boldsymbol{\phi}) = \frac{n}{2} [\ln(2\pi) + 1] - \frac{n}{2} \ln \tilde{\sigma}^2(\boldsymbol{\phi}) + \ln |\mathbf{T}(\boldsymbol{\phi})|. \quad (14)$$

只要进一步通过最小化式(14),即 $\min_{\boldsymbol{\phi} \in \mathbb{R}^m} \ln L(\boldsymbol{\phi})$, 即可得到 $\boldsymbol{\phi}$ 的估计值 $\tilde{\boldsymbol{\phi}} = (\tilde{\phi}_1, \dots, \tilde{\phi}_m)^T$ 。在已知 $\tilde{\boldsymbol{\phi}}$ 的基础上,再返回式(13)即可求解出 $\tilde{\boldsymbol{\eta}}$ 以及 $\mathbf{T}(\tilde{\boldsymbol{\phi}})$, 因此,可得到式(11)相应的估计等式

$$\tilde{\mathbf{Y}} = \mathbf{T}(\tilde{\boldsymbol{\phi}})^{-1} \mathbf{B} \tilde{\boldsymbol{\eta}}. \quad (15)$$

对于本文讨论的模型,将不同的带宽应用在不同系数对应的插值形式中可以更好地适应系数的光滑度差异。因此引入 kernel 参数向量 $\boldsymbol{\theta} = (\theta_1, \dots, \theta_q, \theta_a)^T$, 式(10)分别对应改写为

$$\tilde{\beta}_k(u_i, v_i) = \boldsymbol{\gamma}_k^T \mathbf{b}(u_i, v_i, \theta_k), k = 0, 1, \dots, q \text{ 和}$$

$$\tilde{\rho}(u_i, v_i) = \boldsymbol{\phi}^T \mathbf{b}(u_i, v_i, \theta_a). \quad (16)$$

对带宽 $\boldsymbol{\theta}$ 可以通过变异系数 (coefficient of variation, CV) 方法来估计,对于给定的 $\boldsymbol{\gamma}$ 和 $\boldsymbol{\phi}$, 寻找最小的 CV 值,如下式所示,其中 $\tilde{y}_{-i}(\boldsymbol{\theta})$ 是固定 $\boldsymbol{\theta}$ 已删掉第 i 组观测数据后的数据集得到的估计值

$$\begin{aligned} \tilde{\boldsymbol{\theta}} &= \underset{\boldsymbol{\theta} \in \mathbb{R}^{q+1}}{\text{argmin}} \text{CV}(\boldsymbol{\theta}) = \\ &\underset{\boldsymbol{\theta} \in \mathbb{R}^{q+1}}{\text{argmin}} \frac{1}{n} \sum_{i=1}^n (y_i - \tilde{y}_{-i}(\boldsymbol{\theta}))^2. \end{aligned} \quad (17)$$

2 MGWR-VSLR 模型的相关统计检验

2.1 模型整体的显著性检验

对于回归模型整体的显著性检验,主要是考察模型中所有参数存在的合理性,以及对整个模型的解释力度。若所有参数不能通过显著性检验,则模型本身失去存在的意义。因此,模型的显著性检验中存在零假设 $H_0: \tau = \mathbf{0}$, 其中 $\tau = (\rho, \alpha_1, \dots, \alpha_p, \beta_1^T, \dots, \beta_q^T)^T$, 备择假设 H_1 : 其中至少有一个 $\tau_i \neq 0$ 。在式(13)中得到相应参数的极大似然估计值,全模型的对数似然函数式(12)对应的极值问题转换为

$$\sup_{\tau, \sigma^2} \ln L(\tau, \sigma^2) = -\frac{n}{2} \left[\ln \frac{2\pi}{n} - \frac{2}{n} \ln |T(\tilde{\phi})| + \ln(T(\tilde{\phi})Y)^T S^T S T(\tilde{\phi})Y + 1 \right], \quad (18)$$

其中, $S = I_n - B(B^T B)^{-1} B^T$, $B = (X, \tilde{Z})$, $\tilde{\phi}$ 为 H_0 不成立时 ϕ 的极大似然估计量。

当 $\tau = \mathbf{0}$ 时,代入式(11)得到约简模型

$$Y = \beta_0 + \varepsilon,$$

其中 $\beta_0 = (\beta_0(u_1, v_1), \dots, \beta_0(u_n, v_n))^T$, 由式(10)

可知 $\tilde{\beta}_0(u_i, v_i) = \gamma_0^T b(u_i, v_i)$ 。记 $B_0 = (b(u_1, v_1), \dots, b(u_n, v_n))^T$, 故零假设成立时的对数似然函数为

$$\ln L(\beta_0, \sigma^2) = \ln L(\gamma_0, \sigma^2) = -\frac{n}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} (Y - B_0 \gamma_0)^T (Y - B_0 \gamma_0). \quad (19)$$

则

$$\sup_{\beta_0, \sigma^2} \ln L(\beta_0, \sigma^2) = -\frac{n}{2} \left[\ln \frac{2\pi}{n} + \ln Y^T S_0^T S_0 Y + 1 \right]. \quad (20)$$

其中 $S_0 = I_n - B_0(B_0^T B_0)^{-1} B_0^T$ 。根据式(18)和式(20),定义似然比检验统计量

$$\Lambda_T(B) = 2 \ln \frac{\sup_{\tau, \sigma^2} L(\tau, \sigma^2)}{\sup_{\beta_0, \sigma^2} L(\beta_0, \sigma^2)} = -n \left[\ln(T(\tilde{\phi})Y)^T S^T S T(\tilde{\phi})Y - \frac{2}{n} \ln |T(\tilde{\phi})| - \frac{2}{n} \ln |T(\tilde{\phi})| - \ln Y^T S_0^T S_0 Y \right].$$

当样本量 n 足够大时, $\Lambda_T(B) \rightarrow \chi^2(p + m(q + 1))$ 。对给定的显著性水平 α , 当 $\Lambda_T(B) > \chi_a^2(p + m(q + 1))$ 时,拒绝原假设 H_0 。

+ $m(q + 1)$) 时,拒绝原假设 H_0 。

2.2 模型系数的显著性检验

对于上述整体性检验,如果检验的结果是拒绝原假设,这意味着因变量 Y 的存在依赖于自变量 X 和 Z 以及受到因变量空间滞后项的影响,因此这里需要对引起 Y 变化的每个自变量逐一作显著性检验。对空间滞后项对应的变系数 ρ 进行显著性检验是对因变量 Y 滞后项是否存在独立影响 Y 的证明,需要作单独检验。类似地,对 α 和 β 进行显著性检验也是为了验证自变量 X 和 Z 对 Y 是否存在影响。

1) 空间滞后项对应变系数的显著性检验

考虑零假设 $H_0: \rho = \mathbf{0}$ 的检验问题,对于在式(10)中 ρ 的核插值估计,检验问题近似等价于

$$H_0^*: \phi_1 = \dots = \phi_m = 0 \text{ vs } H_1^*:$$

至少有一个 $\phi_i \neq 0, i = 1, \dots, m$ 。

假设 H_0^* 成立,即模型不存在空间滞后项,此时模型退化为混合地理加权回归模型,记 $\eta = (\alpha^T, \gamma^T)^T$, 相应的对数似然函数为

$$\ln L(\eta, \sigma^2) = -\frac{n}{2} \ln 2\pi\sigma^2 - \frac{[Y - B\eta]^T [Y - B\eta]}{2\sigma^2}. \quad (21)$$

则

$$\sup_{\eta, \sigma^2} \ln L(\eta, \sigma^2) = -\frac{n}{2} \left[\ln \frac{2\pi}{n} + \ln Y^T S^T S Y + 1 \right]. \quad (22)$$

其中 $S = I_n - B(B^T B)^{-1} B^T$ 。该检验所对应的全模型存在时的极大似然估计解为式(13),由式(18)及式(22),定义检验统计量

$$\Lambda_\rho(B) = 2 \ln \frac{\sup_{\phi, \eta, \sigma^2} L(\phi, \eta, \sigma^2)}{\sup_{\eta, \sigma^2} L(\eta, \sigma^2)} = -n \left[\ln(T(\tilde{\phi})Y)^T S^T S T(\tilde{\phi})Y - \frac{2}{n} \ln |T(\tilde{\phi})| - \ln Y^T S^T S Y \right].$$

其中 $\tilde{\phi}$ 为全模型存在时 ϕ 的极大似然估计值。当样本量 n 足够大时, $\Lambda_\rho(B) \rightarrow \chi^2(m)$ 。对给定的显著性水平 α , 当 $\Lambda_\rho(B) > \chi_a^2(m)$ 时,拒绝原假设 H_0 。

2) 自变量对应常系数的显著性检验

下面考虑对于不同自变量对应的系数进行逐个显著性检验。首先讨论常系数 $\alpha_i, i = 1, \dots, p$ 的显著性,记 $\alpha_{-i} = (\alpha_1, \dots, \alpha_{i-1}, \alpha_{i+1}, \dots, \alpha_p)^T$,

$B_{-i} = (X_{-i}, \bar{Z})$, 其中 X_{-i} 为 X 删去系数 α_i 对应的自变量列后的 $n \times (p-1)$ 维矩阵, $i=1, \dots, p$ 。考虑原假设 $H_{0i}: \alpha_i = 0$, 此时原假设下模型对应参数为 $\eta_{-i} = (\alpha_{-i}^T, \gamma^T)^T$, ϕ 及 σ_{-i}^2 , 其中 η_{-i} 和 σ_{-i}^2 的约束极大似然估计值与式(13)相似, 分别为

$$\begin{cases} \tilde{\eta}_{-i} = (B_{-i}^T B_{-i})^{-1} B_{-i}^T T(\tilde{\phi}_{-i}) Y, \\ \widetilde{\sigma_{-i}^2} = \frac{1}{n} (T(\tilde{\phi}_{-i}) Y)^T S_{-i}^T S_{-i} T(\tilde{\phi}_{-i}) Y. \end{cases} \quad (23)$$

其中 $S_{-i} = I_n - B_{-i} (B_{-i}^T B_{-i})^{-1} B_{-i}^T$, $\tilde{\phi}_{-i}$ 为 H_{0i} 成立时 ϕ 的极大似然估计值, 结合式(18), 此时定义似然比统计量为

$$\begin{aligned} \Lambda_{-i}(\mathbf{B}) = 2 \ln \frac{\sup_{\phi, \eta, \sigma^2} L(\phi, \eta, \sigma^2)}{\sup_{\phi, \eta_{-i}, \sigma_{-i}^2} L(\phi, \eta_{-i}, \sigma_{-i}^2)} = \\ -n \left[\ln(T(\tilde{\phi}) Y)^T S^T S T(\tilde{\phi}) Y - \frac{2}{n} \ln |T(\tilde{\phi})| - \right. \\ \left. \ln(T(\tilde{\phi}_{-i}) Y)^T S_{-i}^T S_{-i} T(\tilde{\phi}_{-i}) Y + \frac{2}{n} \ln |T(\tilde{\phi}_{-i})| \right]. \end{aligned}$$

$\tilde{\phi}$ 为全模型存在时 ϕ 的极大似然估计值。当样本量 n 足够大时, $\Lambda_{-i}(\mathbf{B}) \rightarrow \chi^2(1)$ 。对给定的显著性水平 α , 当 $\Lambda_{-i}(\mathbf{B}) > \chi_a^2(1)$ 时, 拒绝原假设 H_{0i} 。

3) 自变量对应变系数的显著性检验

与自变量对应常系数的显著性检验类似, 定义 $\beta_{-j} = (\beta_1, \dots, \beta_{j-1}, \beta_{j+1}, \dots, \beta_q)^T$, $B_{-j} = (X, \bar{Z}_{-j})$, 其中 $\bar{Z}_{-j} = (\bar{Z}_0, \bar{Z}_1, \dots, \bar{Z}_{j-1}, \bar{Z}_{j+1}, \dots, \bar{Z}_q)$ 为 \bar{Z} 删去系数 β_j 对应的自变量列 \bar{Z}_j 后的 $n \times mq$ 维矩阵, $j=1, \dots, q$ 。

令 $\gamma = (\gamma_0^T, \gamma_1^T, \dots, \gamma_q^T)^T$, $\gamma_{-j} = (\gamma_0^T, \gamma_1^T, \dots, \gamma_{j-1}^T, \gamma_{j+1}^T, \dots, \gamma_q^T)^T$, 由式(10)中 β 的核插值估计, 原假设 $H_{0j}: \beta_j = 0$ 近似等价于 $H_{0j}^*: \gamma_j = 0$ 。此时原假设下模型对应参数为 $\eta_{-j} = (\alpha^T, \gamma_{-j}^T)^T$, ϕ 及 σ_{-j}^2 , 其中 η_{-j} 和 σ_{-j}^2 的约束极大似然估计值分别为

$$\begin{cases} \tilde{\eta}_{-j} = (B_{-j}^T B_{-j})^{-1} B_{-j}^T T(\tilde{\phi}_{-j}) Y, \\ \widetilde{\sigma_{-j}^2} = \frac{1}{n} (T(\tilde{\phi}_{-j}) Y)^T S_{-j}^T S_{-j} T(\tilde{\phi}_{-j}) Y. \end{cases} \quad (24)$$

其中 $S_{-j} = I_n - B_{-j} (B_{-j}^T B_{-j})^{-1} B_{-j}^T$, $\tilde{\phi}_{-j}$ 为 H_{0j} 成立时 ϕ 的极大似然估计值, 此时定义似然比统计量为

$$\begin{aligned} \Lambda_{-j}(\mathbf{B}) = 2 \ln \frac{\sup_{\phi, \eta, \sigma^2} L(\phi, \eta, \sigma^2)}{\sup_{\phi, \eta_{-j}, \sigma_{-j}^2} L(\phi, \eta_{-j}, \sigma_{-j}^2)} = \\ -n \left[\ln(T(\tilde{\phi}) Y)^T S^T S T(\tilde{\phi}) Y - \frac{2}{n} \ln |T(\tilde{\phi})| - \right. \end{aligned}$$

$$\left. \ln(T(\tilde{\phi}_{-j}) Y)^T S_{-j}^T S_{-j} T(\tilde{\phi}_{-j}) Y + \frac{2}{n} \ln |T(\tilde{\phi}_{-j})| \right].$$

$\tilde{\phi}$ 为全模型成立时 ϕ 的极大似然估计值。当样本量 n 足够大时, $\Lambda_{-j}(\mathbf{B}) \rightarrow \chi^2(m)$ 。对给定的显著性水平 α , 当 $\Lambda_{-j}(\mathbf{B}) > \chi_a^2(m)$ 时, 拒绝原假设 H_{0j} 。

2.3 模型系数的空间局域性判别检验

由于常系数都可以看作是变系数在每个空间点上的系数值相等, 反映了空间全局性的特征, MGWR 模型、MGWR-CSLR 模型等都是本文中所给出 MGWR-VSLR 模型的特殊情形, 因此, 检验估计的系数是否具有局域性而选取变系数形式, 究竟应该选取常系数还是变系数对于设定适宜的模型形式是非常必要的。

1) 空间滞后项对应系数的空间局域性检验

设置零假设 $H_0: \rho(u_i, v_i) \equiv \rho_0, \forall (u_i, v_i) \in \mathbb{R}^2$, 当 H_0 成立时, 相应系数为空间全局性的常系数, MGWR-VSLR 模型退化为 MGWR-CSLR 模型, 如下所示

$$Y = \rho_0 WY + X\alpha + M + \varepsilon. \quad (25)$$

记 $C(\rho_0) = I - \rho_0 W$, 与前文相同, 取 $B = (X, \bar{Z})$ 以及 $\eta = (\alpha^T, \gamma^T)^T$, 上式对应的对数似然函数为

$$\begin{aligned} \ln L(\rho_0, \eta, \sigma^2) = -\frac{n}{2} \ln 2\pi\sigma^2 - \ln |C(\rho_0)| - \\ \frac{1}{2\sigma^2} \{ [C(\rho_0) Y - B\eta]^T [C(\rho_0) Y - B\eta] \}. \end{aligned}$$

此时上式对应的极值可写为

$$\begin{aligned} \sup_{\rho_0, \eta, \sigma^2} \ln L(\rho_0, \eta, \sigma^2) = -\frac{n}{2} \left[\ln \frac{2\pi}{n} - \ln |C(\tilde{\rho}_0)| + \right. \\ \left. \ln(C(\tilde{\rho}_0) Y)^T S^T S C(\tilde{\rho}_0) Y + 1 \right]. \end{aligned} \quad (26)$$

其中 $\tilde{\rho}_0$ 为 ρ_0 在式(25)下的极大似然估计值, $S = I_n - B$ 。因此可以定义似然比统计量

$$\begin{aligned} \Lambda_{cp}(\mathbf{B}) = 2 \ln \frac{\sup_{\phi, \eta, \sigma^2} L(\phi, \eta, \sigma^2)}{\sup_{\rho_0, \eta, \sigma^2} L(\rho_0, \eta, \sigma^2)} = \\ -n \left[\ln(T(\tilde{\phi}) Y)^T S^T S T(\tilde{\phi}) Y - \frac{2}{n} \ln |T(\tilde{\phi})| - \right. \\ \left. \ln(C(\tilde{\rho}_0) Y)^T S^T S C(\tilde{\rho}_0) Y + \frac{2}{n} \ln |C(\tilde{\rho}_0)| \right]. \end{aligned}$$

其中 $\tilde{\phi}$ 为全模型成立时 ϕ 的极大似然估计值。当样本量 n 足够大时, $\Lambda_{cp}(\mathbf{B}) \rightarrow \chi^2(m-1)$ 。对给定的显著性水平 α , 当 $\Lambda_{cp}(\mathbf{B}) > \chi_a^2(m-1)$ 时, 拒绝原假设 H_0 , 空间滞后项对应的系数形式选取变

系数。

2) 自变量对应系数的空间局域性检验

考虑零假设 $H_{0k}:\beta_k(u,v) \equiv \beta_{k0}, \forall (u,v) \in \mathbb{R}^2, k=0,1,\dots,q$, 当假设 H_{0k} 成立时, 相应系数为空间全局性的常系数, 模型变为

$$Y = \rho_n WY + X\alpha + \beta_{k0}Z_k + M_{-k} + \varepsilon. \quad (27)$$

令 $Z_{-k} = [1, Z_1, \dots, Z_k, \dots, Z_{k+1}, \dots, Z_q], \beta_{-k} = [\beta_0, \dots, \beta_{k-1}, \beta_{k+1}, \dots, \beta_q]^T$, $\{d_{-k,1}, \dots, d_{-k,n}\}$ 为 $Z_{-k}\beta_{-k}$ 的对角线元素, 则 $M_{-k} = \text{diag}\{d_{-k,1}, \dots, d_{-k,n}\}I_n$. 记 $B^* = (X, Z_k, \tilde{Z}_{-k})$, 其中 $\tilde{Z}_{-k} = (\tilde{Z}_0, \tilde{Z}_1, \dots, \tilde{Z}_{k-1}, \tilde{Z}_{k+1}, \dots, \tilde{Z}_q), \tilde{Z}_j = (b(u_1, v_1)z_{1j}, \dots, b(u_n, v_n)z_{nj})^T, j=1, \dots, q$. 另记 $\eta^* = (\alpha^T, \beta_{i0}, \gamma_{-k}^T)^T, \gamma_{-k} = (\gamma_0^T, \gamma_1^T, \dots, \gamma_{k-1}^T, \gamma_{k+1}^T, \dots, \gamma_q^T)^T$, 则 $B^*\eta^* = X\alpha + \beta_{k0}Z_k + M_{-k}$, 则式(27)可以简化为

$$Y = T(\phi)^{-1}(B^*\eta^* + \varepsilon).$$

上式的对数似然函数为

$$\ln L(\phi, \eta^*, \sigma^2) = -\frac{n}{2} \ln 2\pi\sigma^2 + \ln |T(\phi)| - \frac{1}{2\sigma^2} \{ [T(\phi)Y - B^*\eta^*]^T [T(\phi)Y - B^*\eta^*] \}.$$

此时对应的极值可写为

$$\sup_{\phi, \eta^*, \sigma^2} \ln L(\phi, \eta^*, \sigma^2) = -\frac{n}{2} \left[\ln \frac{2\pi}{n} - \ln |T(\tilde{\phi}^*)| + \ln (T(\tilde{\phi}^*)Y)^T S^* T(\tilde{\phi}^*)Y + 1 \right]. \quad (28)$$

其中 $\tilde{\phi}^*$ 为 ϕ 在模型(27)下的极大似然估计值, $S^* = I_n - B^*(B^{*T}B^*)^{-1}B^{*T}$. 根据式(18)及式(28), 定义似然比统计量

$$\Lambda_{ck}(B) = 2 \ln \frac{\sup_{\phi, \eta^*, \sigma^2} L(\phi, \eta^*, \sigma^2)}{\sup_{\phi, \eta^*, \sigma^2} L(\phi, \eta^*, \sigma^2)} = -n \left[\ln (T(\tilde{\phi})Y)^T S^T S T(\tilde{\phi})Y - \frac{2}{n} \ln |T(\tilde{\phi})| - \ln (T(\tilde{\phi}^*)Y)^T S^{*T} S^* T(\tilde{\phi}^*)Y + \frac{2}{n} \ln |T(\tilde{\phi}^*)| \right].$$

其中 $\tilde{\phi}$ 为 ϕ 在全模型下的极大似然估计值。当样本量 n 足够大时, $\Lambda_{ck}(B) \rightarrow \chi^2(m-1)$ 。对给定显著性水平 α , 当 $\Lambda_{ck}(B) > \chi^2_\alpha(m-1)$ 时, 拒绝原假设 H_{0k} , 自变量对应的系数形式选取变系数。

3 蒙特卡罗模拟与实际算例应用

3.1 蒙特卡罗模拟的方案设定

乔宁宁^[18]通过数值模拟实验认为 MGWR-CSLR 模型比传统常系数空间滞后模型在空间异

质性处理上表现出更大的优越性。为考察 MGWR-VSLR 模型的参数估计效果和检验统计量, 本文同样基于数值模拟实验在比较 MGWR-VSLR 模型与 MGWR-CSLR 模型的同时, 也引入了 VSLR 和 GWR-VSLR 2 种变系数的空间滞后回归模型, 通过比较系数估计值的均方根误差 (root mean square error, RMSE) 及因变量预测值的 RMSE, 用以对比不同样本容量下系数估计值偏差和因变量预测能力的差异, 4 种模型的具体形式如下:

a) VSLR 模型

$$y_i = \rho(u_i, v_i) \sum_{j=1}^n w_{ij} y_j + \alpha + \beta x_i + \kappa z_i + \varepsilon_i,$$

b) GWR-VSLR 模型

$$y_i = \rho(u_i, v_i) \sum_{j=1}^n w_{ij} y_j + \alpha(u_i, v_i) + \beta(u_i, v_i) x_i + \kappa(u_i, v_i) z_i + \varepsilon_i,$$

c) MGWR-CSLR 模型

$$y_i = \rho \sum_{j=1}^n w_{ij} y_j + \alpha(u_i, v_i) + \beta x_i + \kappa(u_i, v_i) z_i + \varepsilon_i,$$

d) MGWR-VSLR 模型

$$y_i = \rho(u_i, v_i) \sum_{j=1}^n w_{ij} y_j + \alpha(u_i, v_i) + \beta x_i + \kappa(u_i, v_i) z_i + \varepsilon_i.$$

在数值模拟的方案中, 设定 $Y = (y_1, \dots, y_n)^T$ 为因变量, $\rho \sum_{j=1}^n w_{ij} y_j$ 表示常系数空间滞后项, 而 $\rho(u_i, v_i) \sum_{j=1}^n w_{ij} y_j$ 表示变系数空间滞后项, β 表示自变量 $X = (x_1, \dots, x_n)^T$ 对应的全局性系数, $\beta(u_i, v_i)$ 则为自变量 $X = (x_1, \dots, x_n)^T$ 对应的局域性系数, α 和 κ 分别表示截距项和自变量 $Z = (z_1, \dots, z_n)^T$ 对应的全局性系数, 而 $\alpha(u_i, v_i)$ 和 $\kappa(u_i, v_i)$ 则分别为截距项和自变量 $Z = (z_1, \dots, z_n)^T$ 对应的局域性系数, $\varepsilon_i \sim N(0, \sigma^2)$ 。设定研究空间点 i 的变量二维坐标空间 (u_i, v_i) 且横坐标与纵坐标均在 100 km 范围内, 以 0 值为坐标原点且以 100 km 为坐标单位, 则 $u_i, v_i \in [0, 1]$ 。在 VSLR 模型中 $\alpha = 0.5$ 作为固定的值, W 作为空间权重矩阵其权重元素通过不同空间位置的欧式距离倒数获得。考虑到空间存在的线性与非线性关系, 这里假设不同空间点 i 的可变系数 $\rho(u_i, v_i)$, $\beta(u_i, v_i)$ 和 $\kappa(u_i, v_i)$ 分别满足 3 种关于 u_i, v_i 的空间关系, 即 $\rho(u_i, v_i) = (u_i - 0.5)^2 + v_i^2$; $\beta(u_i, v_i) = 0.5(u_i - 2v_i)$; $\kappa(u_i, v_i) = \exp\{0.6u_i\} + \sin(5u_i v_i)$ 。在实验中 Y, X, Z 变量的观测数值由均匀分布

$U(0,1)$ 随机生成,分别独立生成 $n = 200,500,1\,000$ 和 $3\,000$ 的样本量并重复模拟 $N = 50$ 次,同样由均匀分布 $U(0,1)$ 随机生成的测试集样本量为 $n_{\text{test}} = n/10$,在重构参数化求解的模拟中取插值的节点数 $m = \lceil 4\ln 2n \rceil$, $\lceil \cdot \rceil$ 为向下取整函数。通过模型 a) 至 d) 得到不同模型形式的系数估计值分别为 $\tilde{\alpha}, \tilde{\alpha}(u_i, v_i), \tilde{\beta}, \tilde{\beta}(u_i, v_i), \tilde{\kappa}, \tilde{\kappa}(u_i, v_i), \tilde{\rho}$ 和 $\tilde{\rho}(u_i, v_i)$ 。

3.2 不同样本容量下 4 种模型的模拟结果与分析

采用 Xiong^[20] 的重构参数化方法在 $n = 200, 500, 1\,000$, 及 $3\,000$ 不同样本容量下的数值实验模拟中,对上述 4 种模型进行相关系数估计,其结果为表 1 至表 4,对比发现:

1) 当截距项对应的是全局性系数(VSLR 模型)时,系数估计的偏误明显高于截距项对应的是局域性系数(GWR-VSLR 模型、MGWR-CSLR 模型和 MGWR-VSLR 模型),这一特征在不同样本容量下均表现出一致性。由此可以看出,在处理空间异质性问题时,将截距项对应的系数设为局域性系数比全局性系数有助于减小系数估计的偏误。

2) 在不同样本容量下的数值实验模拟中,4 种模型按照因变量预测值的 RMSE 由大到小依次排序为 VSLR 模型>MGWR-CSLR 模型>GWR-VSLR 模型>MGWR-VSLR 模型,即变系数的空间滞后回归模型对于因变量的预测能力明显弱于其他 3 种模型。可以看出在处理空间异质性问题上,除了考虑因变量自身存在的空间相关性,自变量的空间异质性也决定着因变量的预测能力,自变量也需要考虑是否存在局域性自变量,所对

应的系数用以反映局域性特征,提高模型整体的预测能力。

3) 进一步对比存在局域自变量的 GWR-VSLR 模型、MGWR-CSLR 模型和 MGWR-VSLR 模型发现, MGWR-CSLR 模型的自变量 X 所对应系数估计值 $\tilde{\beta}(u_k, v_k)$ 的偏误小于 GWR-VSLR 模型,但是 MGWR-CSLR 模型因变量空间滞后项对应的变系数估计值 $\tilde{\rho}(u_k, v_k)$ 的偏误大于 GWR-VSLR 模型。由此可以看出自变量 X 与 Z 所对应系数并不都适合反映局域性特征,这也是 MGWR 相比 GWR 具有的优势。同时可以看出 GWR-VSLR 模型采用反映因变量空间滞后项的变系数对整个模型的拟合能力也起着重要作用。

4) 相较于 VSLR 模型、GWR-VSLR 模型和 MGWR-CSLR 模型,本文提出的 MGWR-VSLR 模型在不同样本容量的数值模拟实验中,各项系数估计的偏误整体上表现出较小的特征。并且随着样本容量的增加,系数估计偏误有所下降,测试集上因变量预测值的 RMSE 也在下降,体现了估计的稳定性。其他模型的系数估计偏误与因变量预测值的 RMSE 也随着样本容量增大有所下降,说明本文中的随机试验不同模型的模拟结果对比具有一定的可参考性。

3.3 一个实际算例的应用

经济发展与人口总数和海拔高度具有一定相关性。一般来说,一个区域人口越多,存在的经济活动也就越多,而海拔越高,交通条件受限导致经济发展的自然环境相对越差。考察中国 2015 年人口数与海拔高度对国内生产总值(gross domestic product, GDP)的影响,这里随机抽取 1 200 个栅格样本数据,其中 GDP 作为因变量,人

表 1 MGWR-VSLR 模型与其他模型的系数估计的 RMSE 及因变量预测值的 RMSE($n=200$)

Table 1 Root mean square error of estimation of coefficients of four models and that of prediction of their dependent variables($n=200$)					
模型	$\tilde{\alpha}$ 或 $\tilde{\alpha}(u_k, v_k)$	$\tilde{\beta}$ 或 $\tilde{\beta}(u_k, v_k)$	$\tilde{\kappa}$ 或 $\tilde{\kappa}(u_k, v_k)$	$\tilde{\rho}$ 或 $\tilde{\rho}(u_k, v_k)$	因变量的 RMSE
VSLR	1.054 2[482.885] ($<10e-5$)	0.503 1[428.034] ($<10e-5$)	1.943 2[64.349] (0.201 6)	0.259 5[145.721] ($<10e-5$)	4.387 0
GWR-VSLR	0.220 4[1 264.284] ($<10e-5$)	0.099 1[1 109.680] ($<10e-5$)	0.111 3[602.745] ($<10e-5$)	0.138 2[1 636.278] ($<10e-5$)	0.452 1
MGWR-CSLR	0.442 1[1 086.602] ($<10e-5$)	0.049 5[1 037.082] ($<10e-5$)	0.210 4[618.843] ($<10e-5$)	0.312 2[1 586.801] ($<10e-5$)	1.197 4
MGWR-VSLR	0.212 5[1 264.728] ($<10e-5$)	0.042 0[1 102.873] ($<10e-5$)	0.098 7[600.894] ($<10e-5$)	0.128 7[1 590.716] ($<10e-5$)	0.426 1

注: * 方括号内表示系数似然比检验统计量 λ 的值,圆括号内表示检验的 p 值。下同。

表 2 MGWR-VSLR 模型与其他模型的系数估计的 RMSE 及因变量预测值的 RMSE($n=500$)

Table 2 Root mean square error of estimation of coefficients of four models and that of prediction of their dependent variables($n=500$)

模型	$\tilde{\alpha}$ 或 $\tilde{\alpha}(u_k, v_k)$	$\tilde{\beta}$ 或 $\tilde{\beta}(u_k, v_k)$	$\tilde{\kappa}$ 或 $\tilde{\kappa}(u_k, v_k)$	$\tilde{\rho}$ 或 $\tilde{\rho}(u_k, v_k)$	因变量的 RMSE
VSLR	1.058 6[1 211. 637] ($<10e-5$)	0.498 2[1 055. 930] ($<10e-5$)	1.911 7[153. 636] (0.0720)	0.269 9[356. 266] ($<10e-5$)	3.484 0
GWR-VSLR	0.190 6[3 261. 703] ($<10e-5$)	0.072 5[2 761. 260] ($<10e-5$)	0.079 9[1 494. 665] ($<10e-5$)	0.119 4[3 997. 910] ($<10e-5$)	0.426 3
MGWR-CSLR	0.435 4[2 693. 614] ($<10e-5$)	0.049 3[2 600. 823] ($<10e-5$)	0.196 2[1 549. 625] ($<10e-5$)	0.316 6[3 945. 569] ($<10e-5$)	0.827 9
MGWR-VSLR	0.182 9[3 281. 829] ($<10e-5$)	0.041 3[2 767. 132] ($<10e-5$)	0.075 4[1 501. 916] ($<10e-5$)	0.115 1[3 949. 685] ($<10e-5$)	0.387 7

表 3 MGWR-VSLR 模型与其他模型的系数估计的 RMSE 及因变量预测值的 RMSE($n=1\ 000$)

Table 3 Root mean square error of estimation of coefficients of four models and that of prediction of their dependent variables($n=1\ 000$)

模型	$\tilde{\alpha}$ 或 $\tilde{\alpha}(u_k, v_k)$	$\tilde{\beta}$ 或 $\tilde{\beta}(u_k, v_k)$	$\tilde{\kappa}$ 或 $\tilde{\kappa}(u_k, v_k)$	$\tilde{\rho}$ 或 $\tilde{\rho}(u_k, v_k)$	因变量的 RMSE
VSLR	1.107 8[2 462. 569] ($<10e-5$)	0.506 1[2 062. 607] ($<10e-5$)	1.920 1[169. 8423] (0.058 6)	0.272 9[572. 944] ($<10e-5$)	3.442 8
GWR-VSLR	0.164 2[7 077. 416] ($<10e-5$)	0.056 6[5 613. 243] ($<10e-5$)	0.056 4[3 094. 542] ($<10e-5$)	0.104 3[7 987. 799] ($<10e-5$)	0.361 6
MGWR-CSLR	0.427 6[5 440. 620] ($<10e-5$)	0.031 8[5 336. 517] ($<10e-5$)	0.177 7[3 111. 952] ($<10e-5$)	0.312 2[7 915. 378] ($<10e-5$)	0.892 3
MGWR-VSLR	0.165 0[6 958. 840] ($<10e-5$)	0.027 7[5 609. 487] ($<10e-5$)	0.054 5[3 067. 386] ($<10e-5$)	0.103 4[7 932. 484] ($<10e-5$)	0.360 4

表 4 MGWR-VSLR 模型与其他模型的系数估计 RMSE 及因变量预测值的 RMSE($n=3\ 000$)

Table 4 Root mean square error of estimation of coefficients of four models and that of prediction of their dependent variables($n=3\ 000$)

模型	$\tilde{\alpha}$ 或 $\tilde{\alpha}(u_k, v_k)$	$\tilde{\beta}$ 或 $\tilde{\beta}(u_k, v_k)$	$\tilde{\kappa}$ 或 $\tilde{\kappa}(u_k, v_k)$	$\tilde{\rho}$ 或 $\tilde{\rho}(u_k, v_k)$	因变量的 RMSE
VSLR	1.035 7[6 361. 219] ($<10e-5$)	0.486 4[5 499. 320] ($<10e-5$)	1.877 2[235. 455] ($<10e-5$)	0.274 6[1 517. 495] ($<10e-5$)	3.190 1
GWR-VSLR	0.090 2[19 158. 62] ($<10e-5$)	0.078 4[16 350. 09] ($<10e-5$)	0.066 3[9 048. 33] ($<10e-5$)	0.049 8[23 832. 64] ($<10e-5$)	0.162 2
MGWR-CSLR	0.230 7[15 230. 19] ($<10e-5$)	0.009 5[14463. 64] ($<10e-5$)	0.091 4[9 274. 893] ($<10e-5$)	0.155 4[2 3713. 34] ($<10e-5$)	0.464 8
MGWR-VSLR	0.083 6[16 043. 30] ($<10e-5$)	0.019 2[14 901. 71] ($<10e-5$)	0.069 1[8 785. 518] ($<10e-5$)	0.049 1[23 783. 22] ($<10e-5$)	0.160 2

口数 POP (population) 和海拔高度利用数字高程模型 (digital elevation model, DEM) 数据作为自变量, 每个栅格数据的地理坐标为 (u_i, v_i) , 它们均匀覆盖中国 31 省区市, 代表 1 km^2 范围的平均状况。选取 1 000 个栅格数据用于模型拟合, 其余 200 个栅格数据则作为检验模型的预测能力。数据来源于中国科学院地理科学与资源研究所资源环境科学与数据中心 (<https://www.resdc.cn/>)。从栅格单位面积的人口数来看, POP 即人口密度, 相较于人口密度, 交通状况与经济增长更直接

相关, 而交通状况直接受制于海拔地形等自然条件约束。因此, 可以考虑在 MGWR-VSLR 模型中把 DEM 变量作为局域变量, POP 作为全局变量, 同时各地区的经济发展与周边地区经济相关性存在着差异, 采用变系数空间滞后性也具有一定合理性。对于不同空间点的 GDP, 建立 MGWR-VSLR 模型的具体计算公式为

$$\text{GDP}_i = \rho(u_i, v_i) \sum_{j=1}^n w_{ij} \text{GDP}_j + \beta_0(u_i, v_i) + \beta_1 \text{POP}_i + \beta_2(u_i, v_i) \text{DEM}_i + \varepsilon_i, \quad (29)$$

其中空间权重矩阵 $W = (w_{ij})_{i,j=1}^n$ 同样采用欧式距离倒数获得, $\rho(u_i, v_i), \beta_2(u_i, v_i), \beta_1$ 和 $\beta_0(u_i, v_i)$ 为相关变量的系数和截距项。上述回归分析除采用经典的 OLSR 模型和 GWR 模型外,也引入了 MGWR-CSLR 模型,进行 4 种模型整体性能的对比如表 5)。由表 5 可以看出, MGWR-VSLR 模型通过重构参数化的方法在 GDP 数据上的拟合优度高于其他 3 种模型,同时 AIC 值和均方残差最小,说明 MGWR-VSLR 模型估计方法更加精确。另外,在模型的预测方面, MGWR-VSLR 模型的均方预测误差最小,说明该方法能够更好地预测不同空间单元人口和海拔高度对 GDP 的影响。

表 5 4 种模型对 GDP 数据拟合及预测的效果比较
Table 5 Comparison of the fitting and forecasting effect of four models on GDP data

模型评价指标	拟合优度 R^2	AIC 值	均方残差	均方预测误差
OLSR	0.467 2	8 282.190	15.182 7	9.184 7
GWR	0.472 5	8 277.740	15.106 6	9.064 7
MGWR-CSLR	0.502 8	8 210.250	14.665 9	9.084 9
MGWR-VSLR	0.524 3	8 161.059	14.345 6	8.853 2

4 结论

空间计量模型在国民经济社会发展的今天得到了广泛应用,尤其是在面对处理空间相关性和空间异质性等实际的经济地理学问题时,备受学者们的关注。本研究在现有 GWR 各种模型形式的研究基础上,提出一类新的 MGWR-VSLR 模型,该模型不仅将因变量的空间相关性与自变量的空间异质性同时纳入到模型中考虑,同时刻画了因变量的空间局域性特征。研究结论可以概括如下:

1) MGWR-VSLR 模型对于处理空间异质性问题具有更广泛的适用性。在模型形式的构成上, MGWR-VSLR 模型涵盖了 GWR 的多种模型形式,不仅反映了自变量存在空间全局性与局域性的特征,并且进一步考虑了因变量空间局域性的特征。在蒙特卡洛模拟与实际算例应用的结果中显示出 MGWR-VSLR 模型较好的因变量拟合和预测能力,在样本量增大的情况下也呈现出可靠性与稳健性。

2) MGWR-VSLR 模型能够更加广泛地涵盖 GWR 的多种模型形式,并将 MGWR-CSLR 模型做了进一步扩展。基于似然比检验给出 MGWR-

VSLR 模型系数估计的显著性检验,蒙特卡罗模拟结果显示,显著性检验在样本量增大的情况下具有稳健性。同时也给出了空间全局性常系数与局域性变系数选取的统计检验方法,为空间计量模型的合理设置提供了依据。

概言之,本文提出的 MGWR-VSLR 模型,能够为处理空间效应(空间相关性和空间异质性)问题提供一种更加全面的、更具有一般性的定量分析方法,对于常系数与变系数的显著性检验乃至选取的统计检验,也有望推广到其他非参数空间模型的研究中。

参考文献

[1] Cliff A D, Ord J K. Spatial autocorrelation [M]. London, Pion, 1973.

[2] Anselin L. Local indicators of spatial association: LISA [J]. Geographical Analysis, 1995, 27 (2): 93-115. DOI: 10.1111/j.1538-4632.1995.tb00338.x.

[3] Fisher W D. Econometric estimation with spatial dependence [J]. Regional and Urban Economics, 1971, 1 (1): 19-40. DOI:10.1016/0034-3331(71)90016-9.

[4] Kelejian H H, Prucha I R. A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances [J]. The Journal of Real Estate Finance and Economics, 1998, 17 (1): 99-121. DOI:10.1023/A:1007707430416.

[5] Kelejian H H, Prucha I R. A generalized moments estimator for the autoregressive parameter in a spatial model [J]. International Economic Review, 1999, 40 (2): 509-533. DOI:10.1111/1468-2354.00027.

[6] Kelejian H H, Prucha I R. HAC estimation in a spatial framework [J]. Journal of Econometrics, 2007, 140 (1): 131-154. DOI:10.1016/j.jeconom.2006.09.005.

[7] 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. DOI: 10.1016/j.jeconom.2009.10.025.

[8] Lee L F. Consistency and efficiency of least squares estimation for mixed regressive, spatial autoregressive models [J]. Econometric Theory, 2002, 18 (2): 252-277. DOI: 10.1017/S0266466602182028.

[9] Lee L F. Best spatial two-stage least squares estimators for a spatial autoregressive model with autoregressive disturbances [J]. Econometric Reviews, 2003, 22 (4): 307-335. DOI: 10.1081/ETC-120025891.

[10] Lee L F. Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models [J]. Econometrica, 2004, 72 (6): 1899-1925. DOI:10.1111/j.1468-0262.2004.00558.x.

[11] Lee L F. The method of elimination and substitution in the GMM estimation of mixed regressive, spatial autoregressive models[J]. Journal of Econometrics, 2007, 140(1): 155-189. DOI: 10.1016/j.jeconom.2006.09.006.

[12] Lee L F. GMM and 2SLS estimation of mixed regressive, spatial autoregressive models[J]. Journal of Econometrics, 2007, 137 (2): 489-514. DOI: 10.1016/j.jeconom.2005.10.004.

[13] Brunson C. Estimating probability surfaces for geographical point data: an adaptive Kernel algorithm[J]. Computers & Geosciences, 1995, 21(7): 877-894. DOI:10.1016/0098-3004(95)00020-9.

[14] Brunson C, Fotheringham A S, Charlton M E. Geographically weighted regression: a method for exploring spatial nonstationarity[J]. Geographical Analysis, 1996, 28(4): 281-298. DOI:10.1111/j.1538-4632.1996.tb00936.x.

[15] Brunson C, Fotheringham A S, Charlton M. Spatial nonstationarity and autoregressive models [J]. Environment and Planning A: Economy and Space, 1998, 30(6): 957-973. DOI:10.1068/a300957.

[16] Brunson C, Fotheringham A S, Charlton M. Some notes on parametric significance tests for geographically weighted regression[J]. Journal of Regional Science, 1999, 39(3): 497-524. DOI:10.1111/0022-4146.00146.

[17] Fotheringham A S, Charlton M, Brunson C. Two techniques for exploring non-stationarity in geographical data [J]. Geographical Systems, 1997, 4(1): 59-82.

[18] 乔宁宁. 混合地理加权回归模型中的空间相关性检验和参数估计研究[J]. 数量经济技术经济研究, 2013, 30(8): 93-108. DOI:10.13653/j.cnki.jqte.2013.08.020.

[19] Páez A, Uchida T, Miyamoto K. A general framework for estimation and inference of geographically weighted regression models; 2. Spatial association and model specification tests [J]. Environment and Planning A: Economy and Space, 2002, 34(5): 883-904. DOI:10.1068/a34133.

[20] Xiong S F. The reconstruction approach: from interpolation to regression[J]. Technometrics, 2021, 63 (2): 225-235. DOI:10.1080/00401706.2020.1764869.

[21] 魏传华, 梅长林. 半参数空间变系数回归模型的两步估计方法及其数值模拟[J]. 统计与信息论坛, 2005, 20(1): 16-19, 50. DOI: 10.3969/j.issn.1007-3116.2005.01.004.