

The Scale Dependence of the Local Moran's I

Jinjie Feng*, Huiguo Zhang, Xijian Hu#

College of Mathematics and System Sciences, Xinjiang University, Urumuqi Xinjiang
Email: 957725767@qq.com, #xijianhu@126.com

Received: Aug. 16th, 2015; accepted: Sep. 5th, 2015; published: Sep. 8th, 2015

Copyright © 2015 by authors and Hans Publishers Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY).
<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

The paper focuses on the calculation formula of local Moran's I, and uses three standardized processing methods to construct different spatial distance weighted matrix by choosing different scale d , thus the corresponding values of local Moran's I are obtained. Lately, the significance testing of local Moran's I is proceeded based on changeable scales, and the testing result is positive. So, the values local Moran's I are dependent on scale.

Keywords

Local Moran's I, Spatial Distance Weighted Matrix, Standardized Methods, Scale Dependence

局部Moran's I的尺度依赖性

冯金杰*, 张辉国, 胡锡健#

新疆大学数学与系统科学学院, 新疆 乌鲁木齐
Email: 957725767@qq.com, #xijianhu@126.com

收稿日期: 2015年8月16日; 录用日期: 2015年9月5日; 发布日期: 2015年9月8日

摘要

文章针对局部Moran's I值的计算公式, 通过选取不同的尺度 d , 采用三种不同的标准化处理方法构造出

*第一作者。

#通讯作者。

文章引用: 冯金杰, 张辉国, 胡锡健. 局部 Moran's I 的尺度依赖性[J]. 统计学与应用, 2015, 4(3): 162-168.
<http://dx.doi.org/10.12677/sa.2015.43018>

不同的空间距离权重矩阵,求得相应的局部Moran's I值。最后基于尺度的变化,对局部Moran's I值进行显著性检验,检验结果表明尺度变化对局部Moran's I值的变化影响是显著的,证明了局部Moran's I的值对尺度具有依赖性。

关键词

局部Moran's I, 空间距离权重矩阵, 标准化方法, 尺度依赖性

1. 引言

随着应用的深入,空间自相关性分析逐渐成为人们关注的重点,因而产生了许多用来度量空间自相关性的参数,而在这些参数中,Moran's I 是出现最早、应用最广的一个度量参数。Moran's I 指数是由Moran 于 1950 年首先提出的,经过广大学者的努力,特别是 Anselin [1]提出空间局部自相关分析方法,包括局部 Moran's I 指数和 Moran 散点图。Tobler (1970)曾提出地理学第一定律:“任何事物与其它事物都是有联系的,并且距离越近联系就越密切”。空间自相关是空间统计学的重要研究领域,也是研究空间地理单元之间分布关联性的核心理论方法之一[2] [3]。

空间自相关是指同一个变量在不同空间位置上的相关性。空间现象具有自相关性,空间位置上越靠近的现象就越相似。空间自相关的度量方法可以分为全局空间自相关和局部空间自相关。局部 Moran's I 的高值表示具有相似变量值的面积单元的空间聚集(可以是高或低),而局部 Moran's I 的低值说明不相似值的空间单元的空间聚集[4] [5]。

由于研究区域的面积单元是可变的或可修改的,即研究区域的面积单元的尺度是可以变化的,因而随着尺度的变化所得到的统计结果也是有差异的[6]。所以使用不同的尺度对数据进行模拟分析,可以得到从微观到宏观不同详细程度的信息[7] [8]。所以需要通过计算不同尺度下的局部 Moran's I 值及它们的显著性水平来选择合适的尺度。最优的区域方案往往是以数据的可得性和统计区域的划分为前提的,在综合显示数据的空间变化时,需慎重考虑区域的尺度选择。

2. 空间距离权重矩阵标准化的定义及其影响

空间距离权重矩阵的定义[9]如下:

$$p_{ij} = \begin{cases} 1 & \text{当区域 } i \text{ 和 } j \text{ 的距离小于 } d \text{ 时} \\ 0 & \text{其它} \end{cases}$$

空间距离权重矩阵为:

$$P = \begin{pmatrix} 0 & p_{12} & \cdots & p_{1n} \\ p_{21} & 0 & \cdots & p_{2n} \\ \vdots & \vdots & & \vdots \\ p_{n1} & p_{n2} & \cdots & 0 \end{pmatrix}$$

对 P 进行标准化的三种定义[10]如下:

定义 1 (The row-sum standardized W-coding scheme):

$$d = \left(\sum_j^n p_{1j}, \sum_j^n p_{2j}, \cdots, \sum_j^n p_{nj} \right)$$

$$W = [\text{diag}(d)]^{-1} P$$

定义 2 (The globally standardized C-coding scheme):

$$D = \sum_{ij} p_{ij} \quad (i = 1, 2, \dots, n; j = 1, 2, \dots, n)$$

$$C = \frac{n}{D} P$$

定义 3 (The variance stabilizing S-coding scheme)

$$q = \left[\left(\sum_{j=1}^n p_{1j} \right)^{\frac{1}{2}}, \left(\sum_{j=1}^n p_{2j} \right)^{\frac{1}{2}}, \dots, \left(\sum_{j=1}^n p_{nj} \right)^{\frac{1}{2}} \right]$$

$$S^* = [\text{diag}(q)]^{-1} P$$

$$Q = \sum_{ij} s_{ij}^* \quad (i = 1, 2, \dots, n; j = 1, 2, \dots, n)$$

$$S = \frac{n}{Q} S^*$$

则对应的局部 Moran's I 的值如下:

$$I_i^w = \frac{(y_i - \bar{y}) \sum_{j=1}^n w_{ij} (y_j - \bar{y})}{\frac{1}{n} \sum_{j=1}^n (y_i - \bar{y})^2}, \quad I_i^c = \frac{(y_i - \bar{y}) \sum_{j=1}^n c_{ij} (y_j - \bar{y})}{\frac{1}{n} \sum_{j=1}^n (y_i - \bar{y})^2}, \quad I_i^s = \frac{(y_i - \bar{y}) \sum_{j=1}^n s_{ij} (y_j - \bar{y})}{\frac{1}{n} \sum_{j=1}^n (y_i - \bar{y})^2}$$

本文选用 13*13 的网格，并把它分成了 9 块小的区域(见图 1)，它们的顺序依次为(按从左往右)为区域 a: 4*7 服从 N(0, 1)，区域 b: 4*3 服从 N(2, 1)，区域 c: 4*3 服从 N(0, 1)，区域 d: 5*5 服从 N(3, 1)，区域 e: 5*6 服从 N(0.5, 1)，区域 f: 5*2 服从 N(1, 1)，区域 g: 4*6 服从 N(2, 1)，区域 h: 4*3 服从 N(0, 1)，区域 i: 4*4 服从 N(3, 1)。然后在同一尺度下采用上述三种标准化方法进行模拟。通过模拟结果可以发现：在这 169 个点中的任何一点的 Moran's I 值中，三种标准化方法对应的 Moran's I 值的正负都是相同的。

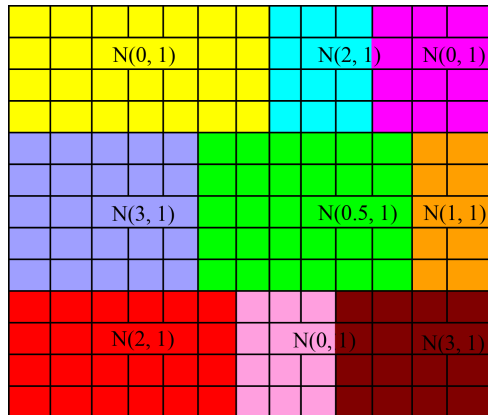


Figure 1. The regional distribution picture
图 1. 区域分布图

这说明了不论采取哪一种标准化方法,若该区域的 Moran's I 为高值则在上述三种标准化处理之下同为高值,表示都具有相似变量值的面积单元的聚集(可以是高或低),若该区域的 Moran's I 为低值则同为低值,表示都具有不相似变量值的面积单元的聚集。但是这三种定义对应的局部 Moran's I 值也有一些小的差异,在每一个点处, I_i^w 是最大的, I_i^c 是最小的, I_i^s 处于二者之间。因此标准化方法 S-coding 是标准化方法 W-coding 与标准化方法 C-coding 的一个综合,所以在对空间矩阵进行标准化时应该首先考虑选 S-coding 进行处理。

3. 对模拟结果进行显著性检验

定义(局部 Moran's I 的期望和方差)

期望:

$$E(I_i^w) = -\frac{\sum_{j=1}^n W_{ij}}{n-1}$$

方差:

$$\text{Var}[I_i^w] = W_i^2 \frac{(n - m_4/m_2^2)}{n-1} + 2W_i(kh) \frac{(2m_4/m_2^2 - n)}{(n-1)^2} - \frac{W_i^2}{(n-1)^2}$$

其中:

$$m_2 = \frac{1}{n} \sum_{j=1}^n (y_j - \bar{y})^2, \quad m_4 = \frac{1}{n} \sum_{j=1}^n (y_j - \bar{y})^4$$

$$w_i^2 = \left(\sum_{j=1}^n w_{ij} \right)^2, \quad w_i^2 = \sum_{j=1}^n w_{ij}^2, \quad 2w_i(kh) = \sum_{i \neq k} \sum_{i \neq h} w_{ik} w_{ih}$$

则对其进行正态标准化有:

$$\bar{I}_i^w = \frac{I_i^w - E(I_i^w)}{\sqrt{\text{Var}[I_i^w]}}$$

同理可得另外两种:

$$\bar{I}_i^c = \frac{I_i^c - E(I_i^c)}{\sqrt{\text{Var}[I_i^c]}}, \quad \bar{I}_i^s = \frac{I_i^s - E(I_i^s)}{\sqrt{\text{Var}[I_i^s]}}$$

经化简后可得: $\bar{I}_{is}^w = \bar{I}_{is}^c = \bar{I}_{is}^s$, 即在同一尺度下它们的 P 值是相同的, 我们取 $\alpha = 0.05$, d 从 2 到 32 以步长 6 递增, 显著性如下图 2~7。

从模拟结果可以发现从 $d = 2$ 到 $d = 8$ 的变化过程中显著性是逐渐增强的, 从 $d = 8$ 到 $d = 32$ 的变化过程中显著性逐渐减弱, 所以我们选择合适的 d 应该在 8 左右。从上述结果可以看出, 在用局部 Moran's I 进行空间相关性分析时, 由于尺度选取的不同会对结果产生比较显著的差异, 如果尺度选取过大就会丢失一些重要的细节, 如果尺度选取过小就无法完全对来自不同空间单元的数据进行比较。

4. 结论

局部 Moran's I 值主要反应了空间单元面积与其邻近单元面积之间的相似程度, 并且它与尺度的选取有着密切的关系。模拟结果显示标准化方法 S-coding 是标准化方法 W-coding 与 C-coding 的一个综合,

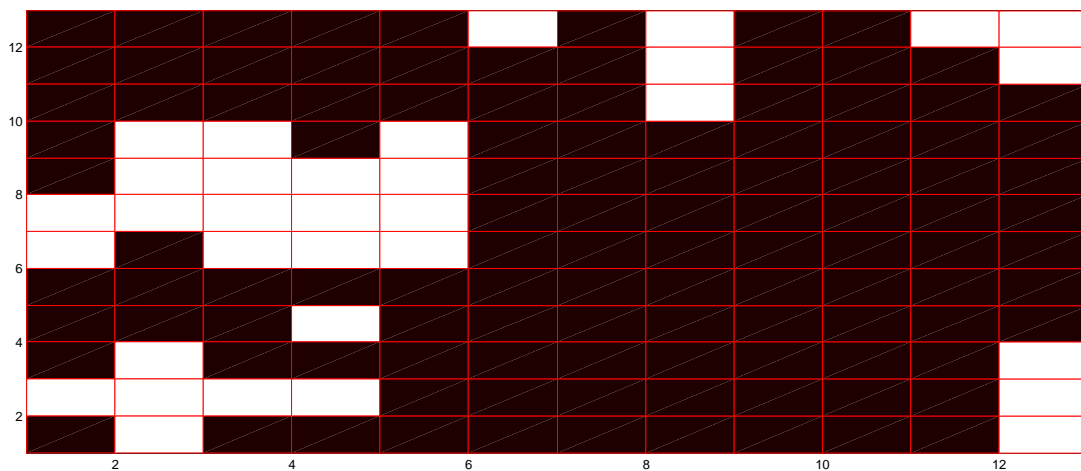


Figure 2. The significance of Local Moran's I when d equals to 2 units length
图 2. 距离 d 为两个单位长度时的局部 Moran's I 显著性

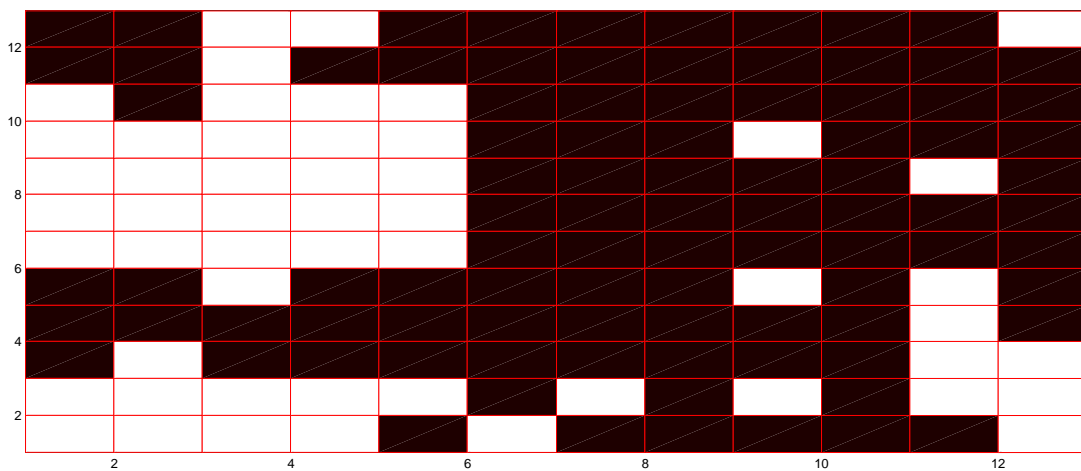


Figure 3. The significance of Local Moran's I when d equals to 8 units length
图 3. 距离 d 为八个单位长度时的局部 Moran's I 显著性

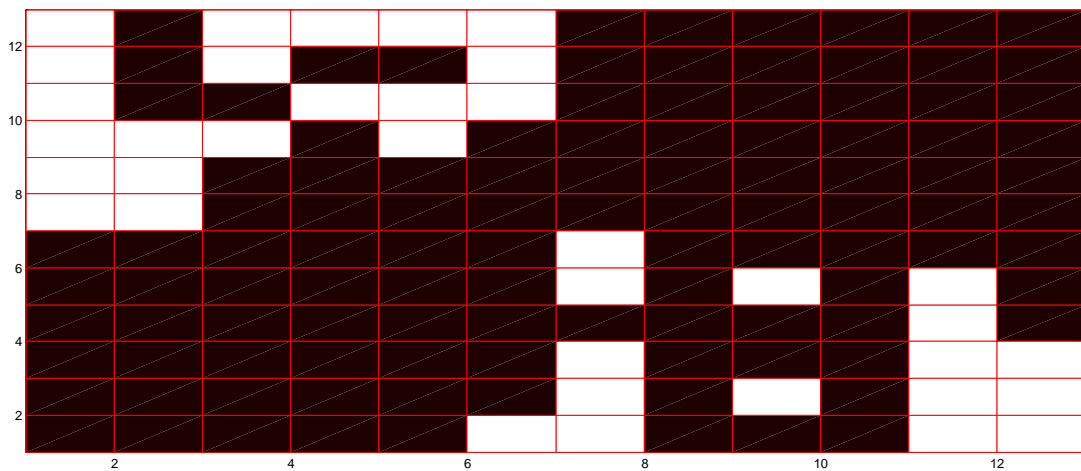


Figure 4. The significance of Local Moran's I when d equals to 14 units length
图 4. 距离 d 为十四单位长度时的局部 Moran's I 显著性

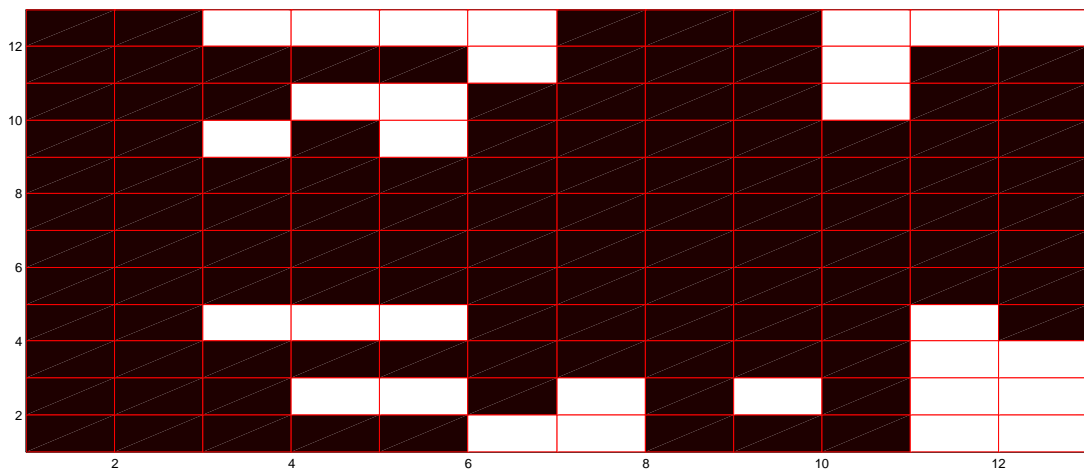


Figure 5. The significance of Local Moran's I when d equals to 20 units length

图 5. 距离 d 为二十个单位长度时的局部 Moran's I 显著性

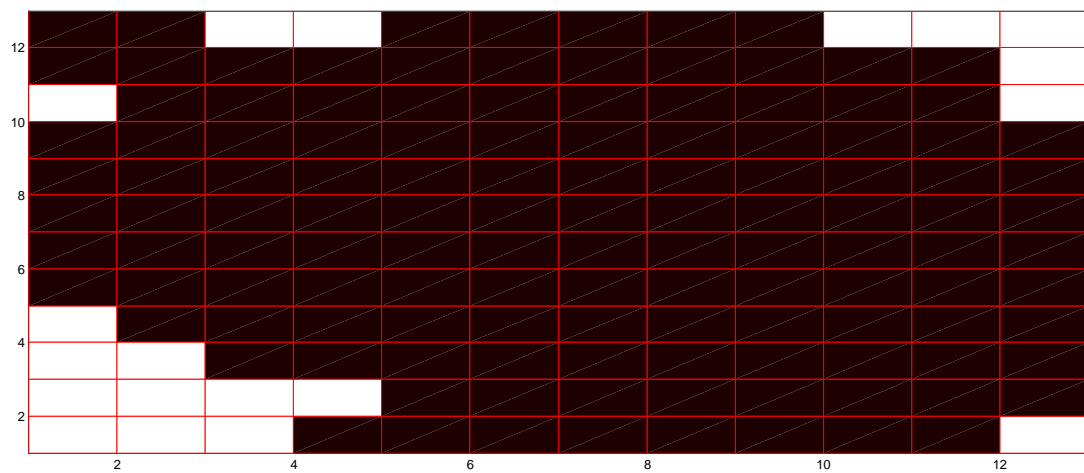


Figure 6. The significance of Local Moran's I when d equals to 26 units length

图 6. 距离 d 为二十六个单位长度时的局部 Moran's I 显著性

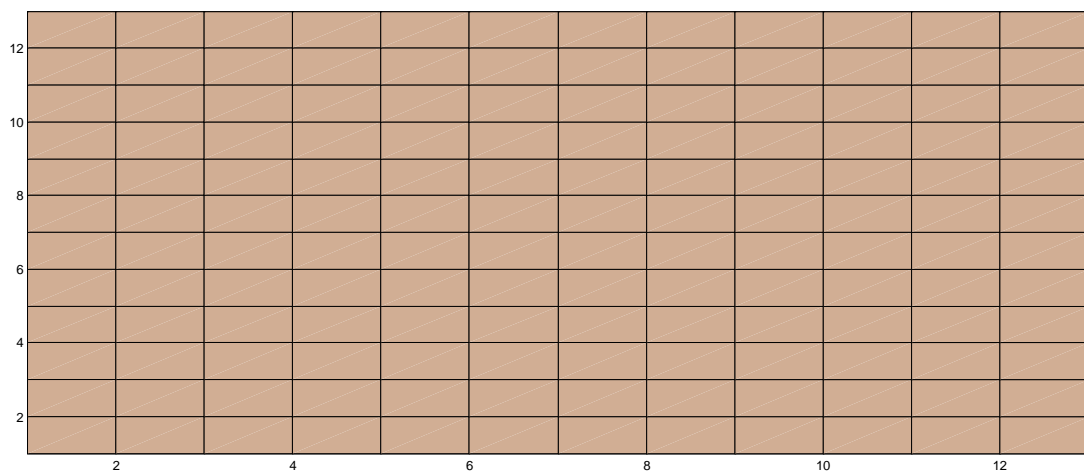


Figure 7. The significance of Local Moran's I when d equals to 32 units length

图 7. 距离 d 为三十二个单位长度时的局部 Moran's I 显著性

所以在对空间矩阵进行标准化时应该首先考虑 S-coding。检验结果表明选取不同的尺度所得到的结果有明显的差异。因此在用局部 Moran's I 值进行空间自相关性检验时, 需慎重考虑区域尺度的选择。

基金项目

国家自然科学基金(41261087), 教育部青年基金(12XJJC910001), 新疆文科基地重大项目基金(0601920)。

参考文献 (References)

- [1] Anselin, L. (1995) Local indicators of spatial association—LISA. *Geographical Analysis*, **27**, 93-115. <http://dx.doi.org/10.1111/j.1538-4632.1995.tb00338.x>
- [2] Cliff, A.D. and Ord, J.K. (1972) Testing for spatial autocorrelation among regression residuals. *Geographical Analysis*, **4**, 267-284. <http://dx.doi.org/10.1111/j.1538-4632.1972.tb00475.x>
- [3] Ord, J.K. and Getis, A. (2001) Testing for local spatial autocorrelation in the presence of global autocorrelation. *Regional Science*, **41**, 411-432.
- [4] Assuncao, R. and Reis, E. (1999) A new proposal to adjust Moran's I for population density. *Statistic in Medicine*, **18**, 2147-2162. [http://dx.doi.org/10.1002/\(SICI\)1097-0258\(19990830\)18:16<2147::AID-SIM179>3.0.CO;2-I](http://dx.doi.org/10.1002/(SICI)1097-0258(19990830)18:16<2147::AID-SIM179>3.0.CO;2-I)
- [5] Cliff, A.D. and Ord, J.K. (1981) *Spatial processes: Models and applications*. London.
- [6] 李海萍 (2009) 空间统计分析中的 MAUP 及其影响. *统计与决策*, **22**, 15-17.
- [7] Tiefelsdorf, M. (1998) Some practical applications of Moran's I exact conditional distribution. *Papers in Regional Science*.
- [8] Tiefelsdorf, M. and Boots, B.N. (2005) The exact distribution of Moran's I. *Environment and Planning*, **27**, 985-999.
- [9] 徐彬 (2007) 空间权重矩阵对 Moran's I 指数影响的模拟分析. 南京师范大学, 南京.
- [10] Tiefelsdorf, M. and Boots, B.N. (1997) A note on the extremities of local Moran's I and their impact on global Moran's I. *Geographical Analysis*, **29**, 248-257. <http://dx.doi.org/10.1111/j.1538-4632.1997.tb00960.x>