

Spatial Econometric

Modern Models and Methods

空间计量经济学

现代模型与方法

王周伟 崔百胜 张元庆 编著



北京大学出版社
PEKING UNIVERSITY PRESS

Spatial Econometrics

Modern Models and Methods

空间计量经济学

现代模型与方法



王桐伟 崔百胜 张元庆 编著



北京大学出版社
PEKING UNIVERSITY PRESS

图书在版编目 (CIP) 数据

空间计量经济学：现代模型与方法 / 王周伟，崔百胜，张元庆编著. —北京：北京大学出版社，
2017.3

ISBN 978-7-301-28067-6

I . ①空… II . ①王… ②崔… ③张… III . ①区位经济学 - 计量经济学 - 研究 IV . ① F224.0

中国版本图书馆 CIP 数据核字 (2017) 第 024491 号

- 书 名** 空间计量经济学：现代模型与方法
Kongjian Jiliang Jingjixue: Xiandai Moxing yu Fangfa
- 著作责任者** 王周伟 崔百胜 张元庆 编著
- 责任编辑** 旷书文
- 标准书号** ISBN 978-7-301-28067-6
- 出版发行** 北京大学出版社
- 地 址** 北京市海淀区成府路 205 号 100871
- 网 址** <http://www.pup.cn> 新浪微博：@北京大学出版社
- 电子信箱** zpup@pup.cn
- 电 话** 邮购部 62752015 发行部 62750672 编辑部 62757857
- 印 刷 者** 三河市北燕印装有限公司
- 经 销 者** 新华书店
- 787 毫米 × 1092 毫米 16 开本 20.25 印张 460 千字
2017 年 3 月第 1 版 2017 年 3 月第 1 次印刷
- 定 价** 59.00 元

未经许可，不得以任何方式复制或抄袭本书之部分或全部内容。

版权所有，侵权必究

举报电话：010-62752024 电子信箱：fd@pup.pku.edu.cn

图书如有印装质量问题，请与出版部联系，电话：010-62756370

第6章至第9章为特殊的单方程空间计量经济模型部分。第6章是空间离散选择模型,包括空间 Probit 模型、空间 Tobit 模型、序数空间概率模型、结构化效应的空间概率单位模型、贝叶斯估计的空间概率模型。第7章是空间分位数回归模型,包括分位数回归概述、空间自回归模型的分位数回归估计、空间分位数回归模型估计的 Stata 软件实现。第8章是贝叶斯空间计量经济模型,包括贝叶斯估计概述、线性单方程计量经济学模型的贝叶斯估计、贝叶斯 FAR 模型、贝叶斯 SAR 模型、贝叶斯空间计量经济模型的 MCMC 估计。第9章是变系数空间计量经济模型,包括变系数空间计量经济模型概述、地理加权回归模型的估计与检验、混合地理加权回归模型的估计与检验、地理加权空间计量模型的估计与检验、面板数据的地理加权空间回归模型。

第10章至第13章为多方程的空间计量经济模型部分。第10章是空间面板模型,包括空间面板模型概述、空间面板混合模型、空间面板变截距模型、空间面板变截距参数模型、动态空间面板数据模型、空间面板似不相关模型、空间面板误差分量模型。第11章是空间联立方程模型,包括空间联立方程模型概述、空间联立方程模型的假设、空间联立方程模型检验、空间联立方程模型的估计。第12章是空间向量自回归模型,包括空间向量自回归模型概述、设定、估计。第13章是空间误差修正模型,包括空间误差修正模型的概述、设定、检验与估计。

第14章至第16章为利用特殊的空间权重矩阵的空间计量经济模型与方法部分。第14章是时空相关经济模型,包括时空调整的空间计量经济模型概述、空间可加经济模型的估计与检验。第15章是空间计量经济交互模型,包括空间计量经济交互模型概述与估计。第16章是矩阵指数空间计量经济模型,包括矩阵指数空间计量经济模型概述、MESS 空间自回归模型的设定与似然估计、MESS 空间自回归模型的贝叶斯估计方法、MESS 空间自回归模型的扩展及其贝叶斯估计、MESS 空间误差模型的设定与似然估计、MESS 空间杜宾模型的设定与似然估计、MESS 空间计量模型的最新研究进展。

本书编写过程中,参阅了大量的国内外文献、教材与专著,特别是詹姆斯·勒沙杰、凯利·佩斯主编的《空间计量经济学导论》,保罗·埃尔霍斯特编著的《空间计量经济学——从横截面数据到空间面板》,沈体雁、冯等田、孙铁山编著的《空间计量经济学》,林光平、龙志和编著的《空间计量经济学:理论与实证》等等。在此我们一并表示最诚挚的谢意。

本书提供每章的教学课件、研究数据和软件实现主要过程及上机练习的数据资料,具体下载地址为上海师范大学金融工程研究中心网页(<http://fb.shnu.edu.cn/Default.aspx?tabid=13755>)。读者可以根据自己的兴趣与需要,使用软件进行操作练习,以熟练掌握空间计量经济模型与分析方法。

经过空间计量系列研讨,本书由王周伟、崔百胜、张元庆三位老师与所指导的研究生们共同合作完成的。由王周伟拟定全书撰写大纲,确定内容框架体系;团队成员分头撰写初稿,相互交叉修改,王周伟、崔百胜、张元庆三位老师与伏开宝博士共同终审,最终定稿。对本书写作分工如下:

1. 胡金群、唐虞、梁士勇、王周伟,第1章,空间经济效应及其统计分析
2. 李超、王周伟,第2章,空间过滤动态合成数据模型
3. 柳闫、王衡、杨欢、欧阳立轩、付妙玉、王周伟,第3章,截面空间计量经济模型
4. 张岚、王周伟,第4章,空间计量经济模型的选择与模拟分析
5. 陶志鹏、王周伟、伏开宝,第5章,空间计量经济模型的估计方法
6. 林丽、赵正丽、孟祥荣、赵星、郭得庆、侯志坚、李昌盛、崔百胜,第6章,空间离散选择模型
7. 胡宇晓、崔百胜,第7章,空间分位数回归模型
8. 梁士勇、张元庆、伏开宝,第8章,贝叶斯空间计量经济模型
9. 赵宗源、张元庆,第9章,变系数空间计量经济模型
10. 吴平、高蓓蓓、杨晓勤、朱麟、徐丽雯、崔百胜,第10章,空间面板模型
11. 谢沐、崔百胜,第11章,空间联立方程模型
12. 王建金、崔百胜,第12章,空间向量自回归模型
13. 杜永康、张元庆,第13章,空间误差修正模型
14. 宋韞贇、张元庆,第14章,时空相关经济模型
15. 刘晓罡、张元庆,第15章,空间计量经济交互模型
16. 周水松、崔百胜、伏开宝,第16章,矩阵指数空间计量经济模型

当然,由于自身水平有限,而空间计量经济学又处在蓬勃发展的态势中,编写过程中难免存在不足、遗漏与错误,恳请同行专家和读者对本书提出宝贵意见和建议,使我们今后能够不断对之加以完善。

编者

2015年11月

目 录

第 1 章 空间经济效应及其统计分析	1
1.1 空间相关性	1
1.2 空间权重矩阵	19
1.3 空间异质性	31
1.4 空间收敛性	37
1.5 空间集聚性	40
1.6 空间扩散效应	47
1.7 空间辐射力与范围分析	48
1.8 空间耦合协调	50
1.9 空间区域网络	53
第 2 章 空间过滤动态合成数据模型	62
2.1 空间过滤动态合成数据模型的建立	62
2.2 空间滤波方法	63
2.3 空间过滤动态合成数据模型的估计与检验	67
第 3 章 截面空间计量经济模型	70
3.1 基本的截面空间计量经济模型	70
3.2 空间杜宾模型	73
3.3 空间杜宾误差模型	74
3.4 截面空间误差分量模型	80
3.5 一般空间模型	81
3.6 条件自回归模型	82
3.7 广义空间自回归(GSAC)模型	88
3.8 广义嵌套式空间模型	89
3.9 动态截面空间计量模型	89
3.10 动态截面空间 SUR 模型	90
第 4 章 空间计量经济模型的选择与模拟分析	92
4.1 引 言	92

4.2	空间计量模型选择概述	93
4.3	基于统计检验的空间计量经济模型选择方法	94
4.4	基于信息准则的空间计量经济模型选择方法	96
4.5	基于模型后验概率的贝叶斯选择方法	97
4.6	基于 MCMC 的空间计量经济模型选择方法	98
第 5 章	空间计量经济模型的估计方法	101
5.1	极大似然估计法	101
5.2	拟极大似然估计法	103
5.3	两阶段最小二乘法与工具变量法	105
5.4	广义矩估计法	107
5.5	贝叶斯估计法	108
5.6	非参数估计方法	113
5.7	有限样本情况下空间计量经济模型的检验与估计	114
第 6 章	空间离散选择模型	119
6.1	空间 Probit 模型	119
6.2	空间 Tobit 模型	123
6.3	序数空间概率模型	127
6.4	结构化效应的空间概率单位模型	132
6.5	贝叶斯估计的空间概率模型	139
第 7 章	空间分位数回归模型	150
7.1	分位数回归概述	150
7.2	空间自回归模型的分位数回归估计	151
7.3	空间分位数回归模型估计的 Stata 软件实现	153
第 8 章	贝叶斯空间计量经济模型	157
8.1	贝叶斯估计概述	157
8.2	线性单方程计量经济学模型的贝叶斯估计	164
8.3	贝叶斯 FAR 模型	169
8.4	贝叶斯 SAR 模型	171
8.5	贝叶斯空间计量经济模型的 MCMC 估计	174
第 9 章	变系数空间计量经济模型	179
9.1	变系数空间计量经济模型概述	179
9.2	地理加权回归模型的估计与检验	179

9.3	混合地理加权回归模型的估计与检验	181
9.4	地理加权空间计量模型的估计与检验	188
9.5	面板数据的地理加权空间回归模型	191
第 10 章	空间面板模型	195
10.1	空间面板模型概述	195
10.2	空间面板混合模型	198
10.3	空间面板变截距模型	201
10.4	空间面板变截距参数模型	209
10.5	动态空间面板数据模型	222
10.6	空间面板似不相关模型	232
10.7	空间面板误差分量模型	237
第 11 章	空间联立方程模型	267
11.1	空间联立方程模型概述	267
11.2	空间联立方程模型的假设	268
11.3	空间联立方程模型的检验	270
11.4	空间联立方程模型的估计	271
第 12 章	空间向量自回归模型	275
12.1	空间向量自回归模型概述	275
12.2	空间向量自回归模型设定	275
12.3	空间向量自回归模型的估计	276
第 13 章	空间误差修正模型	279
13.1	空间误差修正模型概述	279
13.2	空间误差修正模型的设定	280
13.3	空间误差修正模型的检验	281
13.4	空间误差修正模型的估计	283
第 14 章	时空相关经济模型	286
14.1	时空调整的空间计量经济模型概述	286
14.2	时空可加经济模型的估计与检验	291
第 15 章	空间计量经济交互模型	296
15.1	空间计量经济交互模型概述	297
15.2	空间计量经济交互模型的估计	302

第 16 章 矩阵指数空间计量经济模型	307
16.1 矩阵指数空间计量经济模型概述	307
16.2 MESS 空间自回归模型的设定与似然估计	307
16.3 MESS 空间自回归模型的贝叶斯估计方法	308
16.4 MESS 空间自回归模型的扩展及其贝叶斯估计	309
16.5 MESS 空间误差模型的设定与似然估计	311
16.6 MESS 空间杜宾模型的设定与似然估计	312
16.7 MESS 空间计量模型的最新研究进展	313

第1章 ■ 空间经济效应及其统计分析

1.1 空间相关性

空间相关性,也称空间依赖性,是指不同区域的事物和现象之间在空间上的互相依赖、互相制约、互相影响和互相作用,是事物和现象本身所固有的空间经济属性,是地理空间现象和空间过程的本质属性。空间依赖可以定义为观测值及区位之间的依赖性。当相邻区域特征变量的高值或低值在空间上呈现集聚倾向时为正的空间自相关,反之,当相邻区域特征变量取值与本区域变量取值高低相反时则为负的空间自相关。

1.1.1 区域经济溢出视角的空间相关性理论分析

区域经济具有明显的开放性,因为区域经济系统之间存在着广泛的商品贸易技术扩散和要素流动等各种经济联系,所以区域之间广泛存在着溢出反馈关联效应。同时,区域之间的空间溢出效应对相邻区域的经济增长也存在着显著的贡献,不完全竞争和规模报酬递增的条件下,金融外部性有着清晰的福利效应,这种福利效应引起正向或负向的空间溢出效应,正向并外溢效应大的地区会形成良性的“因果循环积累效应”,不仅能有效地促进本地区的经济发展,还能依靠这种正向外部性造成的空间溢出促进周边地区的经济发展;而具有正向性、外部性不明显的地区,则只能促进本地区的经济发展,并不能带动其他地区的经济发展,形成较弱的,区域经济差异明显的区域经济关联;而具有负外部性、溢出效应大的地区会通过空间溢出效应造成周边地区也承担这种非正向的经济外部性,从而造成区域经济之间的恶性循环机制,形成边缘区域。因此,区域经济溢出效应构成了空间相关性的原因。下面将从市场潜能、外部经济、增长极理论、扩散理论、涓滴理论、金融相关六个方面分析区域经济溢出影响:

1. 市场潜能

市场潜能指以空间距离为权重,把具有空间相关性的周边地区国内生产总值加权汇总,来衡量一个地区生产产品和服务的潜在需求规模。这个指标说明了本地区的经济发展同时受本地区到其他地区市场潜能的影响,即当本地区市场潜能高时,相应的其他周边地区市场潜能也会高,这实际上就表现为区域经济的外部性和关联性^①。

2. 外部经济

其含义是一些人的经济活动会给另一些人带来收益或损失,而那些获得收益的人

^① 赵伟光,敬莉. 区域经济关联与经济增长的空间溢出效应[J]. 财经科学, 2015, 3(10):131—140.

无需付费,而蒙受损失的人也无法得到补偿。目前外部经济不再局限于“点”状,而是向区域化方向发展,正是由于外部经济区域化的存在,使得城市与城市之间、城市与郊区之间、郊城镇镇之间的经济联系日趋密切,造成区域经济溢出,使得区域经济之间存在空间相关性。

3. 增长极理论

该理论强调区域经济发展的不平衡性,主张尽量把有限的稀缺资源集中投入到发展潜力大、规模经济和投资效益明显的少数地区或行业,使主导部门或有创新能力的企业或行业在一些地区或大城市聚集,形成一种资本与技术高度集中、具有规模经济效益的增长迅速并能对邻近地区产生强大辐射作用的“增长极”,通过具有“增长极”的地区的优先增长,带动相邻地区的共同发展,形成区域经济溢出效应^①。

4. 扩散理论

扩散理论的出发点主要是特定区域与其他区域的互动关系。从影响因素上看,扩散理论着重探讨特定区域的发展方向,探讨如何在自然资源、资金、技术、交通通讯、人力资源开发等方面受到其他区域的扩散、辐射,或者如何在这些方面依赖于其他区域的,即空间区域经济溢出效应的影响机制。该理论认为随着经济扩张中心地的基础设施的改善等情况的发生,资本、人才等会流向中心地周围地区,进而刺激本地区的发展,逐步赶上中心地^②。

5. 涓滴理论

涓滴效应,也称做“涓滴理论”(又译作利益均沾论、渗漏理论、滴漏理论),指在经济发展过程中并不给予贫困阶层、弱势群体或贫困地区特别的优待,而是由优先发展起来的群体或地区通过消费、就业等方面惠及贫困阶层或地区,带动其发展和富裕。统计显示,2011年北京市人均GDP已超1.2万美元,城市化率在2008年已超80%,这意味着北京应该已进入从“极化效应”向“涓滴效应”过渡的阶段,区域经济溢出效应明显。

6. 金融相关

信息不对称理论是指在市场经济活动中,各类人员对有关信息的了解是有差异的;掌握信息比较充分的人员,往往处于比较有利的地位,而信息贫乏的人员,则处于比较不利的地位。信息腹地对金融集聚的影响主要体现非标准信息对金融机构布局的约束,即信息区位优势是金融集聚的基本条件。由于区域间存在经济溢出效应,由信息不对称理论及信息腹地理论可知,离信息中心区域(经济发达地区)越近的区域,得到的信息质量越高,越有利于区域发展^{③④}。

① 刘芬,邓宏兵,李雪平. 增长极理论、产业集群理论与我国区域经济发展[J]. 华中师范大学学报:自然科学版, 2007, 41(1): 130—133.

② 张敦福. 扩散理论与中国区域发展研究[J]. 山东师大学报:人文社会科学版, 2001, 46(5): 100—102.

③ 李勇,梁琳. 信息腹地、溢出效应与金融集聚研究综述[J]. 金融经济:理论版, 2013(12): 76—78.

④ 王宇伟,范从来. 南京建设区域金融中心的信息腹地战略[J]. 南京社会科学, 2011(1): 143—148.

1.1.2 空间相关性的图形分析

1. 分位数分布图

分位数分布图也称分位数空间分布图,就是将与空间相对应的某一指标值,如某省的人均 GDP,按要求的几分位数等分,并在地图上标出对应的区域,以颜色由浅及深表示数值的由小及大^①。该图可以由 Open Geoda 软件里的 Quantile Map 功能显示。图 1.1 我国各区域经济包容性增长水平四分位图中,由白色到褐色,随着颜色的加深区域经济包容性增长水平也在增大,即白色地区的区域经济包容性增长水平最低,褐色最大,而且可以看到该地区及其相邻地区的区域经济包容性增长水平状况,有助于直观理解该指标的空间相关性。

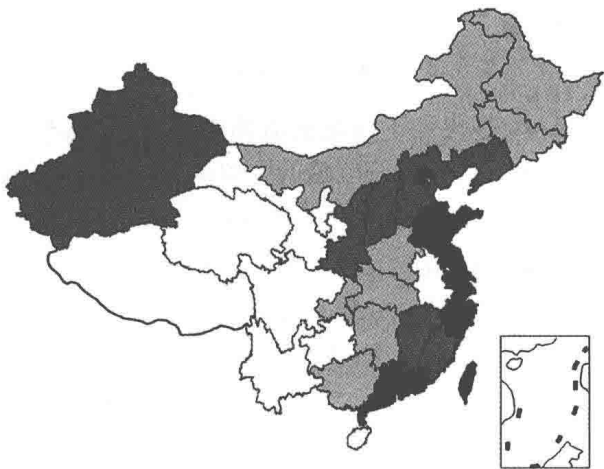


图 1.1 各区域经济包容性增长水平四分位图

2. 热力图

热力图(Heat Map),又叫等值线地图(choropleth map),是超越了传统地图意义的一种新型地图,在现代 GIS(地理空间数据可视化)中使用很广泛。其是传统地图与现代数据结合的产物,最大的特点是它的数据范围有边界,而这个边界,是由传统地图或者说人为划定的。

以中国各地安全指数热力图为例,在图 1.2 上可以很清楚地看到各省及其相邻地区的安全指数,方便直观地理解某个指标是否存在空间相关性。

3. Moran 散点图

Moran 散点图可以进行局部空间自相关分析,对空间滞后因子 W_z 和 z 数据进行可视化的二维图示。Moran 散点图用散点图的形式描述变量 z 与空间滞后(即该观测值周围邻居的加权平均)向量 W_z 间的相互关系。该图的横轴对应变量 z ,纵轴对应空间滞后向量 W_z ,它被分为四个象限,分别识别一个地区及其邻近地区的关系。

Moran 散点图的四个象限分别对应于空间单元与其邻近单元之间四种类型的局部

^① 庞敏. 中国经济包容性增长水平的测度及区域差异分析[D]. 湖南大学, 2013.

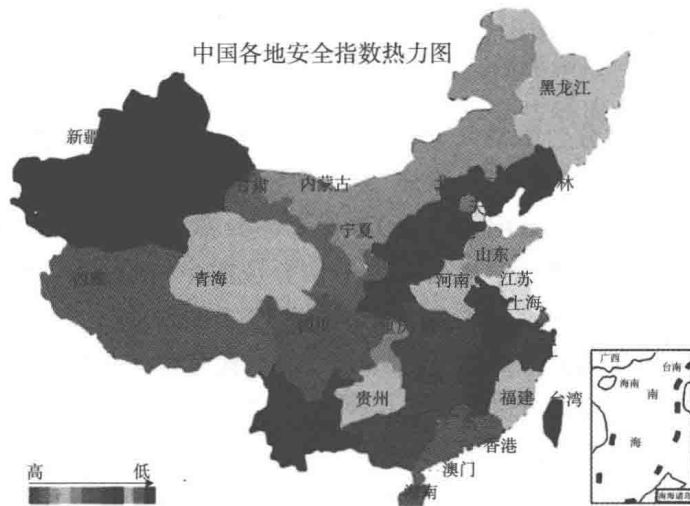


图 1.2 中国各地安全指数热力图

空间联系形式:第一象限(HH)代表了高观测值的空间单元,其相邻区域仍是高值的空间联系形式;第二象限(LH)代表了低观测值的空间单元,其相邻区域是高值的空间联系形式;第三象限(LL)代表了低观测值的空间单元,其相邻区域仍是低值的空间联系形式;第四象限(HL)代表了高观测值的空间单元,其相邻区域是低值的空间联系形式。

以图 1.3 为例,多数点位于第一和第三象限内,即 HH 集聚和 LL 集聚,表现为正相关,说明观测点高值与高值关系紧密,低值与低值关系紧密。

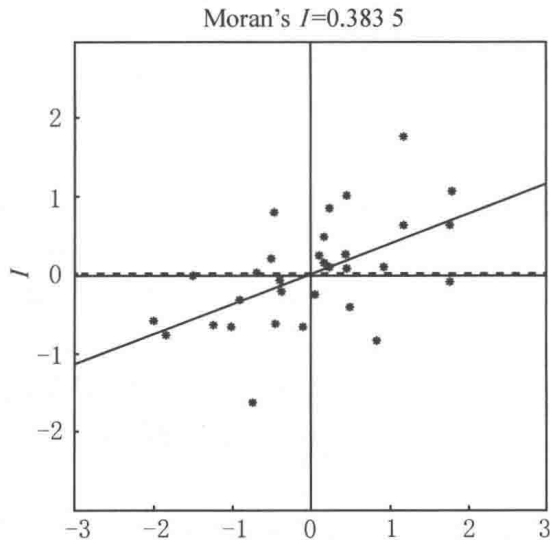


图 1.3 Moran 散点图

与局部 Moran 指数相比,虽然 Moran 散点图不能获得局部空间集聚的显著性指标,但其是形象的二维图像非常易于理解,其重要的优势还在于能够进一步具体区分空间单元和其邻近单元属于高值和高值、低值和低值、高值和低值、低值和高值之中的哪种空间联系形式。并且,对应于 Moran 散点图的不同象限,可识别出空间分布中存

在着哪几种不同的实体。

4. Moran 显著性水平图

将 Moran 散点图与空间联系的局部指标(LISA)显著性水平相结合,可以得到所谓的“Moran 显著性水平图”,图中会显示出显著的 LISA 区域,并分别标识出对应于 Moran 散点图中不同象限的相应区域。

区域 i 的局部 Moran 指数可以写为: $I_i = z_i \sum_{j \neq i} w_{ij} z_j$, 式中, z_i 和 z_j 是区域 i 和 j 上观测值的标准化, w_{ij} 是空间权重, 其中 $\sum_j w_{ij} = 1$ 。

若 I_i 大于 0 且 z_i 大于 0, 则区域 i 位于 HH 象限; 若 I_i 大于 0 而 z_i 小于 0, 则区域 i 位于 LL 象限; 若 I_i 小于 0 而 z_i 大于 0, 则区域 i 位于 HL 象限; 若 I_i 小于 0 而 z_i 小于 0, 则区域 i 位于 LH 象限。以上四个象限的含义与 Moran 散点图相同, I_i 的显著性可以采用 Bonferroni 标准加以判断。当总的显著性水平设定为 α 时, 每一个区域的显著性要根据 α/n 的原则进行判断。在给定显著性水平 α 时, 若 I_i 显著大于 0, 说明区域 i 与周边地区之间的空间差异显著小; 若 I_i 显著小于 0, 说明区域 i 与周边地区之间的空间差异显著大。

5. 分位数分布图的 ArcMap 制作

地图显示在 ArcMap 中丰富多彩, 单一符号地图、唯一值地图、颜色分级地图、符号分级地图、多变量地图, 甚至图标地图、栅格地图、TIN 地表地图, 多种多样的地图显示表明了 ArcMap 作为地理信息系统平台在地图制作中的优势和特点。

虽然每种地图显示方式不同, 但是使用方法都大同小异, 在此为了与 GeoDa 作对比, 仅提供“分级显示”的软件操作方法。

分级显示是在实际操作中经常遇到的分类方法, 用颜色、符号和比例的方式表达分级可以更加清晰地表达数据分级, 在实际操作特别是专题应用中经常遇到。在分级显示中, 有“分级色彩”“分级符号”和“比例符号”。本部分只介绍“分级色彩”, 其它类同。

在对数据进行分类时, 可以根据需求使用 ArcMap 提供的任何一种标准分类方案, 也可以创建自定义分类方案。如果对数据进行主动分类, 只需要选择相应的分类方案, 并设定分类数目即可。如果想自定义分类, 可以通过手工操作来添加类的间断点, 并设置分类范围, 从而创建适合用户数据的分类标准。

用分级颜色对数据进行符号化, 主要是通过改变符号颜色及大小来表达数量的变化, 比如可以用橙色色度的变化, 表达不断变化的省会城市的人均 GDP。

具体操作方法如下:

右击目标图层, 例如本例中目标图层为“BountRegionProv_clip.shp”的面状图层, 在弹出来的快捷菜单中选择【属性】命令, 弹出【图层属性】对话框。

单击【符号系统】标签, 在“显示”选项中单击【数量】→【分级色彩】。

在【字段】的“值”下拉列表中选择“人均 GDP”, 在“归一化”下拉列表框中可以不选择(表示不需要对数据进行标准化处理), 在“类”下拉列表框中选择“4”, 表示将 31 省份的数据值分成 4 类, 如图 1.4 所示。默认分类方法是“自然间断点分级法”, 点击

“分类”选项卡,弹出【分类】对话框,如图 1.5 所示。点击“方法”选项卡后的下拉箭头,如图 1.6 所示,可以看见,分类方法有“手动”“相等间隔”“定义的间隔”“分位数”“自然间断点分类法”“几何间隔”“标准差”。在这里,我们选择“分位数”分类法。“类别”选项卡中填写“4”,这样,我们对 31 省人均 GDP 所做的分类,即为四分位法。完成后,点击“确定”,得到 31 省份人均 GDP 的四分位图,如图 1.7 所示。其他分级显示的方法和效果,读者可以自己尝试。

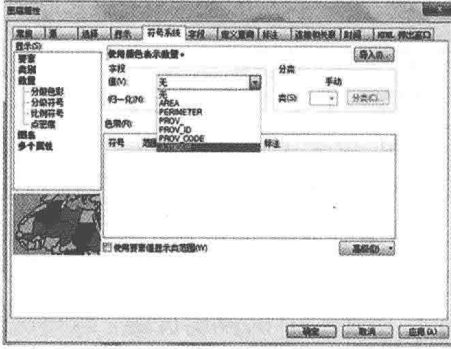


图 1.4 分级色彩设置界面

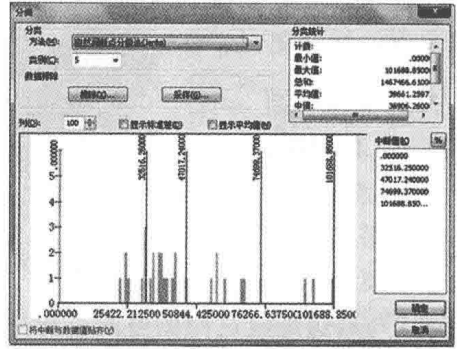


图 1.5 分类方法和类别选择

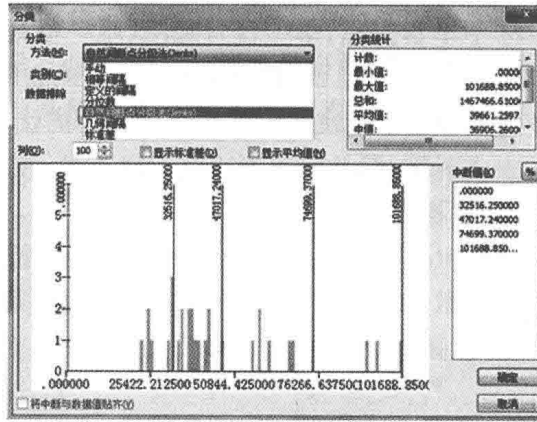


图 1.6 分类方法的选择

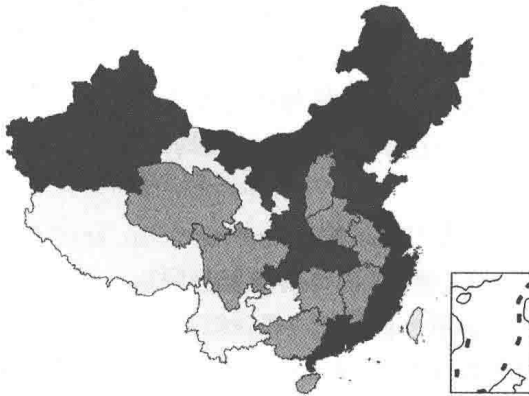


图 1.7 ArcMap 做出的 31 省份人均 GDP 的四分位图

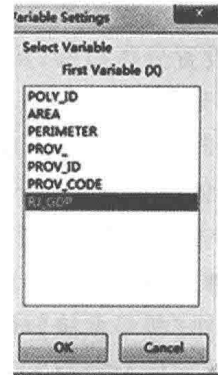


图 1.8 用 GeoDa 生成人均 GDP 的四分位图

6. 分位数分布图的 GeoDa 制作

用 GeoDa 制作 31 省份人均 GDP 的四分位图,以便与 ArcMap 中所制作的四分位图作对比。

在刚才已导入数据的前提下,点击【Map】→【Quantile Map】,在出现的列表中选择“RJ_GDP”表示做人均 GDP 的分位图,【OK】后选择所做的到底是几分位图(默认的是 4 分位图)点击【确定】,得到 GDP 的四分位图,如图 1.9 所示:

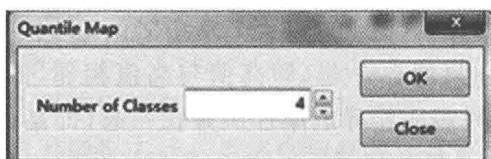


图 1.9 用 GeoDa 生成人均 GDP 的四分位图

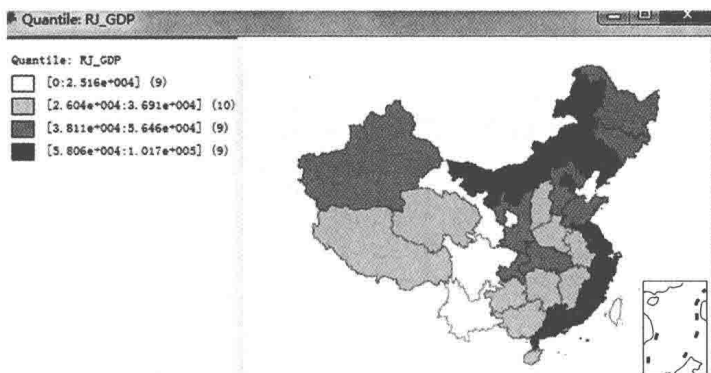


图 1.10 用 GeoDa 生成人均 GDP 的四分位图

1.1.3 空间相关性的度量

探索性空间数据分析空间相关性,主要是利用全局相关性与局部相关性。全局相关性用来分析空间经济数据在整个时空系统中表现的相关性情况;局部相关性则是分析局部区域或子系统表现出的相关性情况^①。

1. 全域空间相关性的度量

(1) Moran's I 指数

Moran's I 指数是最早应用于全局聚类检验的方法(Cliff 和 Ord, 1973)。它检验整个研究区域中邻近地区间是相似、相异(空间正相关、负相关),还是相互独立的。Moran's I 指数的计算公式如下:

$$I = \frac{n \sum_{i=1}^n \sum_{j=1}^n w_{ij} (x_i - \bar{x})}{\sum_{i=1}^n \sum_{j=1}^n w_{ij} (x_i - \bar{x})^2} = \frac{\sum_{i=1}^n \sum_{j \neq 1}^n w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{S^2 \sum_{i=1}^n \sum_{j=1}^n w_{ij}} \quad (1.1)$$

式中, n 是研究区域内地区总数, w_{ij} 是空间权重(如以区域 i 和区域 j 是否相邻设定 w_{ij} : 区域 i 和区域 j 相邻时, $w_{ij} = 1$; 区域 i 和区域 j 不相邻时, $w_{ij} = 0$); x_i 和 x_j 分

^① 本部分整理编辑自: 沈体雁等. 空间计量经济学[M]. 北京大学出版社 2010 年版.

别是区域 i 和区域 j 的属性; $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$, 是属性的平均值; $S^2 = \frac{1}{n} \sum_i (x_i - \bar{x})^2$, 属性的方差。

Moran's I 指数可以看作是观测值与它的空间滞后 (spatial lag) 之间的相关系数。变量 x_i 的空间是 x_i 在邻域 j 的平均值, 定义为:

$$x_{i,-1} = \frac{\sum_j w_{ij} x_{ij}}{\sum_j w_{ij}} \quad (1.2)$$

因此, Moran's I 指数的取值一般在 -1 到 1 之间, 大于 0 表示正相关, 值接近于 1 时表明具有相似的属性集聚在一起 (即高值与高值相邻, 低值与低值相邻); 小于 0 表示负相关, 值接近 -1 时表明相异的属性集聚在一起 (即高值与低值相邻, 低值与高值相邻)。如果 Moran's I 指数接近于 0 , 则表示属性是随机分布的, 或者不存在空间自相关性。

(2) Geary 指数 C

与 Moran's I 指数相似, Geary 指数 C 也是全局聚类检验的一个指数。计算 Moran's I 指数时, 用的是中值离差的叉乘, 但是, Geary 指数 C 强调的是观察值之间的离差, 其公式为:

$$C = \frac{(n-1) \sum_{i=1}^n \sum_{j=1}^n w_{ij} (x_i - x_j)^2}{2 \sum_{i=1}^n \sum_{j=1}^n w_{ij} \sum_{i=1}^n (x_i - \bar{x})^2} \quad (1.3)$$

Geary 指数 C 的取值一般在 0 到 2 之间 (2 不是一个严格的上界), 大于 1 表示负相关, 等于 1 表示不相关, 小于 1 表示正相关。因此, Geary 指数 C 与 Moran's I 指数正好相反。Geary 指数 C 有时也被称为 G 系数 (Getis-general G)。

2. 局部空间相关性的度量

(1) 局部 Moran's I 指数

Anselin (1995) 提出了一个局部 Moran 指数 (local Moran index), 或称 LISA (local indicator of spatial association), 用来检验局部地区是否存在相似或相异的观察值集聚在一起。区域 i 的局部 Moran 指数用来度量区域 i 和它相邻区域之间的关联程度, 定义为:

$$I_i = \frac{(x_i - \bar{x})}{S^2} \sum_{j \neq i} w_{ij} (x_j - \bar{x}) \quad (1.4)$$

正的 I_i 表示一个高值被高值包围, 或者一个低值被低值所包围。负的 I_i 表示一个低值被高值包围, 或者一个高值被低值所包围。

(2) 局部 Geary 指数

Getis 和 Ord (1992) 开发了一个 Geary 指数的局部聚类检验, 称之为 G_i 指数, 用来检验局部地区是否存在统计显著的高值或低值。 G_i 指数的定义如下:

$$G_i = \frac{\sum_{j \neq i} w_{ij} x_j}{\sum_{j \neq i} x_j} \quad (1.5)$$

这个指数用来检验局部地区是否有高值或低值在空间上趋于集聚。高的 G_i 值表