

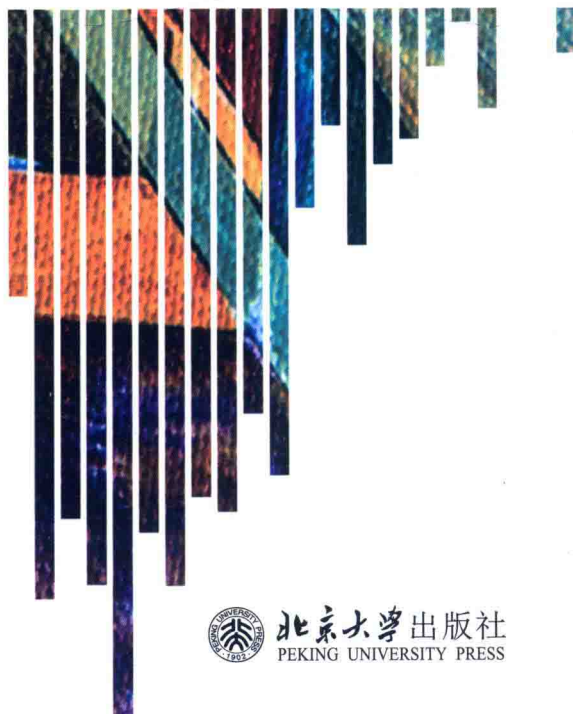
SOFTWARE FOR SPATIAL ECONOMETRICAL ANALYSIS

Operation Manual of GeoDa、GeoDaSpace and PySAL

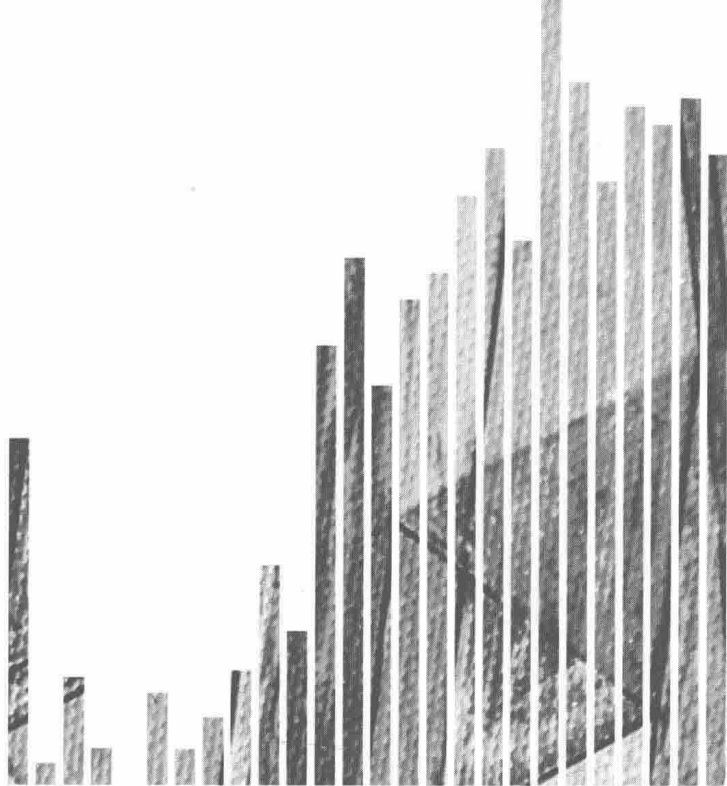
空间计量分析软件

GeoDa、GeoDaSpace 和 PySAL 操作手册

沈体雁 于瀚辰 曹巍韩 何泓浩 / 编著



北京大学出版社
PEKING UNIVERSITY PRESS



SOFTWARE FOR SPATIAL ECONOMETRICAL ANALYSIS

Operation Manual of GeoDa、 GeoDaSpace and PySAL

空间计量分析软件

GeoDa、 GeoDaSpace 和 PySAL 操作手册

沈体雁 于瀚辰 曹巍韡 何泓浩 / 编著

常州大学图书馆
藏书章



北京大学出版社
PEKING UNIVERSITY PRESS

图书在版编目(CIP)数据

空间计量分析软件：GeoDa、GeoDaSpace 和 PySAL 操作手册 / 沈体雁等编著. —北京：北京大学出版社，2019.11

ISBN 978-7-301-30456-3

I. ①空… II. ①沈… III. ①区位经济学—计量经济学—应用软件—研究生—教材
IV. ①F224.0-39

中国版本图书馆 CIP 数据核字(2019)第 084301 号

- 书 名** 空间计量分析软件：GeoDa、GeoDaSpace 和 PySAL 操作手册
KONGJIAN JILIANG FENXI RUANJIAN:GeoDa、GeoDaSpace HE PySAL
CAOZUO SHOUCHE
- 著作责任者** 沈体雁 于瀚辰 曹巍韩 何泓浩 编著
- 责任编辑** 任京雪 徐 冰
- 标准书号** ISBN 978-7-301-30456-3
- 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>
- 电子信箱** em@pup.cn QQ:552063295
- 新浪微博** @北京大学出版社 @北京大学出版社经管图书
- 电 话** 邮购部 010-62752015 发行部 010-62750672 编辑部 010-62752926
- 印 刷 者** 河北滦县鑫华书刊印刷厂
- 经 销 者** 新华书店
- 787 毫米×1092 毫米 16 开本 18 印张 257 千字
2019 年 11 月第 1 版 2019 年 11 月第 1 次印刷
- 定 价** 39.00 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

序

自2010年我们编著出版《空间计量经济学》(第一版)以来,在国内外学者的共同推动下,空间计量经济学在中国的发展取得了长足的进步。这主要体现在三个方面:

第一,中国空间计量经济学的教学体系初步形成。“空间计量经济学”作为一门正式的本科生或研究生课程,被列入经济学、地理学、管理学、规划学等专业的课程体系和学科建设计划。据不完全统计,北京大学、中国社会科学院大学、中国人民大学、武汉大学、华东理工大学、上海师范大学、暨南大学、山东大学、郑州大学、厦门大学、中南财经政法大学、江西财经大学、青岛科技大学等高校已经开设“空间计量经济学”课程。与此同时,空间计量经济学暑期学校、专题讲座和各类培训如火如荼地开展,已成燎原之势。2016—2019年,北京大学与中国区域科学协会等单位连续举办了四届北京大学研究生暑期学校“空间计量经济学前沿”活动,空间计量经济学的创始人之一、美国国家科学院院士、美国芝加哥大学卢克·安索林(Luc Anselin)教授,地理加权回归模型(GWR)的提出者、美国国家科学院院士和英国社会科学院院士、美国亚利桑那州立大学斯提沃特·福瑟林汉姆(Stewart Fotheringham)教授以及国内外著名空间计量经济学家、学者在暑期学校授课,共有1000余名来自国内外各高等院校的青年学者和学生参加了学习和研讨,极大地推动了空间计量经济学在中国的发展。伴随人才培养和教学工作的需要,国内已有十余本空间计量经济学方面的教材和专著(包括译著)相继出版;部分大学的经济学、商学、公共管理学、地理学、规划学等院系已开始招收和培养空间计量经济学方向的硕士和博士研究生;空间计量经济学师资队伍建设也受到有关院系领导的重视,能够开设“空间计量经济学”课程的教师十分紧俏。可以说,到目前为止,中国已经初步形成比较系统的本土化的空间计量经济学教学体系。

第二,中国空间计量经济学理论、方法与应用研究成果丰硕,进步喜人。一方面,空间计量经济学在一些前沿理论与方法的研究上取得了若干标志性成果。北京大学虞吉海教授在面板数据最大似然估计的一致性证明、动态面板模型的拟最大似然估计、空间单位根与伪回归等领域取得了重要突破,香港中文大学黄波教授的团队提出

了时空地理加权回归模型(GTWR),我们团队的于瀚辰博士采用广义加性模型等技术解决了多尺度地理加权回归(MGWR)的统计推断问题。这些研究成果进一步拓展了空间回归和地理加权回归等主流空间计量经济学模型的设定、估计、检验和统计推断,完善了空间计量经济学理论与方法。另一方面,空间计量经济学在人文社会科学领域的应用广度与深度不断拓展,推动了人文社会科学的空间化、定量化和综合化,成为空间综合人文学与社会科学发展的一个重要学科基础。应用空间计量经济学理论与方法分析和解决中国城市与区域发展的现实问题,提出和验证中国城市与区域发展的特征事实、理论范式与科学规律,是中国空间计量经济学的主要任务。近年来,空间计量经济学在各个领域的应用研究及发表的学术论文数量呈几何级数式增长,空间计量分析与地理信息系统正在成为对城市与区域问题展开实证研究的两项必备技能。总体上,中国空间计量经济学领域的科学研究呈现出以应用研究为主、理论与应用研究并进,论文发表数量剧增、学术研究质量不断提升,部分领域进入国际前沿、整体研究水平尚需进一步提高的发展局面。

第三,中国空间计量经济学研究社区已经形成。学术社区的形成是学科成熟的重要标志。过去十年,中国空间计量经济学领域的学术交流和国际合作日渐繁荣,国内学者、学生与国外的学术联系日渐紧密,一些青年学者和学生通过留学、访学和暑期学校等多种形式深度参与国际顶尖学者的研究工作,逐渐步入空间计量经济学理论与方法研究的前沿阵地。中国区域科学协会、中国地理学会等学术组织成立了空间计量经济学方面的专业委员会,推动空间计量经济学在中国的发展,广大学者和学生通过各种“线上”“线下”的学术交流活动日益密切地组织起来,有关空间计量经济学的各种公益性甚至商业性的讲座、论坛、培训等活动在城市与区域研究这个“小花园”里如雨后春笋般层出不穷,分外耀眼。可以非常欣慰地讲,在包括安索林院士和福瑟林汉姆院士在内的国际著名空间计量经济学家的持续支持下,通过包括北京大学在内的国内众多高等院校、科研机构、学术团体和广大学者、学生的共同努力,中国空间计量经济学的发展经历了从“星星之火”到“燎原之势”的“黄金十年”。可以预期,随着中国科技创新从以“跟跑”为主逐步向更多领域的“并跑”“领跑”转变,未来十年中国空间计量经济学也将从国际空间计量经济学“俱乐部”的“跟跑者”发展为“并跑者”,并向局部领域的“领跑者”迈进。

为了尽可能地反映和吸纳空间计量经济学理论与方法的新进展,回应广大读者系统学习空间计量经济学知识与技能的需求,在北京大学出版社的支持下,我们决定对

《空间计量经济学》(第一版)进行大幅修改与调整,重新编写和出版《空间计量经济学》(第二版)。与此同时,为了便于读者在学习和理解第二版中出现的诸多概念术语、理论模型、推导解析、数学公式的过程中,也能够学会各种主流空间计量经济学软件的操作和使用方法,掌握开展空间计量分析和实证研究的技能,我们同步编写了《空间计量分析软件:GeoDa、GeoDaSpace 和 PySAL 操作手册》《空间计量分析软件:R 语言操作手册》《空间计量分析软件:GWR 和 MGWR 操作手册》三本实用性、操作性、手册性导向的工具书,并配有相应的课件和教辅材料,供师生们教学所用。

上述一本教科书、三本软件操作手册的“1+3”教材体系是北京大学空间计量经济学研究团队过去十余年教学科研经验的总结,也是对长期关心与支持我们工作的各位老师、同学的回应和回报。这套教材可以用作“空间计量经济学”“空间数据科学”“地理信息系统”“空间分析与模拟”“地统计学”等课程的教材,也可以作为经济学、管理学、地理学、规划学、人口学、城市与区域科学等领域的教师、学生与实际工作者学习和研究空间计量经济学的参考书,更是在时空大数据时代普及和推广空间计量分析的科普书。

本书是以安索林院士撰写的 *Modern Spatial Econometrics in Practice: A Guide to GeoDa, GeoDaSpace and PySAL* 一书为基础,通过搜集、整理、编译和中国化安索林院士团队所发布的有关 GeoDa 软件的各种资料编写而成的。整个编写工作由安索林院士团队的核心成员——美国芝加哥大学空间数据科学研究中心的李迅博士指导,由北京大学研究团队的于瀚辰、曹巍韡、何泓浩、吕永强、谭丽等同学历时四年共同完成,我负责对全书进行谋划组织、审阅修改和校对,并撰写序和导论等内容。GeoDa 软件家族包括基于图形用户界面的空间计量分析软件 GeoDa 和 GeoDaSpace,以及基于 Python 语言的空间计量分析开源类库——PySAL 等软件,是目前世界上最流行的空间计量分析软件之一,截至 2018 年年底,已在全球 50 多个国家拥有超过 27 万个用户。随着安索林院士团队将 GeoDa 软件进一步模块化、互联网化、移动化和大数据化的研发计划的实施,可以预期,GeoDa 软件家族的用户群体将进一步扩大,软件的使用场景将进一步丰富,软件的运行效率和用户体验将进一步改进,基于核心模块的各种增值软件将不断涌现,因此 GeoDa 软件家族无疑是当前最具有发展潜力和最值得读者学习的空间计量经济学软件之一。本书采用案例研究、软件截图、操作说明等形式,为 GeoDa 软件学习者提供简单明了、直观易学的操作指南,上篇介绍了 GeoDa 和 GeoDaSpace 的地理可视化、探索性空间数据分析、空间权重的创建与操作、空间滞后模型

和空间误差模型的建立与估计等空间计量分析功能,下篇介绍了 PySAL 的数据读入和导出、空间权重矩阵的创建和操作、空间数据分析、各种空间回归模型的建立与估计、空间动力学分析等功能。

在本书即将出版之时,我要衷心地感谢两位空间计量经济学的重要创始人——安索林院士和福瑟林汉姆院士,感谢他们在过去四年中持续不断地指导和支持北京大学空间计量经济学研究团队的工作,使我们的教学、科研和人才培养能够迅速地对接国际前沿,走入理论、方法与应用研究良性循环的道路,成为空间计量经济学在中国发展的重要枢纽和重镇。感谢北京大学研究生暑期学校“空间计量经济学前沿”项目的支持单位和主办单位——北京大学研究生院、中国区域科学协会、中国地理学会、中国地理信息系统产业协会等单位,以及过去四年持续支持我们暑期学校的授课老师们、参加或推荐人员参加我们暑期学校的来自全国各高校的老师 and 同学们。大家共同的努力使得我们的暑期学校已经成为空间计量经济学学习和交流的重要平台,成为与大师“亲密接触”和“面对面”交流的“人性化”场所,成为激发学习兴趣、启发创新智慧、鼓励年轻人走向空间计量经济学研究生涯的灵动空间。感谢北京大学政府管理学院的领导和老师们,特别是我所在的城市与区域管理系的杨开忠教授、李国平教授、陆军教授、薛领教授、张波副教授、孙铁山副教授和刘伦助理教授。感谢北京大学出版社的林君秀主任、刘京编辑和任京雪编辑,没有她们的鼓励和敦促,本书是难以如期出版的。

我们深知空间计量经济学理论、方法与应用背后隐藏着深刻的关于“空间”的哲学命题、科学道理和技术创新“奇点”,揭示“空间之谜”仍将是一条漫长的知识探索道路。本书难免存在许多谬误之处,恳请读者在阅读学习过程中给予批评指正,我们将在今后的工作中进行修正。希望中国空间计量经济学迎来与国际同行“并跑”的新的“黄金十年”!

沈体雁

2019年10月18日

导 论

空间计量经济学是对空间经济系统进行计量分析的经济学科;是以区域科学和空间经济理论为基础,以计量分析和地理计算为手段,以处理空间数据所特有的“空间效应”为核心,研究区域科学或空间经济模型的设定、估计、假设检验、预测和应用的理论与方法体系;是区域科学、计量经济学和地理计算科学的交叉学科。自 Paelinck and Klaassen(1979)首次提出和创立空间计量经济学以来,经过四十年年的发展,空间计量经济学已经成为一门具有比较完善的理论、方法与工具的新兴学科,成为从事空间政治经济分析和空间人文社会科学研究的重要基础学科。

随着现代信息技术的不断进步,作为理论与方法的显性化载体与支撑性平台,特定学科领域软件工具的研究开发已经成为学科进步的重要驱动力和核心标志。与空间计量理论与方法的进步相辅相成,空间计量分析软件的研发取得了长足的进步。到目前为止,空间计量经济学领域已经出现的主流的空间计量分析软件包括:GeoDa 软件家族、R 语言、Metalab、GWR 和 ARCGIS 等。本书旨在采用具体案例和操作说明等形式,具体而翔实地介绍由美国芝加哥大学卢克·安索林院士领衔研究开发的 GeoDa 空间计量分析软件家族,包括 GeoDa、GeoDaSpace 和 PySAL 三个软件包。全书共 22 章,分为上下两篇。其中,上篇 11 章,即第 1 章至 11 章,介绍基于图形用户界面的空间计量分析软件:GeoDa 和 GeoDaSpace;下篇 11 章,即第 12 章至 22 章,介绍基于 Python 的空间计量分析库:PySAL。导论部分,主要介绍三个方面的内容,包括安索林院士及其 GeoDa 软件家族的基本情况、本书案例数据集说明和有关术语说明。

1. GeoDa 软件家族

GeoDa 软件家族由安索林院士及其团队在美国国家自然科学基金、美国空间社会科学研究中心和美国加州大学圣芭芭拉分校、亚利桑那州立大学、芝加哥大学等高校的支持下研究开发和发布使用的空间计量分析软件工具。2002 年,安索林院士团队发布第一个版本的 GeoDa 软件,截至 2018 年 9 月,该软件在全球 50 多个国家拥有 27 万多

个用户,成为最流行的空间计量软件之一。本书介绍的软件版本是 GeoDa1.12.1.161。

GeoDa 软件的主创人安索林院士目前是美国芝加哥大学社会学部教授、空间数据科学研究中心主任,美国国家科学院院士、美国艺术与科学院院士,地理信息科学专业大学联盟会士,Walter Isard 主席。作为世界著名的空间计量和空间分析专家,以及空间计量经济学的重要奠基人之一,安索林院士长期从事空间计量经济学理论和方法的教学、研究和社会服务工作,在空间数据的探索性分析、可视化、建模及其开源软件开发等领域进行了开创性工作。他将先进的空间经济分析理论与方法应用于城市与区域、环境与自然资源经济学、房地产经济学、创新经济学、犯罪学、公共卫生、选举、国际关系等空间社会科学研究领域,取得了丰富的创新性成果,产生了广泛的影响,成为横跨信息科学技术(特别是地理信息科学技术和时空大数据科学)和人文社会科学两大领域的世界顶尖学者,由此获得了 Walter Isard 奖(2005)和 William Alonso 纪念奖(2006)等众多奖项。

迄今为止,安索林院士已经发表 150 余篇学术论文,出版了多本专著和文集。其中,1988 年出版的专著 *Spatial Econometrics: Methods and Models* 已成为在该领域的经典著作,被引用次数超过 9 000 次。他也被世界权威规划杂志和 Google 学术多次评为被引用次数最多的规划学者,总被引用次数接近 50 000 次。通过这些工作,他将早期尚处于边缘地位并且缺乏空间分析软件的空间计量经济学方法论推向了计量经济学的主流,同时也为地理信息科学奠定了坚实的计量经济学基础,为解决经济学“最后的前沿问题”——将“空间”引入主流经济学和推动包括经济、社会和政治等因素在内的更加一般性的区位理论做出了贡献。

在本书所介绍的 GeoDa 软件家族中,GeoDa 和 GeoDaSpace 是基于图形用户界面的空间计量分析软件。其中,GeoDa 是最早的也是最知名的软件包,于 2002 年发布第一版,2003 年发布了一个改进版,即版本 0.9.5-i。2009 年,第一个开源测试版 GeoDa 发布,现在也被称为遗留版 GeoDa,是用 C++ 编程语言,基于美国环境系统研究所(ESRI)的 MapObjects 图形库编写而成的,是只能在 Windows XP 操作系统上运行的桌面软件系统。2011 年,安索林院士团队发布了官方版本基于图形用户界面的 GeoDa 1.0。该版本采用开源的 GPL3.0 许可,支持 Windows、Mac OS X 和 Linux 三种操作系统,而且在每一个操作系统的系统外观和用户体验上都有所不同。GeoDa 具有强大的地理可视化(包括异常值图、速率平滑、统计图和地图动画等),探索性数据分析(包括统计图表、平行坐标图、条件图),特别是探索性空间数据分析(ESDA),创建和操作

空间权重(如邻近权重,距离权重, k 最近邻权重),建立和估计各种空间回归模型,以及构建空间数据集(如点坐标 shape 文件、泰森多边形、质心)等空间计量分析功能,其发布极大地降低了空间计量分析的学习成本,提升了用户体验,推动了空间计量分析的普及和使用。

GeoDaSpace 是一个包含在 PySAL spreg 模块中的有关空间回归代码的图形用户界面软件,其图形是基于 wxWidgets C++ 库的一个 Python 包——wxPython 模块而开发的,它适用于 Windows 和 Mac OS X 操作系统。GeoDaSpace 能够为用户提供先进的空间计量分析功能,包括创建和设置空间权重,以及主流空间计量模型的估计方法和统计检验。

PySAL 是一个空间计量分析的 Python 库,是用 Python 语言开发的用于进行空间计量分析的开源跨平台库。由于该库不是基于图形用户界面的,因此只能基于命令行对它进行操作,相应地,其用户也主要是那些习惯于使用命令行界面的使用者和程序员。该库包含了读入及导出数据、创建和操作空间权重矩阵、空间数据分析、空间回归分析和空间动力学分析等功能。要完成上述功能,我们主要需要用到 PySAL 的 weights 和 spreg 两个模块。

GeoDa 软件的安装包可以从美国芝加哥大学空间数据科学研究中心网站的软件下载页面下载,也可以从 github(开源和私有软件项目托管平台)网站下载。其网址分别是:

- (1) <https://spatial.uchicago.edu/software>
- (2) <https://geodacenter.github.io/>

2. 案例数据集说明

安索林院士团队提供了数十项可用于检验 GeoDa 软件家族可行性的案例数据集。这些数据集可以从美国芝加哥大学空间数据科学研究中心网站的数据下载页面下载,或者从 github 网站下载。其网址分别是:

- (1) <https://spatial.uchicago.edu/>
- (2) <https://geodacenter.github.io/data-and-lab>

本书用来进行操作示范的数据集主要有两个,即美国马里兰州巴尔的摩市的房价数据集和美国县级犯罪数据集。分别说明如下:

巴尔的摩市房价数据集：

巴尔的摩市房价数据集是一个记录了1978年美国马里兰州巴尔的摩市住房销售价格和房屋特征的数据库，存储于 baltim.dbf 文件之中，用于对影响住房价格的各种因素进行空间享乐分析。其原始数据来自美国俄亥俄州克利夫兰凯斯西部研究大学管理学院的罗宾·A.杜宾(Robin A. Dubin)教授。其主要特征如下：

- (1) 文件类型：点形状文件；
- (2) 投影：基于马里兰网格(Maryland grid)的投影坐标，具体投影类型未知；
- (3) 观察值：211个；
- (4) 变量数：17个；
- (5) 年份：1978年。

主要变量描述如下：

表 1 巴尔的摩市房价数据集变量说明

变量	说明
STATION	唯一标识码 ID variable
PRICE	房屋销售总价(单位为千美元) sales price of house in \$ 1 000 (MLS)
NROOM	所含房间数 number of rooms
DWELL	所含卧室数(1表示该房屋有独立卧室) 1 if detached unit, 0 otherwise
NBATH	所含浴室数 number of bathrooms
PATIO	露台的指示变量(1表示该房屋有露台) 1 if patio, 0 otherwise
FIREPL	壁炉的指示变量(1表示该房屋有壁炉) 1 if fireplace, 0 otherwise
AC	空调的指示变量(1表示该房屋有空调) 1 if air conditioning, 0 otherwise
BMENT	地下室的指示变量(1表示该房屋有地下室) 1 if basement, 0 otherwise
NSTOR	房屋层数 number of storeys
GAR	车库的车位数(0表示没有车位) number of car spaces in garage (0 = no garage)

续表

变量	说明
AGE	房产的年龄 age of dwelling in years
CITCOU	地区变量(1 表示该房屋位于巴尔的摩县) 1 if dwelling is in Baltimore County, 0 otherwise
LOTSZ	建筑面积(单位为百平方英尺) lot size in hundreds of square feet
SQFT	室内面积(单位为百平方英尺) interior living space in hundreds of square feet
X	X 坐标 x coordinate on the Maryland grid
Y	Y 坐标 y coordinate on the Maryland grid

美国县级犯罪数据集：

美国县级犯罪数据集是一套记录了美国大陆县(不包括阿拉斯加州和夏威夷州)的犯罪数据,包括 1960 年、1970 年、1980 年和 1990 年四个人口普查年份的凶杀案以及有关的社会经济特征数据^①。其基本特征如下：

- (1) 文件类型:面形状文件；
- (2) 观察值:3 085 个；
- (3) 变量数:73 个；
- (4) 投影:未知投影,经纬度格式。

该数据集以三个不同的文件进行存储,即 NAT.dbf, natregimes.dbf 和 south.dbf。其中,前两者所包含的是所有大陆县的犯罪数据,而后者只包含了美国南部地区的犯罪数据。我们用这套数据来研究美国各县的犯罪率与其潜在社会经济影响因素之间的相互关系。其主要变量描述如下：

表 2 美国县级犯罪数据集变量说明

变量	说明
NAME	县名 county name

^① 详见: Messner, S., Anselin, L., Hawkins, D., Deane, G., Tolnay, S., & Baller R. (2000). An Atlas of the Spatial Patterning of County-Level Homicide, 1960-1990. Pittsburgh: National Consortium on Violence Research (NCOVR).

变量	说明
STATE_NAME	州名 state name
STATE_FIPS ^①	州 FIPS 代码(字符型) statefips code (character)
CNTY_FIPS	县 FIPS 代码(字符型) countyfips code (character)
FIPS	州和县组合 FIPS 代码(字符型) combined state and countyfips code (character)
STFIPS	州 FIPS 代码(数字型) statefips code (numeric)
COFIPS	县 FIPS 代码(数字型) countyfips code (numeric)
FIPSNO	州和县组合 FIPS 代码(数字型) fips code as numeric variable
SOUTH	南方县哑元变量(南方县取值为 1) dummy variable for Southern counties (South = 1)
HR**	每十万人犯罪率(四个年份) homicide rate per 100 000 (1960, 1970, 1980, 1990)
HC**	杀人案件数(以四个年份为基础取前后三年的平均值) homicide count, three year average centered on 1960, 1970, 1980, 1990
PO**	县域人口数量(四个年份) county population, 1960, 1970, 1980, 1990
RD**	资源匮乏(主成分,四个年份) resource deprivation, 1960, 1970, 1980, 1990 (principal component)
PS**	人口结构(主成分,四个年份) population structure, 1960, 1970, 1980, 1990 (principal component)
UE**	失业率(四个年份) unemployment rate, 1960, 1970, 1980, 1990
DV**	14 岁及以上男性离婚百分比(四个年份) divorce rate, 1960, 1970, 1980, 1990 (% males over 14 divorced)
MA**	年龄中位数(四个年份) median age, 1960, 1970, 1980, 1990
POL**	人口的对数值(四个年份) log of population, 1960, 1970, 1980, 1990
DNL**	人口密度的对数值(四个年份) log of population density, 1960, 1970, 1980, 1990

① FIPS,即美国联邦信息处理标准(the Federal Information Processing Standards)。

变量	说明
MFIL**	平均家庭收入的对数值(四个年份) log of median family income, 1960, 1970, 1980, 1990
FP**	贫困家庭百分比(四个年份) % families below poverty, 1960, 1970, 1980, 1990 (see Codebook for details)
BLK**	黑人占比(四个年份) % black, 1960, 1970, 1980, 1990
GI**	家庭收入不平等的基尼系数(四个年份) Gini index of family income inequality, 1960, 1970, 1980, 1990
FH**	女性户主家庭的百分比(四个年份) % female headed households, 1960, 1970, 1980, 1990
WEST	西部地区虚拟变量 west region dummy
REGIONS	地区变量。南部地区的该变量值为 0,北部和中部地区的该变量值为 1,西部地区该变量值为 2. region dummy

注:**指有四个年份的数据。

上述变量说明中没有时间变量说明,事实上所有变量都和时间有关,在本书所有的案例演示中,我们把时间信息放在变量名中加以体现,如 1960 年的犯罪率变量名为 HR60,1970 年的犯罪率变量名为 HR70,如此类推。部分案例出现的 REGIONS 为地区变量,即表示某个县所在的区域的虚拟变量,南部地区的该变量取值为 0,北部和中部地区的该变量取值为 1,而西部地区的该变量则取值为 2。

本书所有案例中所涉及的上述两个案例数据集及其变量,都采取以上的描述,后文不再一一进行说明。

3. 有关术语说明

从简化书写以及方便阅读的目的出发,本书使用了很多计量经济学,特别是空间计量经济学的术语及其简称。为了便于读者将本书与沈体雁教授编著的《空间计量经济学》(第二版)^①整合阅读,在此将本书中出现的各种术语简称以及空间计量经济学常用的术语简称及其中英文对照进行列举,后文不再一一说明。

CLR,经典线性模型(Classical Linear Regression Model)

^① 详见:沈体雁,于瀚辰.空间计量经济学(第二版).北京:北京大学出版社,2019.

FAR, 单变量一阶空间自回归模型 (Simultaneous Univariate First-Order Spatial Autoregression Model)

SEM, 空间误差模型 (Spatial Error Model)

SLM/SAR, 空间滞后模型 (Spatial Lag Model/ Spatial Autoregressive Model)

SLX, 自变量空间滞后模型 (Spatial Lag of X Model)

SDM, 空间杜宾模型 (Spatial Durbin Model)

SDEM, 空间杜宾误差模型 (Spatial Durbin Error Model)

SAC, 广义空间模型 (General Spatial Autocorrelation Model)

GNS, 广义嵌套空间模型 (General Nesting Spatial Model)

SMA, 空间移动平均模型 (Spatial Moving Average Model)

SEC, 空间误差分量模型 (Spatial Error Components Model)

SSUR, 空间似无关回归模型 (Spatial Seemingly Unrelated Regression Model)

OLS, 最小二乘估计 (Ordinary Least Squares)

2SLS, 两阶段最小二乘法 (Two Stage Least Squares)

GLS, 广义最小二乘 (Generalized Least Squares)

WLS, 加权最小二乘 (Weighted Least Squares)

IV, 工具变量法 (Instrumental Variable Estimation)

MLE, 最大似然估计 (Maximum Likelihood Estimation)

GM, 矩估计方法 (Method of Moments)

GMM, 广义矩估计方法 (Generalized Method of Moments)

HAC, 异方差和自相关一致估计方法 (Heteroscedastic and Autocorrelation Consistent)

S2SLS, 空间两阶段最小二乘法 (Spatial Two Stage Least Squares)

FGS2SLS, 可行广义空间两阶段最小二乘估计 (Feasible Generalized Spatial Two-Stage Least Squares)

BLUE, 最小方差的线性无偏估计量 (Best Linear Unbiased Estimator)

LM, 拉格朗日乘子检验 (Lagrange Multiplier Test)

RLM, 稳健的拉格朗日乘子检验 (Robust Lagrange Multiplier Test)

LR, 似然比检验 (Likelihood Ratio Test)

MCMC, 马尔可夫链蒙特卡洛方法 (Markov Chain Monte Carlo Method)

ESF,特征向量空间滤波法(Eigenvector Spatial Filter)

EDM,特征函数分解法(Eigenfunction Decomposition Method)

GWR,地理加权回归(Geographically Weighted Regression)

SGWR,半参数地理加权回归(Semi-Parameter Geographically Weighted Regression/Mix Geographically Weighted Regression)

MGWR,多尺度地理加权回归(Muti-Scale Geographically Weighted Regression)

AIC,赤池信息量准则(Akaike Information Criterion)

AICc,修正的赤池信息量准则(Corrected Akaike Information Criterion)

GIS,地理信息系统(Geographic Information System)

R&D,研究与开发(Research and Development)

MAR,马歇尔-阿罗-罗默外部性(Marshall-Arrow-Romer Externalities)

LISA,局部空间关联指标(Local Indicators of Spatial Association)

NEG,新经济地理学(New Economic Geography)

EDA,探索性数据分析(Exploratory Data Analysis)

ESDA,探索性空间数据分析(Exploratory Spatial Data Analysis)

SP,空间过程(Spatial Processes/ Spatial Stochastic Processes)

MAUP,可变面元问题(Modifiable Areal Unit Problem)

GWPR,地理加权泊松回归(Geographical Weighted Poisson Regression)

GWLR,地理加权 Logistic 回归(Geographical Weighted Logistic Regression)

AGWR,地理加权空间自回归(Autoregressive Geographical Weighted Regression)

PGW,地理加权 Probit 回归(Probit Model with Geographical Weighted)

GWRR,地理加权岭回归(Geographical Weighted Ridge Regression)

GWL,地理加权套索回归(Geographical Weighted Lasso)

SWIM,空间加权相互作用模型(Spatially Weighted Interaction Models)

GTWR,时空地理加权回归(Geographical and Temporal Weighted Regression)

GAM,广义加性模型(Generalized Additive Model)

ENP,有效参数数量(Effective Number of Parameters)

ICT,信息、通信和技术(Information,Communication and Technology)

BIC,贝叶斯信息准则(Bayesian Information Criterion)

HQC, 汉南-奎因准则(Hannan-Quinn Criterion)
空间计量经济学(Spatial Econometrics)
空间效应(Spatial Effect)
空间溢出效应(Spatial Spillover Effects)
计量经济学(Econometrics)
经典计量经济学(Classical Econometrics)
高斯-马尔科夫假定(Gauss-Markov Assumptions)
非经典计量经济学(Non-Classical Econometrics)
现代时间序列计量经济学(Time Series Econometrics)
微观计量经济学(Micro-Econometrics)
非参数计量经济学(Nonparametric Econometrics)
面板数据计量经济学(Panel Data Econometrics)
国际区域科学协会(Regional Science Association International)
《区域科学杂志》(*Journal of Regional Science*)
空间统计学(Spatial Statistics)
计算科学(Computational Science)
联立方程(Simultaneous Equation)
离散选择(Discrete Choice)
截面数据(Cross-Sectional Data)
时间序列数据(Time-Series Data)
面板数据(Panel Data)
截断数据(Truncation Data)
归并数据(Censored Data)
持续时间数据(Duration Data)
离散选择数据(Discrete Choice Data)
计数数据(Count Data)
空间数据(Spatial Data)
空间计量经济学会(Spatial Econometrics Association)
美国国家地理信息与分析中心(National Center for Geographic Information and Analysis)