

基于GIS的 数量方法与应用

〔美〕王法辉 著
姜世国 滕骏华 译



商務印書館
The Commercial Press

首都师范大学“211 工程”、地理科学与技术国家级实验
教学示范中心、三维信息获取与应用教育部重点实验室资助出版

基于 GIS 的数量方法与应用

〔美〕王法辉 著

姜世国 滕骏华 译

商务印书馆

2009 年·北京

图书在版编目(CIP)数据

基于 GIS 的数量方法与应用/(美)王法辉著;姜世国,滕骏华
译. —北京:商务印书馆,2009

ISBN 978-7-100-06092-9

I. 基… II. ①王…②姜…③滕… III. 地理信息系统-应
用-社会科学 IV. C39

中国版本图书馆 CIP 数据核字(2008)第 161590 号

所有权利保留。

未经许可,不得以任何方式使用。

基于 GIS 的数量方法与应用

[美] 王法辉 著

姜世国 滕骏华 译

商务印书馆出版

(北京王府井大街36号 邮政编码 100710)

商务印书馆发行

北京瑞古冠中印刷厂印刷

ISBN 978-7-100-06092-9

2009年6月第1版

开本 787×960 1/16

2009年6月北京第1次印刷

印张 21¼

定价: 45.00 元

Fahui Wang

Quantitative Methods and Applications in GIS

Taylor & Francis Group 2006

(根据泰勒-弗朗西斯出版集团 2006 年版译出)

献给齐磊和我们的孩子(梦妮、成齐和安妮)
并以此纪念我们永世难忘的真妮

序 言

迈克·巴迪(Michael Batty)

伦敦大学学院高等空间分析中心主任

这是一本优秀的著作。贯穿全书的一条主线就是:要搞好政策性较强的社会科学研究,数量方法是必不可少的,而这些方法及背后的理论一定要空间化。作者证明,计量地理方法要应用得有实效,当前的新潮流就是与地理信息系统(GIS)和地理信息科学(GISc)结合起来。本书的重点是计量方法在城市及区域研究中的应用。全书的一个主要优点是,所有介绍的方法都是基于实际的应用问题,并以案例的形式加以演示,从而使各种层次的读者都可以通过学习设计解答自己的研究问题。

在过去十年中,GIS技术日渐成熟,并与空间分析和计量地理的结合与共同发展,创生了地理信息科学(GISc)。地理信息科学不仅仅是关于方法或技术的科学,它也跟地理学中的各种理论密切相关,而这些理论一般是从自然科学和社会科学借鉴过来的。本书主要介绍了社会、经济研究中的应用问题,但所用方法并不局限于社会科学领域,只要涉及地理空间(如拓扑空间、欧式空间、各种维数的空间),空间分析方法都得到了发展。书中介绍的几种应用于社会科学的方法,最初就是起源于物理学等自然科学,如地球物理学、医学、生态学等领域。本书整合的计量社会科学的各种方法,是理解社会发展及社会政策的基本工具。

本书介绍的几个主题都围绕空间这个核心概念,特别是距离、邻接性、可达性等概念对于定义空间集聚和扩展过程中的聚类、密度、均质及腹地来说十分重要。这些都有利于揭示不同尺度上城市和区域系统的形态和功能,以此为基础发展的方法和技术可用于测算、分析区域中城市的自然和社会形态。这正是发展中的地理信息科学。通观全书,作者强调,

2 基于 GIS 的数量方法与应用

那些乍看起来完全不同的形态和功能,其实可以统一到更一般的系统和模型中。作者开发的应用研究范围很广,包括从健康服务与犯罪到交通与零售等多个主题,涉及城市的多个部门和多种空间尺度。其核心目标并不仅仅限于简单的数据测量和概念理解,所有问题都有其公共政策方面的应用价值。特别是最后几章,探讨了各种空间问题尤其是交通规划中的空间优化方法。

事实上,本书的一个重要目标就是介绍那些基本的政策性问题,不管是寻找商业中心的最佳区位,还是发现疾病的集聚中心,都是通过空间分析来有效解决的。这些问题越来越适宜于用空间分析的方法来解决,是与高质量、多变量、小尺度数据的日益增加、空间数据整合手段和显示技术的不断发展密切相关的。这正是 GIS 对社会科学研究的贡献,也是定量分析方法在社会科学研究中的应用日益空间化的主要原因。现在,研究疾病分布模式及其公共卫生的对策方案很难不用到空间数据。在我们这个资源短缺的时代,随着新技术的发展、个人机会的增加,社会越来越复杂,为了寻求问题的最优解决方案,空间分析方法已经不可或缺,这也是本书强调的一个主题。读者可能发现,可靠的空間分析方法常常让那些用传统方法很难解决的问题迎刃而解。犯罪分析是一个很好的例子,正如作者在书中演示的那样,那些原来从土壤科学和地质学发展出来的聚类方法,可以很好地用于犯罪分析。聚类分析除了用于分析犯罪热点地区之外,还可以用于识别疾病集聚区,这是政策研究中十分重要的问题。作者通过案例有效地证明,一些常规方法可以通过适当改变来解决广泛的政策问题。

书中多处介绍了一些复杂、完整的模型,如从零售业服务区模型到可达性分析、从人口密度方程的拟合到全书最后的土地利用与交通网络互动模型,这是很值得称许的。由此证明,地理信息科学不是单纯的技术工具箱,它包括越来越一般性的系统问题,本书介绍的方法及其应用触及到这些系统问题的本质。虽然全书主要集中于与政策有关的空间分析,但是揭开了城市与区域系统更广范围复杂性的序幕,而这正是地理信息科学亟待解决的问题。

前 言

近年来,社会科学研究(包括应用社会科学研究和公共政策研究)的一项重要进展,就是定量或计算方法在研究复杂的人类社会系统中的应用。美国许多大学都设有相关的计量社会科学研究中心,其中,芝加哥大学、华盛顿大学、加州大学洛杉矶分校、乔治·梅森(George Mason)大学近年来兴办的这类中心,在提升社会科学交叉领域研究的贡献尤其突出。围绕这一主题的相关学术会议也层出不穷,地理信息系统(GIS)在其中扮演了重要角色,因为 GIS 在整合、分析各种数据尤其是空间数据方面具有独特优势。在加州大学圣巴巴拉分校的社会科学空间综合研究中心(CSISS)得到了美国国家自然科学基金的大力资助,为促进 GIS 技术在各种社会科学中的应用发挥了重要作用。随着 GIS 技术的扩展,地理信息科学(GISc)逐渐成型,它涵盖的范围更广,如空间数据的质量与不确定性、空间数据结构的设计与发展、GIS 相关的社会与法律问题等等。2005 年 10 月 20 日,在取消地理系半个多世纪之后的哈佛大学,宣布新成立一个地理分析研究中心。是什么促使地理学研究重返哈佛?原因就在于“空间分析和地理信息系统”(哈佛大学教务部空间分析委员会 2003 年公布的“关于哈佛大学空间分析致教务长的报告”)。

当前,地理学及其他相关社会科学领域(如社会学、人类学、商学、城市与区域规划、公共管理)的许多学生都对 GIS 有着浓厚的兴趣。但是,如果他们学习的内容仅限于空间数据管理和制图的话,那么对 GIS 的兴趣很快就会消退。与此同时,很多学生抱怨统计学、定量方法及空间分析等课程十分枯燥,好像与自己的兴趣关系不大。我在多年讲授 GIS、空间分析与定量方法的教学实践中,感到把他们结合起来,并用现实世界中的数据进行案例教学的好处不少。例如,通过将一些 GIS 技术应用于犯罪

4 基于 GIS 的数量方法与应用

热点地区的识别,学生很好地掌握了相关的 GIS 技能;通过 GIS 来分析城市用地的空间分布模式,学生加深了对经典城市用地理论的理解。当学生们意识到可以用一定的计算方法和 GIS 技术来解决实际问题时,在课堂上的学习积极性也得到了极大提高。换言之,GIS 技术技巧和定量方法是通过解决实际问题来学习的,而掌握好了 GIS 技术和定量方法,学生在求职中也有了更强的竞争力。

本书是作者多年整合 GIS 和计量方法的初步成果,同时也是二者在社会科学领域中各种应用的集中演示。

对书中选取的应用案例,作者主要基于三个方面的考虑。首先是展示 GIS 在加强社会问题与公共政策应用研究中的多样性。本书涉及的 GIS 应用,涵盖从典型的城市与区域分析(如区域增长模式、零售业服务区分析)到犯罪与健康问题等多个主题。其次是举例讲解各种计量方法。有些方法,如果没有 GIS 的帮助是很难实施的,或者应用起来非常麻烦;而有些方法一经整合到 GIS 中,自动化程度大大提高。第三是涵盖空间分析中常见的任务(如距离与交通时间的估算、空间平滑与插值、可达性测量等)和主要问题(如可变地域单元问题、小人口基数小概率事件分析、空间自相关等)。

本书的一个重要特点是,各章都是任务导向的。在解决实际问题的背景下,读者更容易领会各种方法的妙处。虽然每种方法都是基于某个特定的应用进行介绍的,但它们都可以用于分析不同的问题。每章都作为一个主题,介绍与之密切相关的方法(或一组方法),例如,线性规划用于解决浪费性通勤,线性方程组用来分析和预测城市土地利用模式,空间回归用于检验就业可达性和谋杀犯罪率之间的关系,空间聚类分析用于考察癌症分布模式等。

本书另外一个重要特点是强调方法的实用。所有 GIS 任务都通过 ArcGIS 平台进行演示,大部分统计分析(包括线性规划)都用 SAS 来实现。换言之,读者只要能用 ArcGIS 和 SAS 即可操作本书讲解的内容,并可以进行类似的研究。之所以选取 ArcGIS 和 SAS,是因为它们分别

是 GIS 和统计分析方面的领先软件。有些特殊的任务如空间聚类和空间回归用一些可以从网上下载的免费软件实现。本书使用的大部分数据都可以从公开渠道获取(例如,从网上免费下载)。老师或熟练的读者可以用本书介绍的这些数据和技术来设计他们的课程项目或草拟自己的研究项目。随书赠送的光盘包含所需的全部数据及电脑程序示例(参见数据文件列表)。

本书主要适用于地理、城市与区域规划及相关领域的学生,可以用于如下一些课程:空间分析、区位分析、GIS 在商业及社会科学中的应用、地理中的数学方法。本书也可供地理和规划之外的研究者在进行 GIS 和空间分析时参考,从事城市经济研究的人也可能从本书关于城市结构及浪费性通勤的章节中获得帮助,商业领域的研究者可能会觉得有关服务区分析和可达性测量的章节有用,犯罪学家可能会对本书有关犯罪分布模式的案例研究感兴趣,流行病学家可能会喜欢本书关于癌症空间聚类分析的章节。

全书分三部分共 11 章。第一部分包括 3 章,主要介绍一些常规的问题,如 GIS 中数据管理概述及基本的空间分析工具(第一章)、距离及交通时间的测量(第二章)以及空间平滑和空间插值(第三章)。第二部分包括 4 章,主要介绍一些基本的数量方法,不需要或很少用到编程,它们分别是:服务区分析(第四章)、可达性测量(第五章)、方程拟合(第六章)及因子分析(第七章)。第三部分包括 4 章,主要介绍一些高级操作:小人口样本中的频率分析(第八章)、空间聚类和回归(第九章)、线性规划(第十章)以及线性方程组的求解(第十一章)。第一二两部分可以作为本科生高年级课程使用,第三部分可以作为研究生课程使用。本书假设读者具有相当于 GIS 入门和初级统计的基础。

除第一章之外,每章都集中介绍一种计量方法。总体来说,每章的结构大致如下:开始是介绍方法,接着讨论该方法的应用问题,最后以案例研究的形式在 GIS 上实现前面介绍的方法。那些与各章主题不直接相关但是很有意义的问题,则以附录的形式附于各章之后。许多重要的技

6 基于 GIS 的数量方法与应用

术和方法在不同案例中反复出现以加深印象、增强学习效果(可参阅“空间分析技术和数量方法速查目录”)。

我对计量方法的兴趣起源于在北大本科毕业前,杨齐老师讲的计量地理生动活泼又深入浅出,使我受益匪浅。我后来又有幸选修朱德威老师给研究生开的计量地理方法,感知良多。1988年年初,胡兆量老师受香港大学梁怡先生的启发,指导我大学毕业论文研究中国城市位序规模分布的统计模型及历史演变。胡老师把自己几十年从原建设部收集的中國城市人口资料(堆起来两尺多高)完全交给我,另外还有他关于城市研究的手稿,都是极其珍贵的唯一版本,我现在回想起来才体会到先生当时对我信任的分量。我欣喜地建数据库,改写测试杨齐老师编的回归程序,度过了我整个求学生涯中最愉快的半年多时光。这一经历使我的神经兴奋点集中在社会经济现象中呈现的科学规律,大部分要借用数理模型来刻画。1991年留美,从师于俄亥俄州立大学(以下简称俄大)城市与区域规划系的金-米切尔·古德曼(Jean-Michel Guldmann)教授。我在他的静态和动态规划两门课程上学会了线性规划及线性方程组的求解,更重要的是,他让我在异国他乡找到了自信和求学的乐趣。我们常常在他的办公室里海阔天空地聊天,鼓吹计量方法的精深。我也从俄大经济系的唐纳德·哈伦(Donald Haurin)教授那里学到不少东西,本书中关于城市与区域密度分布模式及浪费性通勤的内容,可以追溯到他关于城市经济学的课程。同为俄大城市与区域规划系的菲利普·维顿(Philip Viton)教授,教了我不少计量经济学,我真希望自己当时做学生时更用心地听好他的课。

决定撰写本书需要很大的勇气,甚至可以说是有些乐观地估计了自己。自动手写这本书来,我发现自己把许多重要事情一拖再拖,更是疏于家务。我的妻子除了全职上班,还要承担大部分家务劳动。我也经常听到几个小家伙轻声耳语:“安静点儿!爸爸在写书呢。”因此,我要首先感谢他们对我的支持。我二女儿真妮生病去世后,我消沉过一段时期。写这本书时,常以思念换取力量,由此克服不少困难。

我要感谢原来供职的北伊利诺伊大学,2004年秋季批给我半年假以便我专心撰写本书。我也要感谢当时的同事理查德·格林(Richard Greene)、安德鲁·卡门勒克(Andrew Krmeneć)和罗卫与我许多有益的讨论和提供的不少建设性意见。我要感谢伊利诺伊大学香槟分校(UIUC)地理系的穆兰博士,她开发了第八章所用的尺度空间聚类工具。北伊利诺伊大学的伦纳德·瓦尔热(Leonard Walther)帮我设计、修改和完善了部分图件。伊利诺伊州日内瓦市公共事务部的刘玉红清绘了第十一章中假想城市的电子地图,她的慷慨帮助和优秀的绘图工作帮我顺利设计成这一案例。感谢迈克·巴迪(Michael Batty,英国皇家科学院院士)为本书撰写热情洋溢的序言,我觉得好多话他比我说的更准确、更深刻。

我还要感谢CRC出版社的编辑小组:采编编辑兰蒂·科恩(Randi Cohen)和曾田·泰助(Taisuke Soda),项目协调人特丽萨·德·福恩(Theresa Delforn),项目编辑克里斯蒂·纳让诺(Khrysti Nazzaro),以及许多其他工作人员,如排版、校阅、制图、电脑等,感谢他们在本书出版过程中的大力帮助。

我本人以及在北伊利诺伊大学听过我区位分析、城市地理和交通地理课程的学生都曾反复检验过书中的案例。最近,在本书清样校对阶段,我受朋友之邀,在清华大学建筑学院、东北师范大学城市与环境学院开设了“基于GIS的数量方法及其在社会经济规划科学中的应用”短期培训课程,用的就是这本书作教材,得到了许多有价值的意见。后来英文版正式出版后,我又在北京大学城市与环境学院、北京大学深圳研究生院、华东师范大学资源与环境科学学院和首都师范大学资源环境与旅游学院,分别举办了类似的短训班,也是用它作教材,反响都不错。当然,本书可能还会存在一些错误,尤其是新版的ArcGIS界面又有更新的地方。作者诚恳地欢迎学者、老师、同学在使用本书时提出宝贵意见,这对本书将来可能的修订、再版大有裨益。

最后,我想就中文版的出版补充几句。我旅美多年,回国来才体会到对这块土地的感情。给学生讲课时,本来就高的话音又高半度,常常每天

工作十几个小时也不觉得累,就是早晨跑步也觉得步伐轻了。我近几年与国内大学和中科院交流比较多,总体的印象是我们在基于 GIS 的计量方法以及人文地理和规划上的应用与国际上的先进水平还有比较大的差距,在其他社会科学和公共政策方面的应用更是少见。我希望本书的中文版对推动这一方向在中国的发展起到一些积极的作用。

感谢姜世国、滕骏华把这本书翻成中文。我认识姜世国时,他还是北大周一星老师的硕士生(现为俄大地理系博士生)。迈克·巴迪在北大作报告时,我亲身体会他作临场翻译,英文好,技术也熟悉,我印象很深。我一提议他翻译此书,他就欣然答应。我大学同学滕骏华在北伊利诺伊大学看我时,又慷慨伸出援手,翻译了较难的第八至十一章。全书译稿完成后,我又作了详细的校对修订。许多地方为了帮助理解,我又补充了不少背景知识。我写中文提笔忘字,也不熟悉许多专业词汇的中文对应,汉字的电子输入又不熟练,真是多亏两位译者译出初稿,中科院地理科学与资源研究所的王姣娥也帮我全面校对了一遍。首都师范大学资源环境与旅游学院的李小娟也多次鼓励我把这本书的中文版尽早推出,为此调动不少资源,提供具体的帮助。

在本书中文版的校订过程中,我又基于新版的 ArcGIS 补充、更新了各章案例的少数步骤。2008 年 12 月 8~16 日,我在首都师范大学开设“基于 GIS 的定量分析在人文地理和规划中应用”课程,面向全国招收了近 40 名学员,由首都师范大学资源环境与旅游学院的八名地理信息系统专业的研究生担任我的助教。课程所有案例,除案例 7A 外,都来源于这本书(这里案例的序号与本书略有出入)。助教们认真地钻研了各个案例,制作了详尽的上机操作演示。本中文版将他们的制作成果也收入其中(包括原英文版没有的案例 7A 的数据),统一放在所附 CD 的 Workshop 文件夹内,希望对大家上机实验有所帮助。我在此再次感谢这些助教:孙婧一(案例 1A、1B)、王冠雄(案例 2A、2B、2C)、马兰艳(案例 3A、3B、3C)、王一涵(案例 4A)、麻莉(案例 5A)、王安琪(案例 6A)、李雅箐(案例 7A、7B、7C)、胡姝婧(案例 8A、8B、8C)。

最后,感谢商务印书馆的李平和田文祝,他们的信任和支持促使本书的中文版得以尽早面世。

王法辉

2009年3月

目 录

第一部分 GIS 和空间分析的基本方法

第一章 ArcGIS 入门:数据管理及基本的空间分析工具	3
1.1 ArcGIS 中的空间和属性数据管理	4
1.1.1 地图投影及空间数据模型	4
1.1.2 属性数据管理及属性连接	6
1.2 案例 1A:绘制俄亥俄州库亚霍加县人口密度图	8
1.3 ArcGIS 中的空间分析工具:查询、空间连接、地图叠加	14
1.4 案例 1B:提取克利夫兰市的普查小区、分析多边形的邻接关系	19
1.4.1 提取克利夫兰市的普查小区	20
1.4.2 识别邻接多边形	22
1.5 小结	25
附录 1 用 ArcGIS 输入、输出 ASCII 文件	26
第二章 距离和时间的测算	29
2.1 距离的测算	29
2.2 测算路网距离和网络时间	31
2.2.1 最短路径的标号设定算法	31
2.2.2 用 ArcGIS 测算路网距离和时间	34
2.3 案例 2:测算中国东北地区各县到四大中心城市之间的距离	35
2.3.1 测算欧式距离和曼哈顿距离	37
2.3.2 测算交通路网距离	38

2 基于 GIS 的数量方法与应用

2.3.3 测算交通时间	44
2.4 小结	44
附录 2 用赋值图法求解最短路径问题	45
第三章 空间平滑和空间插值	47
3.1 空间平滑	47
3.1.1 移动搜索法	48
3.1.2 核密度估计法	49
3.2 案例 3A:用空间平滑法分析中国南方的台语地名分布 ...	51
3.2.1 基于移动搜索法的空间平滑	52
3.2.2 基于核密度估计法的空间平滑	55
3.3 基于点的空间插值	56
3.3.1 整体插值法	56
3.3.2 局部插值法	57
3.4 案例 3B:表面建模及中国南方台语地名图的绘制	59
3.4.1 用趋势面分析法制图	59
3.4.2 用局部插值法绘制分布图	61
3.5 基于面域的空间插值	62
3.6 案例 3C:将克利夫兰地区普查数据从普查小区转到邻里 单元和校区	64
3.6.1 用简单整合将普查小区数据转到邻里单元	65
3.6.2 用面积权重插值将普查小区数据转到校区	66
3.7 小结	68
附录 3 空间平滑的经验贝叶斯估计	69

第二部分 初级数量方法及应用

第四章 基于 GIS 的服务区分析及其在商业地理和区域规划中的 应用	73
4.1 服务区分析的基本方法	74

4.1.1 类比法及回归模型·····	74
4.1.2 邻域法·····	75
4.2 划分服务区的引力模型·····	76
4.2.1 赖利定律·····	76
4.2.2 哈夫模型·····	77
4.2.3 赖利定律与哈夫模型的关系·····	79
4.2.4 哈夫模型的推广·····	80
4.2.5 引力模型中 β 值的估算·····	81
4.3 案例 4A:确定芝加哥小熊队和白袜队的球迷范围·····	82
4.3.1 用邻域法确定球迷范围·····	83
4.3.2 用哈夫模型确定球迷范围、绘制概率面·····	86
4.3.3 讨论·····	87
4.4 案例 4B:确定中国东北主要城市的腹地·····	89
4.4.1 用铁路旅程确定邻域区·····	90
4.4.2 用哈夫模型确定腹地·····	92
4.4.3 讨论·····	94
4.5 结论·····	94
附录 4 引力模型的经济基础·····	95
第五章 基于 GIS 的空间可达性测量及其在医疗服务研究中的 应用·····	98
5.1 可达性问题·····	98
5.2 移动搜索法·····	101
5.2.1 移动搜索法的早期模型·····	101
5.2.2 两步移动搜索法(2SFCA)·····	102
5.3 引力法·····	105
5.3.1 引力可达性指数·····	105
5.3.2 2SFCA 法和引力法的比较·····	106
5.4 案例 5:测算芝加哥地区基本医疗服务的空间可达性·····	107