



地理信息系统理论与应用丛书

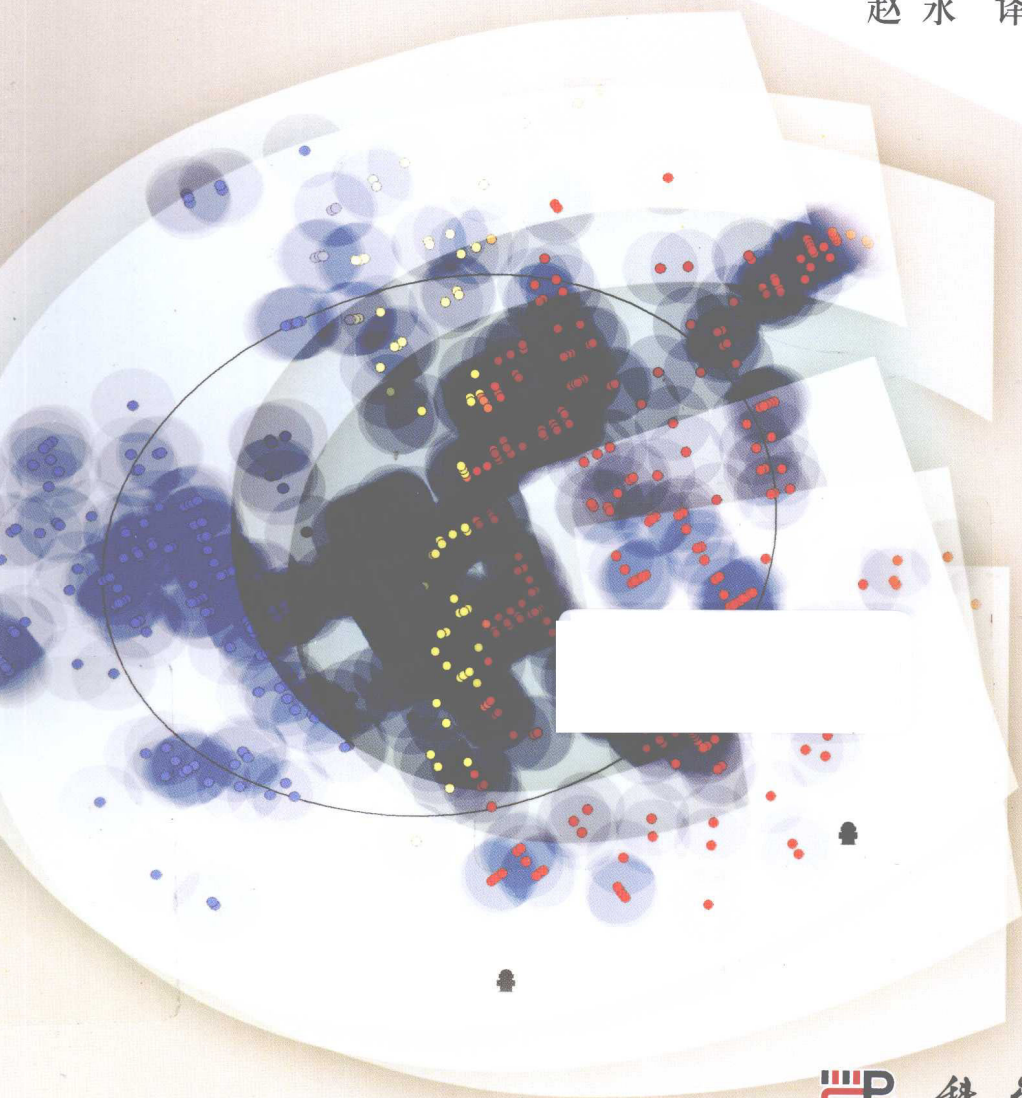
GEOGRAPHIC INFORMATION ANALYSIS

地理信息分析

(原著第二版)

〔新西〕 David O'Sullivan 著
〔英〕 David J. Unwin

赵永 译 王劲峰 审校



科学出版社

GEOGRAPHIC INFORMATION ANALYSIS

地理信息分析

(原著第二版)

[新西]David O'Sullivan 著
[英]David J. Unwin

赵 永 译
王劲峰 审校

河南省国家重点学科培育计划
国家自然科学基金(40901284) 联合资助

科 学 出 版 社

北 京

图字:01-2011-4977

内 容 简 介

本书是2010年John Wiley & Sons出版图书*Geographic Information Analysis (Second Edition)*的翻译版。内容涵盖空间分析的基本问题与基本方法,涉及制图基础、过程与模式、点模式分析、面对象与空间自相关、局部统计量、场数据分析、地图叠置,以及地理计算、空间模型、网格和云计算等空间分析的前沿问题与新方法。

本书可作为空间分析、空间数据分析、空间统计、空间模型与建模等地理信息科学领域的本科生和研究生教材,也可供该领域和相关学科的研究人员参考。

© 2010 by John Wiley & Sons, Inc.

All Rights Reserved. This translation published under license. Authorized translation from the English language edition, entitled *Geographic Information Analysis*, 2nd, ISBN 978-0-470-28857-3/0-470-28857-4, by David O'Sullivan and David J. Unwin, Published by John Wiley & Sons. No part of this book may be reproduced in any form without the written permission of the original copyrights holder.

图书在版编目(CIP)数据

地理信息分析:第2版/(新西兰)奥沙利文(O'Sullivan,D)等著;赵永译. —北京:科学出版社,2013.3

(地理信息系统理论与应用丛书)

书名原文:Geographic Information Analysis

ISBN 978-7-03-036735-8

I. ①地… II. ①奥…②赵… III. ①地理信息系统—分析方法 IV. ①P208

中国版本图书馆CIP数据核字(2013)第033924号

责任编辑:朱海燕 陈婷婷 / 责任校对:鲁 素

责任印制:钱玉芬 / 封面设计:王 浩

科 学 出 版 社 出 版

北京东黄城根北街16号

邮政编码:100717

<http://www.sciencep.com>

源海印刷有限责任公司印刷

科学出版社发行 各地新华书店经销

*

2013年3月第 一 版 开本:787×1092 1/16

2013年3月第一次印刷 印张:18 1/4

字数:436 000

定价:59.00 元

(如有印装质量问题,我社负责调换)

第二版序言

本书第一版 (O'Sullivan and Unwin, 2003) 写于 21 世纪的头两年, 其基本框架依托于在地理信息科学中具有悠久历史的两个关键思想——制图学思想和统计学思想。敏锐和年长的读者也许会注意到书中某些章节来自于 *Introductory Spatial Analysis* (Unwin, 1981), 那是本书的作者之一在约 30 年以前写的一本小册子。其第一个关键思想是在长度维的框架下, 把地理对象描述为点、线、面和连续性表面 (场); 第二个关键思想是把图示分布看作是某种空间随机过程的实现。

与其他任何框架一样, 这些观点是不完美的。如严重依赖于固定的几何实体, 并高度重视统计假设检验, 从空间表达和统计推断发展的角度来看, 这可能有些过时。在第二版中, 关于与时俱进和采用不同的替代框架, 我们进行了一段时间的思考。最终没有轻易坚持使用原方案, 但我们确信我们的决定是对的。自从该书第一版出版后, 我们基于这一框架, 在美国、英国、新西兰以及通过互联网在全球, 已经教授了高年级本科生和初级研究生水平的课程班, 并发现其是清晰的和富有弹性的教学方法。对于大多数学生而言, 无论是空间数据还是统计推理都是新的或比较新的内容, 重要的是, 我们为这样很大一部分读者打开了一道通向更高级知识领域的“大门”。对于少数熟悉这两部分内容的读者, 我们希望该书内容的覆盖面足够广, 并且能够获得读者对于书中保留的更高级内容的充分认同。

机 会

尽管第二版与第一版之间具有广泛的一致性, 但通过逐章比较, 便会看出我们在点模式分析、空间自相关、克里金和空间数据回归处理的内容上进行了大量更新和许多较大改动。其中一些是为了强调, 而另一些则是为了使内容更加充实。熟悉第一版的读者会发现, 传统的统计假设检验进一步让位于更受重视的蒙特卡洛/随机化方法, 因为在大多数实际工作中使用的杂乱数据通常需要经过预先处理, 在我们看来, 这一方法提供了与空间数据有关的众所周知的问题的很多方面。甚至在本书第一版出版 7 年后, 计算能力的提高使这一方法的应用更加切合实际且更容易实现。此外, 读者还会发现第二版中更重视对局部的描述。我们相信, 这反映了科学方法论上的一个重要的变化, 并且其在当今计算环境中的实现已经成为可能。

这两个主要的改变, 遵循并追溯这些重要的变化, 因为我们增加了关于地理可视化 (第 3 章) 和局部统计量 (第 8 章) 的全新篇章。事实上, 还有为本书第一版准备的关于地图和制图的一章, 但最终并没有采用。我们自己的解释是: 要注意压缩篇幅, 并考虑到大多数读者已经熟悉制图科学与艺术的一些核心概念。随后, 针对被称为“半途出家的地理学者” (Unwin, 2005) 的地理信息分析课程的教学经验表明, 这一疏忽是个错误。“半

路出家的地理学者”一词，是指那些空间数据分析的新手，他们对地理科学的理解在很大程度上是基于地理信息系统（GIS）软件进行的可能操作。有三个原因增加了制图（或者如果你愿意，可称之为地理可视化）的切题性。首先，即使现在有为数众多的统计方法可用，某些形式的制图或许仍然是主要的分析策略。其次，强调局部统计量并进行制图，增大了对理解制图基本原则的需求。第三，到几乎任何的 GIS 贸易展览会上走一走，你都将发现，即使有经验的 GIS 用户也会继续犯很基本的制图错误。新的第 3 章与原先拟定的内容相似性很小。新材料在很大程度上依赖于使用互联网搜索引擎来查找和评论例子，这是差不多半个世纪以前本书年长的作者在地图欣赏课堂上学习到的东西。但即便如此，我们仍采用悠久的、有时会令人遗憾地被忽略的传统制图方法，确定这一章的大部分内容。

第二个主要的补充是关于局部统计量的一章。同样，这也有素材组织上的问题，因为正如我们已经发现的那样，最初在不同背景下形成的资料，其中有相当大的比例可以理所当然地纳入这个框架。如包括所有与距离、邻接和邻域概念相关的定义空间权重矩阵 (W) 的资料；利用控制点数据对场平均高度的估计——空间插值；点过程强度估计中，局部峰值的识别——聚集；以及通过对空间自相关全局 Moran's I 的分解，识别相似的空间区域——Moran 散点图。读者无疑会发现其他的例子，它们是很多空间分析中的核心概念。该章提供了很多更为明确的空间相关性局部指标，并包括了对地理加权回归（GWR）思想的介绍。尽管也可以很容易地把核密度估计（KDE）放在同一章中，但我们认为它在地理可视化的背景下最为常用，所以把它从原来的第 3 章移到了第 5 章的点模式分析中。我们意识到，这些变化会使读者有必要不时地返过头来参考以前的材料，当有必要使用思考练习的时候，我们尝试发出这种信号。

通过删减一些材料，本版中增加的内容在篇幅上得到了平衡。首先，出于完全务实的原因，我们删除了几乎所有的线对象分析内容。虽然该书第一版中用一章讨论了线对象分析的一些基本思想，但由于篇幅的原因，它没有、也不可能反映几乎所有科学分支中网络分析的日益重要性，我们两位作者对这一章很不满意。读过那一章的读者都会认识到，当处理线对象时，我们竭力采用基本的随机过程方法。某个人、在某个地方、某一天，会写这样一本所需要的书——一本关于用网络表达的地理信息分析的书，但其内容要远远超出在该书第一版中那一章所涵盖的篇幅。本书第一版中关于多元统计量的一章，感觉把 n 维数据当作空间数据有些不适宜，故将其删去了。只保留其中的一些材料，在第二版关于地理可视化的一章中以“空间化”的标题形式出现。除了这些较大的调整，我们还取消了对描述空间自相关的连接数分析方法的扩展处理，它虽然有益于教学，但似乎与当今的实践越来越不密切。最后，为了保证本书的篇幅适宜，放弃了介绍基本统计概念的一个附录，我们假设读者能够从众多优秀的入门教科书中获得这些知识。

自从本书第一版出版后，情况已经发生了很大变化，计算科学、统计学和地理信息科学的一般领域都有很大的发展。在材料更新过程中，我们试图尽力反映这类新成果和它所处的日益“局部觉醒的”科学和社会环境，但我们知道很多内容被略去了。如果你很失望没有看到你想要的东西，我们只能表示歉意。

软 件

我们力图反映的一个重大变化是，学术型空间分析者所采用的方法与大多数商业 GIS 中所提供的嵌入型空间分析功能之间的差距在日益扩大。诚然，如果你知道自己在做什么，并且不总是依赖于软件的默认设置，那么我们所描述的许多方法可以在这种软件系统中应用，但这种应用是不理想的。在过去的十年中，当今最领先的研究人员已经在公共 R 语言环境中 (Ihaka and Gentleman, 1996) 进行他们的研究工作，而这已变得越来越明显。希望实践我们所介绍的方法的读者应当注意，几乎所有方法以及更多其他方法，已经在 R 语言环境中实现 (Baddeley and Turner, 2005; Bivand et al., 2008)，强烈建议希望开发新的和创新的地理信息分析方法的读者加入这个学者们的社区。

David O'Sullivan

奥克兰大学

Dave Unwin

伦敦大学伯贝克学院

2009 年毛利新年

参 考 文 献

- Baddeley, A. and Turner, R. (2005) Spatstat: an R package for analyzing spatial point patterns. *Journal of Statistical Software*, 12: 1-42.
- Bivand, R. S., Pebesma, E. J., and Gomez-Rubio, V. (2008) *Applied Spatial Data Analysis with R* (New York: Springer) .
- Ihaka, R. and Gentleman, R. (1996) . R: a language for data analysis and graphics. *Journal of Computational and Graphical Statistics*, 5: 299-314.
- O'Sullivan, D. and Unwin, D. J. (2003) *Geographic Information Analysis* (Hoboken, NJ: Wiley) .
- Unwin, D. J. (1981) *Introductory Spatial Analysis* (London: Methuen) .
- Unwin, D. J. (2005) *Fiddling on a different planet*. *Geoforum*, 36: 681-684.

致 谢

David O'Sullivan

在本版中，宾夕法尼亚州地理学课程（编号：GEOG 454、455、586）和奥克兰地理学课程（编号：GEOG 318、771）中的学生和同事以多种方式，或多或少地帮助完善了材料。我对 David Dibiase, Mark Gahegan, Frank Hardisty, Jim Harper 和 George Perry 通过各种形式给予的支持表示感谢。我也要感谢奥克兰大学的一个研修奖项和空间认知教学（SpLinT）协会（英国莱斯特大学与诺丁汉大学、伦敦大学学院），尤其是 Paul Longley 和 Nick Tate，感谢他们在 2008 年提供的研究奖金。研修和 SpLinT 奖使得 Dave 和我一起为本书工作了愉快而富有成效的几天，并帮助设计了这个版本。

我也感谢 Gill，感谢她一直以来的宽容和支持。最后，周六早上 Fintan 和 Malachy 学院在板球场和足球场上的过招，更不用说深邃的虚拟 LEGO™ 星系，随时随地都提醒着我事物的真正中心在哪里。

Dave Unwin

华盛顿“空间培训班”点燃了我的研究生的空间分析热情，而最终遇到他们几个是一件很高兴和荣幸的事。伦敦阿伯里斯特威斯、莱斯特和伯贝克地理系提供了研究基地，并且有知识渊博的同事。特别感谢在莱斯特中部地区研究实验室与我一同讨论众多问题的以前的同事，他们是 Alan Strachan, Mike Worboys, David Maguire, Jo Wood, Jason Dykes，以及最近认识的 Pete Fisher。所有教师都应该向他们的学生学习，我非常愉快地承认他们对我思想形成的贡献，有时是在不知不觉中，有时是直接受到了如下研究生班的启发：英国的莱斯特大学、伦敦伯贝克学院和伦敦大学学院，新西兰的怀卡托大学和坎特伯雷大学，以及美国的雷德兰兹大学。一个最后和特别形成的经验是，在互联网上对“半路出家的地理学者”进行的教学和辅导班，已经证实了我的信念——在科学中应用地理学的价值。像我的同事一样，我很感谢 SpLinT 的发起者在 2008 年给予的支持。

仍然存在的可能的错误，当然由我们自己负责。

David O'Sullivan

奥克兰，新西兰

Dave Unwin

梅德韦尔，英格兰

第一版序言

这本书由一本小册子“自由成长”起来，那本小册子是我们中的一位作者从 1979 年至 1981 年写的 (*Introductory Spatial Analysis*, 伦敦: 梅休因出版社)。虽然它在第一个商业地理信息系统 (GIS) 出现后整整十年, 并差不多与第一台微型计算机问世的时间一致, 但那本书的传统根植于 20 世纪 60 年代的计量地理学方法, 并且所采用的工具比一个便携式计算器复杂不了多少。从 1983 年起我们尝试出版第二版的计划总是被其他项目搁置, 这些项目几乎总是与 GISs 在当代的迅速发展有关。与此同时, 电脑在发达国家普及到几乎每个人, 在科学研究和商业中许多人发现了应用 GIS 软件的地理学潜力。到 20 世纪 90 年代末, 很明显, 只有一本全新的教材才能够如此, 是在这一点上我们两人开始联手做这项工作, 其结果就是现在的这本书。

本书所涵盖的内容已经实践了相当长的时间, 并已在我们任教的莱斯特、伦敦 (伦敦伯贝克学院和伦敦大学学院)、宾夕法尼亚 (宾夕法尼亚州)、怀卡托、坎特伯雷 (新西兰) 和其他地方的大学高年级本科生和研究生课程中试用过。我们热爱在几乎任何工作中介绍的与地理数据相关的有用概念和技术, 并且希望能成功地传达给读者。关于这些思想狂热的或率真的应用, 我们也希望当中所陈述的保留意见不损害我们根本的积极性。我们的许多保留意见源于一个原因, 即利用当前 (或许将来?) 可能的技术数字化地表达外部现实的限制性。我们觉得, 如果 GIS 是用来作为支持地理学中各种研究方法的一种有效工具, 而不是作为“回答”每个地理问题的一种万全之策, 那么即使是对该领域的新人, 揭示这种不确定性也是恰当的。

虽然计划上不是这种方式, 但经过再三考虑, 我们从精心制订、全面阐述基础知识开始, 并基于 *Introductory Spatial Analysis* 的理性传统进行讲述, 一直到对最近的计算密集型方法比较散漫的描述。早期的材料强调基本概念和问题的重要性, 这对把统计方法应用于空间数据的任何尝试而言很常见, 并应当为后面章节中更一般地讨论的更高级方法的进一步研究, 提供一个坚实的基础。我们提供的一些较老的参考文献表明这样的一个事实, 即至少现在被称为地理信息科学的一些慧根, 牢牢地根植于 20 世纪 60 年代的地理学和各种不同的学科。近些年在地理数据分析中看到大量的技术创新, 我们希望能够在该书和建议阅读的材料当中, 捕捉到一些令人兴奋的内容。

通篇让我们纠结的两个问题是数学和符号的应用。这两者联系在一起, 都需要谨慎对待。数学上, 我们试图尽可能严谨并与目标受众相一致。经验表明, 找到他们的 GIS 分析方法, 并希望更深入地探讨一些问题的学生, 具有特别不同的数学背景。因此, 我们的“严谨”, 经常表现为仅仅采用正式的符号。除了第 9 章单独的“这可以证明”外, 我们设法避免使用微积分, 但在整个第 5 章及以后使用的矩阵和向量符号, 对于完整地理解相关方法, 大抵是必不可少的。附录 B 提供了一个关于矩阵和向量的指南, 对大多数读者而言应该足够了。如果把该书作为一门课程使用, 我们强烈建议教师在讲述正文之前, 在适当

时候花些时间介绍本附录的相关内容。我们假设读者具有基本的统计学基础，但为了保险起见，我们包含了一个类似的附录，概述了所用的主要统计思想，以及类似解释。

差的符号有着混淆的巨大可能性，而空间分析是以很多变量称道（也许是让人讨厌）的领域。绝对的一致性是难以保证的，并且在任何情况下这都可能是一种过于乐观的估计。我们力图尽可能地一致和明确。也许在这方面最不和谐的是在第 8 章和第 9 章中介绍的第三维位置坐标，它用字母 z 表示。当在 9.3 节讨论空间坐标回归时，这将会导致一些尴尬和稍微的符号变化。经过权衡，比起太多的书页中下标林立（快速地翻一下书，确保你们不应被保留下来的许多下标吓倒），我们宁愿使用 (x, y, z) 并容忍不一致的指责。就像该书的前身，这种实用方式应作为一种醒示，提醒这是关于地理信息实际分析的一本书，而不是作为一本空间统计学专著。最主要的，这是一本地理书！

涵盖这么多基础的这样篇幅的一本书，可能永远不会只是两个人的工作，而没有其他人的帮助。多年来，我们中的一位得益于与教育领域的同事和地理信息企业的接触，他们太多而不能一一列出。他对所有意见、讨论与心平气和的争论表示感谢。在这个迅速发展的领域，它是开放和建设性氛围的一种证明，这种伙伴关系中年轻的一方已经受益于许多类似的接触，这也很难一一列出。我们只想说，英国伦敦大学学院的高级空间分析中心和宾夕法尼亚州州立大学地理系的支持对此帮助很大。照例，存在的错误仍由我们负责。

David O'Sullivan
宾夕法尼亚州立大学
2002 年圣基兰节

Dave Unwin
伦敦，英格兰
2002 年情人节

目 录

第二版序言

致谢

第一版序言

第 1 章 地理信息分析与空间数据	1
本章目标	1
1.1 引言	1
1.2 空间数据的类型	3
1.3 一些困难	6
1.4 属性描述的尺度	12
1.5 GIS 与空间数据操作	16
1.6 前面的路	19
本章回顾	19
参考文献	20
第 2 章 空间数据的陷阱和潜力	22
本章目标	22
2.1 引言	22
2.2 坏消息：空间数据的陷阱	23
2.3 好消息：空间数据的潜力	27
本章回顾	35
参考文献	35
第 3 章 基础——把它图示出来	37
本章目标	37
3.1 引言：制图学传统	37
3.2 地理可视化与分析	39
3.3 雅克·伯汀的图形变量	40
3.4 新的图形变量	42
3.5 地理可视化中的问题	43
3.6 制图和探索点数据	44
3.7 制图和探索面数据	49
3.8 制图和探索场数据	54
3.9 非空间数据的空间化	57
3.10 总结	58
本章回顾	59

参考文献	59
第4章 基础——把地图作为过程的结果	63
本章目标	63
4.1 引言：地图和过程	63
4.2 过程及其所产生的模式	64
4.3 预测由过程产生的模式	68
4.4 更多定义	72
4.5 线、面和场的随机过程	73
4.6 总结	78
本章回顾	79
参考文献	80
第5章 点模式分析	82
本章目标	82
5.1 引言	82
5.2 描述一个点模式	83
5.3 统计地评估点模式	94
5.4 蒙特卡洛检验	100
5.5 总结	103
本章回顾	104
参考文献	104
第6章 实用点模式分析	106
本章目标	106
6.1 引言：空间统计分析的问题	107
6.2 常规统计推断的替代方法	108
6.3 IRP/CSR 的替代模型	109
6.4 现实世界中的点模式分析	112
6.5 处理非齐次性	114
6.6 焦点检验	116
6.7 聚集探测：扫描统计量	117
6.8 应用密度和距离：近邻多边形	120
6.9 距离矩阵和点模式分析的一个注记	122
本章回顾	123
参考文献	124
第7章 面对象与空间自相关	127
本章目标	127
7.1 引言：面对象复习	127
7.2 面对象的类型	127
7.3 面的几何特性	130

7.4 测度空间自相关	135
7.5 一个例子：2001~2006 年奥克兰的肺结核	140
7.6 其他方法	143
本章回顾	144
参考文献	145
第 8 章 局部统计量	147
本章目标	147
8.1 引言：地理地思考，局部地测度	147
8.2 局部性定义：再次提到空间结构	149
8.3 一个例子：Getis-Ord G_i^* 和 G_i^* 统计量	149
8.4 局部统计量推断	152
8.5 其他局部统计量	154
8.6 总结：局部地看世界	159
本章回顾	160
参考文献	160
第 9 章 场数据描述与分析	162
本章目标	162
9.1 引言：标量和向量场回顾	162
9.2 模拟与存储场数据	164
9.3 空间插值	169
9.4 表面的衍生测度	178
9.5 地图代数	183
9.6 总结	185
本章回顾	186
参考文献	186
第 10 章 认识未知——场统计学	188
本章目标	188
10.1 引言	188
10.2 空间坐标回归：趋势面分析	189
10.3 平方根差云图和（半）变异函数	195
10.4 统计插值方法：克里金	199
10.5 总结	211
本章回顾	212
参考文献	212
第 11 章 把地图放在一起——地图叠置	214
本章目标	214
11.1 引言	214
11.2 布尔叠置与筛选叠图	216

11.3	替代布尔叠置的一般模型	221
11.4	指数叠置与加权线性组合	222
11.5	证据权重	225
11.6	应用回归的模型驱动叠置	226
11.7	总结	227
	本章回顾	228
	参考文献	228
第 12 章	空间分析新方法	231
	本章目标	231
12.1	不断变化的技术环境	231
12.2	不断变化的科学环境	233
12.3	地理计算	234
12.4	空间模型	239
12.5	网格和云：超级计算简介	245
12.6	总结：新地理信息分析？	247
	本章回顾	248
	参考文献	248
附录 A	符号、矩阵和矩阵数学	253
A.1	简介	253
A.2	符号的一些基本说明	253
A.3	矩阵基础与符号表示	255
A.4	简单的矩阵数学	257
A.5	用矩阵求解联立方程组	261
A.6	矩阵、向量和几何结构	264
A.7	特征向量和特征值	266
	参考文献	267
	词汇表	268
	译后记	276

第1章 地理信息分析与空间数据

本章目标

在本章中，我们：

- 定义本书中的地理信息分析。
- 当涉及地理信息分析和基于 GIS 的空间数据操作时，能够把两者区分开。
- 概述空间数据的实体-属性模型，它由点、线、面和场组成，具有相关的名义量、次序量、间隔量和比率量。
- 注意针对空间数据模型观点中的某些困难，特别是不同尺度上的多重表达、时间、具有不确定边界的对象、模糊对象和可能是分形的对象。
- 概述空间数据操作处理，并强调它们的重要性。
- 探讨地理实体表达形式之间的不同转换，指出它们在地理信息分析中的用途。

读过本章后，你应该能：

- 列出 4 种不同的空间分析方法，以及它们之间的不同。
- 说出为什么现代空间分析方法没有很好地嵌入典型 GIS 工具包中的原因。
- 区分空间对象和空间场，并讨论为什么 GIS 中矢量与栅格的争论，实际上是如何选择表达这些实体类型的问题。
- 区分并举例说明点、线和面对象。
- 列出描述一个场的基本的数据特征。
- 提供不易纳入该体系的现实世界实体的例子。
- 对现实世界实体在数字化数据库中的表达，以及它在地图上的显示保持清醒的区分。
- 区分并举例说明名义量、次序量、间隔量和比率量 4 种属性数据。
- 给出至少 12 种空间数据组合类型的例子。
- 列出典型 GIS 中一些基本的几何数据操作。
- 概述可以对实体表达进行转换的方法，并解释为什么这对地理信息分析有用。

1.1 引言

地理信息分析不是一个已经建立起来的学科。实际上，它是一个比较新的概念。为了定义我们用该术语所要表达的意思，首先有必要定义一个较老的术语——空间分析，然后描述我们如何看待这两者之间的关系。但是，简洁地定义空间分析也不容易。空间分析在不同的背景中提及，文献中分别以不同的方式应用该术语，它至少有四个可以识别的宽泛领域：

(1) 空间数据操作, 一般应用于地理信息系统 (geographic information system, GIS), 通常称为“空间分析”, 在 GIS 公司的宣传材料中更是如此。GIS 手册将让你明确这些技术的范围, 如 Tomlin (1990) 和 Mitchell (1999)。

(2) 空间数据分析, 具有描述性和探索性。这是所有空间分析中重要的第一步, 并往往可应用于非常庞大和复杂的数据集。地理学家, 如 Unwin (1982)、Bailey 和 Gatrell (1995)、Fotheringham 等 (1999) 的书遵循这一传统。

(3) 空间统计分析, 应用统计方法研究空间数据, 以确定数据是否可以用一个统计模型来表示。上述提及的地理教科书涉及这些问题, 并有少数统计学家对空间数据分析感兴趣, 特别是 Ripley (1981, 1988)、Diggle (1983) 和 Cressie (1991)。

(4) 空间建模, 需要构建模型预测空间结果。在人文地理学中, 模型用于预测物流和人流, 或者进行设施位置的优化 (Wilson, 2000), 而在环境科学中, 可利用模型尝试模拟自然过程的动力学 (Ford, 1999)。建模技术是空间分析的一个自然扩展, 但超出了本书的研究范围。

实际上, 通常很难区分这些方法, 很多正式研究包含所有这四种方法。首先, 收集数据、可视化和进行描述性分析; 接着, 基于探索性技术可能提出问题, 并就感兴趣的现象给出理论建议; 然后, 用空间统计技术对这些理论进行统计检验。于是, 当前理论可以作为感兴趣现象的模型构建基础, 其结果, 反过来, 可能会需要进行更多的统计研究和分析。

不可能只考虑地理信息, 而不考虑日益成为其分析基础的技术——地理信息系统 (GISs)。虽然 GISs 不是像 (比如说) 字处理器那样无处不在, 但它们已经渗透到越来越多的企业、政府部门和其他决策机构。即使这是你第一次阅读一本地理教科书, 在不知情的情况下, 也许当你在网站上查找一个度假目的地地图时, 或者是寻找行车路线时, 你可能已经使用了 GIS。

在以上的列示中, 当前的 GISs 通常包含条目 (1) 作为其基本内容 (因为一个没有这些功能的 GIS 将只是一个普通的 IS!), 并包括一些简单的数据分析能力, 特别是图示探索性分析 [条目 (2)]。最近, GISs 开始包含条目 (3) 的一些统计方法, 而只包括很少的模型构建并确定其可能结果的功能 [条目 (4)]。实际上, 扩展 GIS 完成这种分析是困难的, 这是为什么许多地理信息分析者应用其他软件的原因, 这些软件属于条目 (3) 和条目 (4)。在本书中, 我们主要关注条目 (2) 和条目 (3)。在实践中你将发现, 尽管可用的工具在迅速发展, 但空间数据的统计检验仍然相对缺乏。统计方法可以很好地解决和认识某些类型的数据, 而对许多其他数据则并非如此。

如果空间分析是如此需要 (甚至值得) 写一本书, 那么它为什么不是标准的 GIS 工具包中的一部分呢? 我们给出一些原因, 其中包括:

(1) 空间数据的 GIS 视角与空间数据的空间分析视角是不同的。空间数据的空间分析视角更关注过程和模式, 而不是数据库管理和操作; 但大多数 GIS 购买者 (政府部门、公共机构) 对空间数据库的基本要求更为重视, 而不是要求完成复杂的和有时是模糊的空间分析。

(2) 空间分析没有被广泛地认识。空间分析不是很简单或特别容易, 虽然在本书中我

们的目标是解决这一问题。这个明显的困难意味着，很难说服软件生产商将空间分析工具作为标准产品。相反，空间分析工具可能是 GIS 经常忽视的一个附加产品。近些年，这种情况变得不太明显，因为软件工程方法使得 GIS 供应商可以提供独立出售的“扩展模块”给需要的用户。同时，第三方生产商可以比以前更容易地提供附加组件，并且开源软件在一些人群中已日益成为一种重要的选择。

(3) 空间分析的视角有时掩盖了 GIS 的优势。通过应用空间分析技术，我们往往提出尴尬的问题：“看起来那儿有一个模式，但它在统计上是显著的吗？也许不是。”这是一个很困难的 GIS 软件产品卖点！

除了对上述几点的关注之外，不要低估 GIS 提供的空间数据操作函数的重要性，比如缓冲区、点在多边形内的查询等。这些是产生问题和形成假设的重要前提。为了强调它们的重要性，我们将在 1.5 节中综述这些主题，并考虑如何用统计方法加以改善。更广泛地说，空间数据的存储方式，或者说如何在 GIS 中表达地理现象，对分析变得日益重要。因此，我们将在 1.2 节和 1.3 节中花一些时间讨论该问题。

基于这些原因，我们使用了更广泛的术语地理信息分析涵盖本书涉及的内容。该术语的一个实用定义是：它关注模式的研究，而模式可能是过程的一个结果。其中，表示、描述、测度、比较和产生空间模式的技术与方法，对地理信息分析的研究至关重要。当然，在这一点上我们的定义没有多大用处，因为它提出了一个问题，即这里的模式与过程是何含义。现在，我们将接受你对关键术语含义的任何直观理解。当我们在第 4 章和第 5 章阐述完点模式分析的概念后，这两个术语的意思将会变得更加清晰。现在，我们将注意力集中于了解可能遇到的一般空间数据类型。

1.2 空间数据的类型

思考练习：地理表达

贯穿本书，你将发现思考练习会帮助你以更具体的方式理解教材。通常，我们会要求你做某件事，并利用这些结果得出一些结论。你应该会发现，这些练习能帮助你记忆书中介绍过的内容。本书的第一个思考练习关注于如何在计算机里表达地理现象：

(1) 假设你为一个道路养护机构工作，你的职责扩大到县以上大小区域内的道路。你的 GIS 需要支持一些操作，比如表面重建，避免与同样在路上挖洞的其他机构（比如公共事业公司）发生冲突，并改善路面结构。

思考并写出你将如何在数据库中记录路网的几何结构。你会收集什么样的道路属性？

(2) 想象一下，你现在在同一地区为一个巴士公司工作。现在，GIS 必须支持这样的操作，如编制时间表、提出对现有和潜在巴士新路线的需求、优化站点配置。

路网的几何结构及其属性与 (1) 中的建议有什么不同？

(3) 从这里我们可以得到什么简单结论？应该明确的是，如何表达相同的地理实体依赖于表达的目的。这是显而易见的，但它很容易被忘记。

撇开有关技术问题，社会对地理信息分析的批评往往基于这样一个事实，即分析常常局限于可以很容易地进行数字化表达的现实世界的那些方面（Fisher and Unwin, 2005）。

当用地图思考世界时，你如何看待它呢？在早期的 GIS 文献中，通常区分两种系统，它们以如何数字化表达现实实体为特征：

(1) 矢量系统，它记录构成地图的特征位置坐标 (x, y) 。在矢量观点下，我们列出特征，并把每个特征表示为一个点、线或面对象。矢量 GIS 起源于基于数字化数据用计算机绘制地图，并受到特别重视，因为当时计算机存储器是一种昂贵的商品。虽然匹配不精确，但矢量模型最接近现实世界由对象组成的观点，它认为空间是一种空的容器，由不同类型的对象填充其中。

(2) 栅格系统。它不是从地面上的对象开始，而是定义一个地表上称为“像素”的小单位网格。对于每个像素，记录下它的值，或者说是感兴趣的事物存在或不存在。如此，我们把地图划分为相同的离散单元，并列出每个单元的内容。因为空间中的每个位置都有一个值（即使是零或空），所以栅格方法通常比矢量方法占用更多的存储空间。栅格 GIS 主要起源于图像处理，其中多数数据来源于遥感。

在本节中，我们希望能够说服你，在更高的抽象层次上矢量/栅格的区分不是非常有效，并且它模糊了一个更为重要的界限，即我们所说的对象观点和场观点之间的界限。

对象观点

在对象观点中，我们把世界看作是位于空间中的一系列实体。实体（通常）是实际存在的，你可以触摸它们，站在它们当中，甚至移动它们。一个对象是一个实体的全部或部分的数字化表示。对象可以划分为不同的类型，如点对象、线对象和面对象，在具体应用中，这些类型由特定的对象进行实例化。例如，在一个环境 GIS 中，森林和农田可能是面对象的实例。在对象观点中，空间能被任何数量的对象占据。普查小区里有房子，其中可能还包括灯柱、巴士站、路段、公园等。

因为对象观点也可能与对象的行为联系在一起，当具有良好定义的对象随时间发生变化时，对象观点具有其优势——例如，一个普查对象区域内一系列人口普查数据的变化。请注意，我们没有说任何关于计算机科学意义上面向对象的概念。Worboys 等（1990）给出了与空间数据相关的这一概念的简单描述。

场观点

在场观点中，世界由空间上连续变化的属性组成。一个例子是地球本身的表面，其中，场变量是海拔高度。同样，地面上有或者没有一座房子，我们可以用一个格网单元对其进行编码，其结果也是一个场。在这个二进制数表达的情况下，其中，1 = “有一座房子”，0 = “没有一座房子”。如果一座房子足够大或它的轮廓穿过了网格单元的边界，可