

第 1 章 导 论

1.1 引 言

Paelinck 和 Klaassen(1979)的《空间计量经济学》全面阐述了空间计量经济学的研究对象和内容,说明空间计量经济学已成为计量经济学的一个重要分支,标志着它作为一门学科的诞生。空间计量经济学将研究主体的地理单元空间关系引入计量经济学之中,把空间效应分为空间依赖性和空间异质性,能够为区域和城市计量模型的构建提供方法论基础。空间计量经济学近 40 年的发展历程,从一个较为冷门的学科发展成为较热的研究学科,显示出其强大的生命力。近年来,关于空间计量经济学应用的文章如雨后春笋般涌现。孙酒文和姚鹏(2014)将计量经济学的研究领域概括为模型识别、模型估计、模型检验以及空间预测。

当前,国内学者编著出版了不少关于空间计量经济学方面的教材,如《空间计量经济学:现代模型与方法》《空间计量经济学的前沿理论及应用》《空间计量经济学理论及其方法应用——基于 R&D 溢出效应测度的视角》《空间计量经济学的理论与实践》《空间计量经济学理论及其在中国的实践应用》《空间计量经济学》等。同时也出现了一些国外空间计量经济学译本,如《空间计量经济学:从横截面数据到空间面板》《空间计量经济学:空间数据的分位数回归》《空间计量经济学导论》。这些教材或专业书籍极大地丰富了国内空间计量经济学理论研究及应用的开展,本书作者将以上书籍大致分为两类:一类是侧重系统介绍空间计量经济学理论、估计方法及推断等问题,同时也不乏提及一些前沿的空间计量方法;另一类是在梳理相关计量经济学理论方法的同时,侧重空间计量经济学模型的应用。例如用于研究 FDI(外商直接投资)对区域空间知识溢出、持续性创新、经济增长等;用于研究区域层面上的 R&D(研究与开发)溢出效应测度;用于研究产业集中对区域协同发展的影响;用于研究空间视角下区域产业之间的耦合关系。大量的应用研究都有一个共同的特性,即整体研究的内部变量之间会通过某种空间关系联结在一起,使研究变得更加复杂化,但这也是计量模型发展从特殊到一般的体现。当然,为了学者进行应用和理论研究,上述书籍也给出了大部分模型在不同软件下的实现方法,但正如我们所知,空间计量经济学是当前计量经济学研究中的一个热门领域,近年来快速发展,在已有研究的基础上,空间计量经济学的方法理论得到了较大的改进和延伸,如加入了非线性变量的半参数空间计量经济学模型以及多国空间计量经济模型发展——GVAR(全局向量自回归)模型等。这些模型在以往的书本中较少涉及,更多的是出现在相关的学术论文之中。因此,我们对近年来出现的一些新的空间计量经济学模型进行整理,更有助于学习者进行理论研究及应用。

虽然本书作者在 2015 年编著的《空间计量经济学》仍然具有一定的适用性,但也存在以下几方面的问题:一是一部分重要的空间计量经济学模型方法没有纳入其中,近年来出现的新方法也无法在其中体现;二是相对侧重空间计量经济学方法及理论描述,虽然附有相对丰富的案例,但对于模型的软件实现缺乏足够的说明,许多人往往因未能熟练掌握软件对空间计量经济学的应用望而却步;三是“重建模思想而不是数学推导”的原则没能贯穿全书,因此,对于本科生和研究生来说,理解起来难度偏大。鉴于此,我们决定重新编著一本适合经济学、管理学本科生和研究生学习的空间计量经济学教材,同时,这也可以作为学习者进行现代空间计量经济学应用研究的基础参考书,并将它命名为《应用空间计量经济学:软件操作和建模实例》。

1.2 本书框架介绍

我们将本书定位为经济学、管理学本科生和研究生学习空间计量经济学的适用教材,也作为进行现代空间计量经济学应用研究的基础参考书。本书的前半部分仍是一些相对基础的空间计量模型,包括空间相关性及其检验、空间权重、三代空间计量经济学模型(横截面数据模型、静态面板数据模型、动态面板数据模型)的设定、参数估计及统计检验等,同时,也附有大量详尽的操作案例。我们对这部分内容进行了系统性整合,详细解释了不同模型的设定及其相关估计方法,说明了不同模型的适用性及如何对估计结果进行解释,有助于经济学、管理学专业的本科生和研究生对空间计量经济学由浅入深地进行理解与学习,掌握并应用空间计量经济学的基本研究方法。本书的后半部分包括扩展空间计量经济学模型和半参数空间计量经济学,适合高等院校和科研机构中从事空间计量经济学研究的学者与部分硕士和博士研究生进行深入学习。

本书共有 6 章内容,遵循由浅入深的撰写原则,方便读者理解并掌握。同时,为了便于读者进行针对性或目的性的研究,我们接下来将介绍本书的基本框架及各章对应的主要内容,具体如下。

第 1 章为导论部分,分为三节。1.1 节主要介绍空间计量经济学研究的基本内容,对现有空间计量经济学相关参考书的简要介绍并分类,言明撰写本书的目的。1.2 节为本书基本框架的介绍,方便读者对具体模型的查阅与深入研究。1.3 节介绍空间计量经济学的相关软件基础。

第 2 章主要介绍空间计量经济学的经济理论基础。空间计量经济学从某种程度上来说是地理学、经济学和空间统计学的有机结合,因此本章首先引出了空间效应的概念及来源,度量空间相关性,并对空间相关性进行统计检验,使读者能够深入理解经济社会中普遍存在的空间相关性;然后,说明空间权重设定的基本原则及设定方法。

第 3 章主要介绍了第一代空间计量经济学模型(截面数据空间计量经济学模型),相比于以截面数据为样本构建的经典计量经济学模型,空间计量经济学认为样本的空间相关性客观存在,故有必要在经典模型中正确引入空间相关性、发展空间计量经济学模型理论与方法。本章内容包括广义空间自回归模型、空间误差模型、空间杜宾模型、广义嵌套

空间模型、空间滞后模型、空间杜宾误差模型和矩阵指数空间模型。对于上述模型,我们先给出了具体模型的数学表达并整合了现有对该模型的估计方法,同时对不同模型进行了案例分析,并交代了相关的软件操作。由于模型种类繁多,错误的模型选取会导致估计结果出现严重的偏误,因此,在本章的最后,我们针对不同的模型,给出了具体模型的选取标准。

第4章主要介绍了第二、三代空间计量经济学模型(面板和动态面板数据空间计量经济学模型)。截面数据空间计量经济学,只考虑了空间单元之间的相关性,而忽略了具有时空演变特征的时间尺度之间的相关性,而本章介绍的面板数据空间计量经济学不仅能克服这方面的不足,而且可以有效地解决时间序列分析受多重共线性的困扰。本章内容具体包括面板数据空间滞后模型、面板数据空间误差模型、面板数据空间杜宾模型、空间变系数回归模型、面板数据空间杜宾误差模型及动态空间回归模型等。本章是以往的横截面类空间模型的进一步推广,需要注意的是在模型构建过程中,我们不仅要关注固定效应和随机效应的选取,同时也要注意区分不同种类的固定效应。

第5章主要讲述扩展空间计量经济学模型。较以往的空间计量经济学有所不同,本章所讲部分内容是经典空间计量经济学横向延伸,如空间离散选择模型、空间分位数回归模型、空间联立方程模型不仅能够保留离散选择模型、分位数回归模型、联立方程模型所研究问题的特性,同时还能在此基础上考察其空间效应,其估计结果将更加真实有效。此外,本章的空间向量自回归模型以及全局向量自回归模型均可以看作空间模型和向量自回归模型的结合,但两者间又存在较大的区别,全局向量自回归模型不仅考虑了全局变量对各个地区内生变量的影响,而且其在进行系数估计时是从单个地区的模型开始的。该章的学习能够增强读者对空间计量经济学模型的广度认识,为空间计量经济学的进一步扩展打下一定的理论基础。

第6章的模型专门讲述了三类空间滞后模型同非参数模型结合后所形成的半参数模型,在模型估计过程中需要用到局部线性估计法和广义矩估计法,因此,要求读者较为熟练地掌握这两种估计方法。在模型的解释上,对于模型的线性项之间的关系,其解释方式同原有线性模型一致,对于模型中的非线性项,则通过偏导数图来解释。

1.3 相关软件基础

对于同一个空间计量经济学模型,往往可以有不同的软件实现方式。本书为方便读者在学习过程中重现相关的操作过程,我们会采用 Matlab 软件、Stata 软件、GeoDa 软件和 R 软件等计量工具来实现本书各种模型的估计、检验和预测。下面将对本书用到的主要软件进行简单的介绍。

1. Matlab 软件

Matlab 是 matrix laboratory 的缩写,俗称矩阵实验室。最早的 Matlab 程序是 1967 年由 Clere Maler 用 Fortran 编写的,之后逐步变成由 Mathworks 用 C 语言完成。经过多次版本的升级,Matlab 已经具有强大的数值计算能力,可以高效地完成各种矩阵的运

算。同时, Matlab 已经从一个简单的矩阵分析软件发展成为具有高性能的运算工作平台, 所附带的工具箱涉及 30 多个领域, 包含数百个内部函数。例如, 图像处理、金融建模与分析、信号处理与通信、工程计算、控制设计、管理和调度优化计算等。

正是因为 Matlab 具有强大的功能, 所以其深受国内外大学的欢迎, 尤其是在美国大学中, 受到了广大师生的好评。为使用者省去了烦琐的计算过程, 可以让广大学生把关注点集中到数学含义的基本理解上, 这就使它成为许多大学生和研究生课程中的有效运算工具。此外, 所有 Matlab 主包文件和各种工具包都是可读可修改的文件, 用户通过对源程序的修改或加入自己编写的程序构造新的程序包。

2. Stata 软件

Stata 是用于 Windows、Macintosh 以及 Unix 电脑系统下的一种功能完备的统计软件包。其作为一款数据分析、数据管理、绘制专业图表、计量分析的完整及整合性统计软件, 具有快捷、灵活且易用的特点, 现在已越来越受人们的重视和欢迎。它和 SAS(统计分析系统)、SPSS(统计产品与服务解决方案)一起被称为三大权威统计软件, 已经被广泛应用于金融、经济、生物医疗、心理学等众多学科领域, 亦可供高等院校相关专业的本科生、研究生及从事统计分析的学者使用。

当然, 该软件也允许用户根据需要来创建自己的程序, 添加更多的功能。大部分的操作可以通过下拉菜单来完成, 也可以直接通过调用命令来实现, 初学者还可以通过阅读菜单中的帮助选项来进行相关学习。随着该软件的普及, 越来越多的计量经济学模型也开始用该软件来实现, 如 Cox 比例风险回归、指数与 Weibull 回归、多类结果与有序结果的 logistic 回归、poisson 回归、负二项回归及广义负二项回归、随机效应模型以及本书所讲的空间计量经济学模型等。

3. GeoDa 软件

GeoDa 是一个设计实现栅格数据探求性空间数据分析(ESDA)的软件工具集合体的最新成果。它向用户提供一个友好的和图示的界面用以描述空间数据分析, 如自相关性统计和异常值指示等。

GeoDa 的设计包含一个由地图和统计图表相联合的相互作用的环境, 使用强大的连接窗口技术。GeoDa 能在任何风格的微软公司的操作系统下运行, 适用于区域经济学、法学、政治学等社会学科, 以及医学、地理学、植物学、土壤学、地质学、水文学和气候学等领域。

4. R 软件

R 具有一套完整的数据处理、计算和制图软件系统。其功能包括: 数据存储和处理系统; 数组运算工具(其向量、矩阵运算方面功能尤其强大); 完整连贯的统计分析工具; 优秀的统计制图功能; 简便而强大的编程语言: 可操纵数据的输入和输出, 可实现分支、循环, 用户可自定义功能。它还提供了一些集成的统计工具, 但更重要的是它提供各种数学计算、统计计算的函数, 从而使使用者能灵活机动地进行数据分析, 甚至创造出符合需要的新的统计计算方法。该语言的语法表面上类似 C, 但在语义上是函数设计语言(functional programming language)的变种。

第2章 空间计量经济学基础

空间计量经济学起源于区域科学和计量经济学的共同发展,研究的是如何在横截面数据和面板数据中处理空间相互作用和空间结构问题,是计量经济学的一个分支。Anselin(1988)将空间计量经济学定义为:“在区域科学模型的统计分析中,研究由空间引起的各种特性的一系列方法。”本章主要介绍数据中的空间效应及其常见检验方法,以及用于表达空间关系的空间权重矩阵的性质和常用设定。

2.1 空间效应

空间计量经济学与传统计量经济学的最大区别就是引入了空间效应,空间效应是空间计量经济学的基本特征,它反映着空间因素的影响,是空间计量经济学从传统计量经济领域独立出来的根本原因。空间效应可分为空间相关性(spatial dependence,即空间依赖性)和空间异质性(spatial heterogeneity)(Anselin,1988a)。空间依赖性是指主体行为间的空间交互作用而产生的一种截面依赖性,这意味着不同区位随机变量之间的相关性或者协方差结构主要来自空间组织形式,这些空间组织形式是由地理空间中主体之间空间相对位置(距离、空间排序)决定的。空间异质性是指空间结构的非均衡性,表现为主体行为之间存在明显的空间结构性差异。

与时间序列模型引入时间滞后反映序列相关一样,在空间计量模型中,空间相关和空间异质是通过引入空间滞后来实现的。空间滞后是相应变量在邻近区域的加权平均值,它可以是因变量的滞后、自变量的滞后、误差项的滞后,还可以是三者的不同组合,属于空间平滑的一种方式。因为空间异质性很多时候可以用传统的计量经济学方法进行处理,如处理异方差性的方法,所以在本章中只关注空间相关性。

2.1.1 空间相关性

空间相关性是指空间中各变量之间存在相互影响。Goodchild(1992)指出,几乎所有的空间数据都具有空间依赖(空间自相关)特征,也就是说一个地区空间单元的某种经济地理现象或者某一属性值与邻近地区空间单元上同一现象或属性值是相关的。空间依赖是事物和现象在空间上的相互依赖、相互制约、相互影响和相互作用,是事物和现象本身所固有的属性,是地理空间现象和空间过程的本质特征。它是指不同位置的观测值在空间上非独立,呈现出某种非随机的空间模式(Lesage,1999)。一般来说,空间相关性主要表现在两个方面。

空间实质相关性:空间外部性、邻近效应等因素造成的计量模型中解释变量的空间相关性。

空间扰动相关性：忽视了一定的空间影响，如存在空间影响的因素没有被考虑在模型中，造成的模型残差存在空间相关性。

空间依赖性打破了大多数传统经典统计学和计量经济学中相互独立的基本假设，是对传统方法的继承和发展。由于空间相关性违反了经典计量模型假设中有关观测值不相关的假定前提，传统方法对独立样本的统计推断将不再有效。粗略来说，与相同大小的独立样本相比，存在空间相关性的样本将导致较大的方差估计、假设检验的低显著水平，以及估计模型较低的拟合度。简言之，空间相关性会导致数据信息失真和传统计量经济分析有偏。

2.1.2 空间相关性的检验

在处理空间数据过程中，空间相关性检验是一步非常重要的工作。如果不存在空间相关，则使用标准的计量方法即可；如果存在空间相关，则要使用空间计量方法。空间相关性检验大概分成两类：第一类，包括空间误差自相关或空间误差移动平均的误差相关检验，如 LMERR, R-LMERR；第二类，空间滞后相关检验，如 LMLAG, R-LMLAG。此外，部分统计量既可以检验对象间的空间误差相关关系，又可以检验空间滞后相关关系。例如，空间相关性 Moran's I 检验和 Geary 检验。迄今为止，Moran's I 检验是最常见的空间相关性检验方法，本节将重点介绍 Moran's I 检验和 Geary 检验。

1. 全局空间自相关指标

Moran 指数和 Geary 指数是两个用来度量空间自相关的全局指标。

1) Moran 指数

Moran 指数反映的是空间邻接或空间邻近的区域单元属性值的相似程度。如果 Y 是位置(区域)的观察值，则该变量的全局 Moran'I 值用如下公式计算：

$$\text{Moran}'I = \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} (Y_i - \bar{Y})(Y_j - \bar{Y})}{S^2 \sum_{i=1}^n \sum_{j=1}^n w_{ij}} \quad (2.1.1)$$

其中， $S^2 = \frac{1}{n} \sum_{i=1}^n (Y_i - \bar{Y})^2$ ； $\bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i$ ； Y_i 表示第 i 地区的观测值； n 为地区总数； w_{ij}

为空间权值矩阵第 i 行第 j 列元素。标准化的 Moran'I 统计量为

$$Z = \frac{\text{Moran}'I - E(I)}{\sqrt{\text{Var}(I)}} \quad (2.1.2)$$

其中， $E(I) = -\frac{1}{n-1}$ ， $\text{Var}(I) = \frac{n^2 \omega_1 + n \omega_2 + 3 \omega_0^2}{\omega_0^2 (n^2 - 1)} - E^2(I)$ ， $\omega_0 = \sum_{i=1}^n \sum_{j=1}^n w_{ij}$ ，

$\omega_1 = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (w_{ij} + w_{ji})^2$ ， $\omega_2 = \sum_{i=1}^n \sum_{j=1}^n (w_{i.} + w_{.j})$ ， $w_{i.}$ 和 $w_{.j}$ 分别为空间权值矩阵

中 i 行和 j 列之和。在不存在空间相关性的原假设下， Z 服从标准正态分布。

Moran 指数 I 的取值一般在 $[-1, 1]$ ，小于 0 表示负相关，等于 0 表示不相关，大于 0 表示正相关。越接近 -1 表示单元间的差异越大或分布越不集中；越接近 1，则代表单元

间的关系越密切,性质越相似(高值集聚或者低值集聚);接近 0,则表示单元间不相关。

2) Geary 指数

由于 Moran 指数不能判断空间数据是高值集聚还是低值集聚,Getis 和 Ord 于 1992 年提出了全局 Geary 指数。Geary 指数与 Moran 指数存在负相关关系。Geary 指数 C 计算公式如下:

$$C = \frac{(n-1) \sum_{i=1}^n \sum_{j=1}^n w_{ij} (x_i - x_j)^2}{2 \sum_{i=1}^n \sum_{j=1}^n w_{ij} \sum_{k=1}^n (x_k - \bar{x})^2} \quad (2.1.3)$$

式中, C 为 Geary 指数;其他变量同上式。Geary 指数 C 的取值一般在 $[0, 2]$, 大于 1 表示负相关,等于 1 表示不相关,而小于 1 表示正相关。也可以对 Geary 指数进行标准化:

$$Z(C) = [C - E(C)] / \sqrt{\text{Var}(C)} \quad (2.1.4)$$

式中, $E(C)$ 为数学期望, $\text{Var}(C)$ 为方差。正的 $Z(C)$ 表示存在高值集聚,负的 $Z(C)$ 表示低值集聚。

2. 局部空间自相关指标

局部空间自相关指标包括空间联系的局部指标(LISA 集聚图)和 Moran 散点图。

1) 空间联系的局部指标

LISA 包括局部 Moran 指数(local Moran)和局部 Geary 指数(local Geary)。

局部 Moran 指数被定义为

$$I_i = \frac{(x_i - \bar{x})}{S^2} \sum_j w_{ij} (x_j - \bar{x}) \quad (2.1.5)$$

正的 I_i 表示该空间单元与邻近单元的属性相似(“高一高”或“低—低”),负的 I_i 表示该空间单元与邻近单元的属性不相似(“高一低”或“低—高”)。

局部 Geary 指数由 Getis 和 Ord(1992)提出,是一种基于距离权重矩阵的局部空间自相关指标,能探测出高值集聚和低值集聚,计算公式为

$$G_i^* = \frac{\sum_j w_{ij} x_j}{\sum_k x_k} \quad (2.1.6)$$

在各区域不存在空间相关下,Getis 和 Ord 简化了 G_i^* 的数学期望和方差的表达式:

$$E(G_i^*) = \frac{\sum_j w_{ij}}{n-1} = \frac{W_i}{n-1}, \quad \text{Var}(G_i^*) = \frac{W_i(n-1-W_i)}{(n-1)^2(n-2)} \frac{Y_{i2}}{Y_{i1}^2}$$

式中, $Y_{i1} = \frac{\sum_j w_j}{n-1}$, $Y_{i2} = \frac{\sum_j x_j^2}{n-1} - Y_{i1}^2$ 。

将 G_i^* 标准化,得到

$$Z_i = \frac{G_i^* - E(G_i^*)}{\sqrt{\text{Var}(G_i^*)}} \quad (2.1.7)$$

此时,显著的正 Z_i 表示邻近单元的观测值高,显著的负 Z_i 则表示邻近单元的观测值低。

2) Moran 散点图

以 (Wz, z) 为坐标点的 Moran 散点图,常用来研究局部的空间特征。它对空间滞后因子 Wz 和 z 数据进行了可视化的二维图示。Moran's I 指数可看作各地区观测值的乘积和,其取值范围在 -1 到 1 之间,若各地区间经济行为为空间正相关,其数值应当较大;负相关则较小。当目标区域数据在空间区位上相似的同时也有相似的属性值时,空间模式整体上就显示出正的空间自相关性;而当在空间上邻接的目标区域数据不同寻常地具有不相似的属性值时,就呈现为负的空间自相关性;零空间自相关性出现在当属性值的分布与区位数据的分布相互独立时。Moran 散点图中的第 1、3 象限代表观测值的正空间相关性,第 2、4 象限代表观测值的负空间相关性,并且第 1 象限代表了观测值高的区域单元被高值区域所包围(HH);第 2 象限代表了观测值低的区域单元被高值区域所包围(LH);第 3 象限代表了观测值低的区域单元被低值区域所包围(LL);第 4 象限代表了观测值高的区域单元被低值区域所包围(HL)。

2.2 空间权重

2.2.1 空间权重及其设定方法

将空间效应纳入计量模型,是通过引入权重矩阵来实现的。建立空间计量模型进行空间统计分析时,一般要用空间权重矩阵 W 来表达 n 个位置的空间区域的邻近关系,其形式如下:

$$W = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1} & w_{n2} & \cdots & w_{nn} \end{bmatrix} \quad (2.2.1)$$

式中第 i 行第 j 列元素 w_{ij} 表示区域 i 与 j 的邻近关系。对于空间矩阵的构造,一直是一个有争议的问题。因为无法找到一个完全描述空间相关结构的空间矩阵,也就是说,理论上讲,不存在最优的空间矩阵。

Tobler 地理学第一定律指出:任何事物与别的事物之间都是相关的,但近处的事物比远处的事物的相关性更强。这是权重矩阵设置的一个基本依据。一般来讲,空间矩阵的构造必须满足“空间相关性随着‘距离’的增加而减少”的原则。这里的“距离”是广义的,可以是地理上的距离,也可以是经济意义上合作关系的远近,甚至可以是社会学意义上的人际关系亲疏。

空间权重矩阵 $W = (w_{ij})_{n \times n}$ 主要用于表达空间的相互依赖性,它是外生信息。 W 中对角线上的元素被设定为 0,而 $w_{ij} (i \neq j)$ 表示区域 i 和区域 j 空间上的紧密程度,为了减少或者消除区域间的外生影响,权重矩阵在进行其他运算前通常被标准化,矩阵 W 的行和为 1,标准化意味着每一个矩阵元素仅仅表示邻接空间的加权平均数。

2.2.2 常用的空间权重矩阵

空间矩阵的常规设定有两种：一种是简单的二进制邻接矩阵，另一种是基于距离的二进制空间权重矩阵。简单的二进制邻接矩阵的第 i 行第 j 列元素为

$$w_{ij} = \begin{cases} 1, & \text{当区域 } i \text{ 和区域 } j \text{ 相邻接} \\ 0, & \text{其他} \end{cases} \quad (2.2.2)$$

邻接性空间依赖关系相对应空间权重矩阵元素值的定义是：邻接为 1，不邻接为 0。参照国际象棋中棋子的行子规则，相邻关系可以分为以下三种。

“车式”邻接(rook contiguity)：两个相邻区域有共同的边。

“象式”邻接(bishop contiguity)：两个相邻区域有共同的顶点，但没有共同的边。

“后式”邻接(queen contiguity)：两个相邻区域有共同的边或顶点。

常用相邻关系如图 2.2.1 所示。

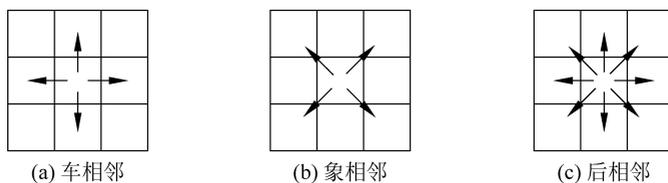


图 2.2.1 常用相邻关系

在实践中，为了区分边与点，可以设定一个最小距离，在此距离以下为点，而在此距离以上为边。另外究竟使用车、象还是后相邻，取决于具体情况。例如，区域 i 与区域 j 仅在一点相交(象相邻)，但有一条主要高速公路通过此点连接两区域，则不宜使用车相邻。

与邻接性相比，距离性空间依赖关系的空间权重矩阵元素值定义要复杂得多。从距离变量属性来看，可将空间分为非几何距离和几何距离两种。

基于距离的二进制邻接矩阵的第 i 行第 j 列元素为

$$w_{ij} = \begin{cases} 1, & \text{当区域 } i \text{ 和区域 } j \text{ 的距离小于 } d \text{ 时} \\ 0, & \text{其他} \end{cases} \quad (2.2.3)$$

在实际的区域分析中，空间权重矩阵的选择设定是外生的，原因是 $(n \times n)$ 阶矩阵 W 包含了关于区域 i 和区域 j 之间相关的空间邻接的外生信息，不需要通过模型来估计得到它。权重矩阵中对角线上的元素 w_{ii} 被设定为 0。为了减少或消除区域间的外在影响，权重矩阵被标准化 $(w_{ij}^* = w_{ij} / \sum_{k=1}^n w_{ik})$ ，使得行元素之和为 1。

1. 基于邻近概念的空间权重矩阵

基于邻近概念的空间权重矩阵(contiguity based spatial weights)有一阶邻近矩阵(the first order contiguity matrix)和高阶邻近矩阵两种。

(1) 一阶邻近矩阵。一阶邻近矩阵是假定两个地区有共同边界时空间关联才会发生，即当相邻地区 i 和 j 有共同边界时用 1 表示，否则以 0 表示。一般有 Rook 邻近和

Queen 邻近两种计算方法 (Anselin, 2003)。Rook 邻近用仅有共同边界来定义邻居, 而 Queen 邻近则除了共有边界邻居外还包括共同顶点的邻居。由此可见, 基于 Queen 邻近的空间矩阵常常与周围地区具有更加紧密的关联结构 (拥有更多的邻居)。当然, 如果假定区域间公共边界的长度不同 (如 10km 和 100km), 其空间作用的强度也不一样, 则还可以通过将公用边界的长度纳入权重计算过程中, 使这种邻近指标更加准确一些。

(2) 高阶邻近矩阵。空间权重矩阵不仅仅局限于第一阶邻近矩阵, 也可以计算和使用更高阶的邻近矩阵。Anselin 和 Smirnov (1996) 提出了高阶邻近矩阵的算法。二阶邻近矩阵 (the second order contiguity matrix) 表示了一种空间滞后的邻近矩阵。也就是说, 该矩阵表达了邻近的相邻地区的空间信息。当使用时空数据并假设随着时间推移产生空间溢出效应时, 这种类型的空间权重矩阵将非常有用, 在这种情况下, 特定地区的初始效应或随机冲击不仅会影响其邻近地区, 而且随着时间的推移会影响其邻近地区的相邻地区。当然, 这种影响是呈几何递减的。

可以看出, 邻近空间权重矩阵因其对称和计算简单而较为常用, 适合于测算地理空间效应的影响。

2. 其他经济社会空间权重矩阵

除了使用真实的地理坐标计算地理距离外, 还有包括经济和社会因素的更加复杂的权重矩阵设定方法。例如, 根据区域间交通运输流、通信量、GDP (国内生产总值) 总额、贸易流动、资本流动、人口迁移、劳动力流动等确定空间权值, 计算各个地区任何两个变量之间的距离。例如以下几种。

1) 基于万有引力定律的空间邻接矩阵

近年来, 一些学者采用诺贝尔经济学奖获得者 Tinbergen (1962) 提出的引力模型研究区域贸易问题, 该模型直接把地区间的距离作为解释变量引入模型中。其思想源自物理学中的万有引力定律, 即两个物体之间的引力与它们的质量乘积成正比, 与它们之间的距离平方成反比。尽管引力模型已经得到了广泛应用, 但是基于万有引力定律构造空间邻接矩阵并不多见。我们认为, 技术溢出效应是广泛存在的, 而不仅仅局限于有共同边界的地区之间, 并且两个地区之间的经济实力越强, 技术交流与合作的吸引力往往越大, 相应的技术溢出效应也越大。为此, 基于万有引力定律构建如下空间邻接矩阵 $\mathbf{W}=(w_{ij})$:

$$w_{ij} = \begin{cases} \frac{m_i m_j}{r_{ij}^2}, & i \neq j \\ 0, & i = j \end{cases} \quad (2.2.4)$$

其中, r_{ij} 为地区 i 与地区 j 的地理距离, 可由两个地区的经纬度计算; m_i 为地区 i 的经济实力, 以样本期内的人均实际 GDP 衡量。为了消除单位选取的影响, 邻接矩阵需要标准化使行元素之和为 1。

2) 基于地理距离标准构造空间权重矩阵

空间邻接标准认为空间单元之间的联系仅仅取决于二者相邻与否, 即只要不同空间单元相邻, 则认为它们之间具有相同的影响强度, 这在区域创新经济研究中是不符合客观事实的。例如, 用空间邻接标准衡量的区域的地理位置, 与北京邻接的只有天津、河北两