空间面板数据模型及其应用研究

胡亚权 著

Kongjian Mianban Shuju Moxing Jiqi Yingyong Yanjiu

WUHAN UNIVERSITY PRESS
武汉大学出版社

课题研究获国家自然科学青年基金项目（71503073）支持

空间面板数据模型及其应用研究

胡亚权　著

Kongjian Mianban Shuju Moxing
Jiqi Yingyong Yanjiu

WUHAN UNIVERSITY PRESS
武汉大学出版社

图书在版编目(CIP)数据

空间面板数据模型及其应用研究/胡亚权著．—武汉：武汉大学出版社，2019.12

ISBN 978-7-307-21222-0

Ⅰ.空…　Ⅱ.胡…　Ⅲ.经济模型—研究　Ⅳ.F224.0

中国版本图书馆 CIP 数据核字(2019)第 244194 号

责任编辑：林　莉　喻　叶　　　责任校对：汪欣怡　　　版式设计：马　佳

出版发行：**武汉大学出版社**　（430072　武昌　珞珈山）

（电子邮箱：cbs22@whu.edu.cn　网址：www.wdp.com.cn）

印刷：北京虎彩文化传播有限公司

开本：720×1000　1/16　印张：5.75　字数：116 千字　插页：1

版次：2019 年 12 月第 1 版　　2019 年 12 月第 1 次印刷

ISBN 978-7-307-21222-0　　定价：29.00 元

前　言

Anselin(2006)将空间计量经济学定义为计量经济学的一个分支，处理与区位、距离和空间位置相关的空间变量有关的横截面模型或时空模型及其模型设定、参数估计、假设检验和预测，最初的研究都是基于横截面数据。近些年来，有关空间面板的模型设定和估计方法越来越多地出现在空间计量经济学文献中。本书的目的就是梳理空间计量经济学模型的设定、估计方法的发展过程，介绍空间面板数据模型最新的研究进展，并将空间面板数据模型应用于中国的经济研究。

本书共分为七章，第一章是本书的绪论，介绍了研究背景、研究内容及本书的创新点；第二章介绍了空间计量经济学的发展过程；第三章对静态和动态空间面板数据模型的模型设定、参数的估计方法和假设检验进行了文献综述，并介绍了空间面板数据模型最近的一些研究进展；第四章分析了中国地区之间的知识溢出效应。第五章对影响中国地区间失业率差异的因素进行了空间计量分析；第六章分析了中国地区间经济增长与二氧化碳排放的空间关系；第七章用空间面板数据模型分析了中国地区之间气候变化、农业生产条件与农业净收益的关系。

本书在编写过程中引用了相关专家学者的研究成果，虽然已有标注和说明，但不可避免存在遗漏之处，谨向他们表示感谢！由于时间和水平的限制，书中疏漏、不妥之处，恳请广大读者批评指正。

作　者

2019 年 7 月

目　　录

1 绪　论

1.1 研究背景、目的和意义

自从 Paelinck and Klaassen（1979）出版《空间计量经济学》（*Spatial Econometrics*），已经过去三十几年了，空间计量经济学经历了启蒙—发展—成熟三个阶段。早期的理论和方法上的进展主要出现在区域科学和数量地理杂志，主要的经济学和计量经济学期刊文献缺乏空间视角。形势发生改变也是在近些年，一般性的空间分析和专业性的空间计量分析几乎呈指数增长，特别是在社会科学领域（Goodchild et al. 2000；Bivand，2008）。总之，计量经济学中对空间效应的关注不再是模糊不清的，但是在理论研究和实证应用上缺乏系统性，主要原因在于越来越多的地理数据的出现和利用，以及界面友好的地理信息系统对地理数据的处理。同样重要的是，理论上越来越关注空间视角，经济学中代理人互动模型、社会相互作用模型改变了传统经济学理论中的行为方式，经典假设之一的数据相互独立性得到放松，尤其是空间效应可以在方程的右边包含因变量的空间滞后项，解释变量的空间滞后项，以及空间误差项也可以考虑空间滞后项，从而形成了空间计量经济学的研究和发展，本书的目的就是梳理空间计量经济学模型的设定、估计方法的发展过程，并将空间面板数据模型应用于中国的经济研究。

1.2 本书的研究内容及框架结构

空间横截面计量经济学模型的理论研究以及模型应用已经广泛存在于国内外文献中，空间面板数据模型的理论研究及其实证分析越来越多地引起国外学者的广泛关注，国内学者开始将空间面板数据模型应用于中国的经济研究，本书的主要内容是对空间计量经济学的发展过程进行概述，然后对空间面板数据模型的模型设定、参数估计以及假设检验进行文献综述，最后将空间面板数据模型应用于我国的几个经济问题研究。

本书分为 7 章，第 1 章是本书的绪论，介绍了本书的研究背景、研究内容及本书的创新点；第 2 章介绍了空间计量经济学的发展过程；第 3 章对静态和动态空间

面板数据模型的模型设定、参数的估计方法和假设检验进行了文献综述，并介绍了空间面板数据模型最近的一些研究进展；第 4 章分析了中国地区之间的知识溢出效应；第 5 章对影响中国地区间失业率差异的因素进行了空间计量分析；第 6 章分析了中国地区间经济增长与二氧化碳排放的空间关系；第 7 章用空间面板数据模型分析了中国地区之间气候变化、农业生产条件与农业净收益的关系。

1.3　本书的创新点

(1)理论上回顾了空间计量经济学的发展过程，以及空间面板数据模型的模型设定、估计方法以及假设检验问题，综述了现在的一些研究进展。

(2)应用上，将空间面板数据模型应用于以下中国经济领域。

①将知识资本分为可衡量的知识资本和不可衡量的知识资本。可衡量的知识资本存量用专利申请授权数作为代理变量，假设可衡量的知识资本与不可衡量的知识资本均服从空间自回归过程，以 C-D 生产函数为桥梁，推导出了全要素生产率与可衡量知识资本存量的空间计量经济学模型。关于知识资本存量与全要素生产率之间关系的已有文献更多地采用企业、高校等微观个体的数据，本书选取全国 30 个地区 1996—2010 年的面板数据，用 SDM 模型分析了知识资本存量对全要素生产率的溢出效应，结果表明，这种地区之间的溢出效应是存在的。因此，我们应该鼓励地区之间的合作，鼓励人才和技术在地区之间的流动，同时提高本地区的学习吸收能力，从而减少地区之间知识资本之间的差距，这对缩小地区之间的收入差距是很有帮助的。

②影响地区失业率的因素和失业率一样，呈现空间集聚现象，这意味着地区失业率的空间演化是各地区城市化特征不同的结果，因为各地区的经济，社会人口属性也存在空间集聚效应。基于我国各地区 1999—2010 年的面板数据，使用空间计量经济学方法分析了地区失业率差异性的决定因素。正如 Elhorst(2010) 和 Fingleton and Gallo(2010) 所认为的，当模型的内生性问题是因空间滞后变量的省略所引起的情况下，空间计量经济学方法可以给出系数的一致性估计，并能正确估计地区之间一般均衡效应的大小。分析结果表明，就业增长、第一、二产业就业人数、受教育程度是影响失业率地区差异的重要因素，这些因素的变化不仅影响本地劳动力市场，还通过溢出效应影响相邻地区。

③一个国家或地区的环境质量指标是否受相邻国家或地区人均污染物排放量的影响，空间计量经济学方法为该问题提供了研究的工具。当我们使用与研究个体地理位置相关的数据集时，如果忽视本已存在的空间依赖性，那么惯常的估计结果将是有偏且非一致的。本书选取中国 29 个地区 1995—2009 年的面板数据，研究了人均 CO_2 排放量与人均 GDP 之间的关系，在考虑了地区之间 CO_2 排放量的空间相关

性的情况下，估计结果表明 EKC 曲线是存在的，进一步根据估计的参数计算出 EKC 曲线的转折点为人均收入等于 73630 元。2011 年我国有些地区的人均 GDP 已经达到这个临界点，这说明我国的经济发展已经到了从注重 GDP 转向关注环境质量的阶段，这和我国的国情是相符合的。

④本书利用 1997—2009 年我国 31 个省份的面板数据，分析了气候变化和非气候因素对我国农业净收益的影响情况。结果表明，我国 31 个省份的农业净收益之间存在正的空间自相关效应，气候变量中，夏季气温的升高会减少农业净收益，但是夏季气温和降雨量共同对农业净收益有促进作用，这说明夏季气候变化比较极端，只有气温和降雨量比较协调才能提高农业净收益。另外，非气候因素(化肥施用量，有效灌溉面积，机械化程度和人口密度)都和农业净收益正相关。基于上述结果，本书最后从水利基础设施，农业综合生产能力和调整农业结构三个方面提出了政策建议。

2 空间计量经济学概述

作为计量经济学的分支之一，空间计量经济学是一门处理横截面或面板数据回归模型中的空间相互作用(spatial autocorrelation，空间自相关)和空间结构问题(spatial heterogeneity，空间异质性)的学科。过去，空间计量经济学的应用主要集中在以下领域：区域科学、城市和房地产经济学、经济地理学。近些年来，空间计量经济学大量应用于传统经济学领域：需求分析研究、国际经济学、劳动经济学、公共经济学、农业和环境经济学。同时，越来越多方法论方面的文献开始研究空间计量模型的模型设定、估计方法以及假设检验等问题。这种关注可以归结为两方面的因素：一方面是理论经济学模型中出现的代理人相互作用机制，马歇尔外部性以及集聚经济等模型对计量经济学提出了新的要求；另一方面，实践中出现的空间数据(房地产经济学、资源环境经济学、发展经济学中空间数据的使用)也需要新的估计方法。

空间计量经济学(spatial econometrics)这一概念是由 Paelinck 在 1974 年提出的，并在其与 Klaassen 合著的 *Spatial Econometrics*(1979)中指出空间计量经济学是用来处理多区域模型中空间关系的一种方法。但是 Paelinck 并没有给空间计量经济学下定义，至少给出了模型设定的五个主要原则：①空间依赖性的作用；②空间关系的非对称性；③空间距离解释因素的重要性；④事前和事后相互作用的差别；⑤建模中的空间因素的显著重要性。Anselin(1988)发表的 *Spatial Econometrics: Methods and Models* 成为空间计量经济发展的里程碑，Anselin 将空间计量经济学定义为，在区域科学模型中由空间因素的特殊性质而引出的一系列特殊的处理方法。不管是前者给出的建模原则还是后者给出的初步定义，都没能想到空间计量经济学方法还越来越多地应用到经济学以及其他社会科学中。Anselin(2006)给出了新的定义：空间计量经济学是计量经济学的一个分支，处理与区位、距离和空间位置相关的空间变量有关的横截面模型或时空模型及其模型设定、参数估计、假设检验和预测。

具体来说，空间计量经济学处理两个方面的空间效应(spatial effects)：空间相关性(spatial dependence)和空间异质性(spatial heterogeneity)。

2.1　基本概念

2.1.1　空间效应

空间回归分析中，空间效应指的是空间依赖性和空间异质性。空间依赖性是指地理空间上的观测值缺乏相互独立性，这种相对位置可以是距离上的，也可以是空间位置上的。空间依赖性被认为是区域科学和地理学的核心，被表述成地理学第一定律：任何事物都和周围事物相联系，距离越近，联系越紧密(Tobler，1979)。从这种意义上讲，空间依赖性由相对空间或相对位置来决定，强调距离效应。空间依赖性来源之一是测量问题，在应用区域科学的绝大多数实证分析中，数据来源于空间单元基于坐标或距离的观测值。类似的例子出现在按照行政单位(例如省份、县市、调查跟踪点)划分的人口数据、就业数据以及其他的经济活动数据中。空间依赖性的来源之二是对人类行为解释的空间因素。区域科学和人文地理学的本质是指与位置和距离相关的问题，这些问题导致各种空间上的相互依赖性，即空间上点与点之间发生的事情相互关联。与时间序列的相关性不一样，空间依赖性需要特殊的一套方法和技巧。空间依赖性用公式表示如下：

$$\operatorname{cov}(y_i,\ y_j)=E(y_iy_j)-E(y_i)E(y_j)\neq 0,\qquad i\neq j \tag{2.1}$$

其中，i、j 指不同的任意两个空间位置，y_i、y_j 是指对应位置上的观测值。

在区域科学和经济地理学的文献中，有充分证据表明空间效应存在异质性。诸多因素，像中心位置层次化，先进与落后地区共存等问题需要考虑每个空间单元的特质，计量经济学中用时变参数、随机系数或结构变动等方法来建模。除了结构稳定性的缺乏，观测值对应的空间单元自身之间也存在异质性。例如，空间单元在面积和形状上存在差异，人口和收入的不对等。这种异质性程度反映在导致异方差的测量误差上。简单来讲，空间异质性指模型误差项方差或变量系数形式上结构不稳定性。正式地，假设集合 S 包含 N 个地理单元(例如：省份、县市或调查跟踪点)，S 被分成 R 个互不重叠的子集 $S_r(r=1,\ 2,\ \cdots,\ R)$，即对任意的 r、$s(r\neq s)$，$S_r\cap S_s=\varnothing$ 且 $\bigcup_{r=1}^{R}S_r=S$。如果当 $i\in S_r$ 时，$Var(\varepsilon_i)=\sigma_r^2$，即当观测值的误差方差呈现空间集聚差异时，称为空间组际异质性；类似地，如果当 $i\in S_r$ 时，$\beta_i=\beta_r$，即回归系数随观测值所在子集的变化而变化，称之为斜率系数异质性。考察空间异质性的重要性在于：①不稳定结构某种意义上来看是空间的，观测值的位置是决定其形式的关键因素；②由于上述结构的不稳定性，空间异质性通常还会伴随着空间自相关性，传统的计量经济学方法此时无能为力；③横截面模型中，空间自相关和空间异质性可能拥有共同的观测值。

2.1.2 空间权重矩阵

空间权重矩阵(spatial weights matrix)是一个 n 阶方阵 $W\,(w_{ij})_{n\times n}$，通过 W 来设定观测值之间的位置关系，如果 $w_{ij}\neq 0$，则表明观测值 i，j 存在空间上的相邻关系，习惯上我们假定观测值自身之间不相邻，即 $w_{ij}=0$。为了模型解释上的方便，通常需要对权重矩阵的元素进行标准化处理：$w_{ij}^{s}=w_{ij}\Big/\sum\limits_{j}w_{ij}$。在模型设定中，没有严格的原则来指导权重矩阵的选择，比较典型的做法是基于地理标准来确定相邻关系。

(1)0-1 相邻矩阵(binary contiguity matrix)：①rook 标准：如果空间单元 i 与 j 拥有共同的边界，则 $w_{ij}=1$，否则 $w_{ij}=0$；②bishop 标准：如果空间单元 i 与 j 拥有共同的顶点，则 $w_{ij}=1$，否则 $w_{ij}=0$；③queen 标准：如果空间单元 i 与 j 拥有共同的边界或共同的顶点，则 $w_{ij}=1$，否则 $w_{ij}=0$。

如图 2-1 所示，对于空间单元 3 来说，按照 rook 标准，4 和 5 都是其邻居；按照 bishop 标准，只有 2 是其邻居；按照 queen 标准，2，4，5 都是其邻居。

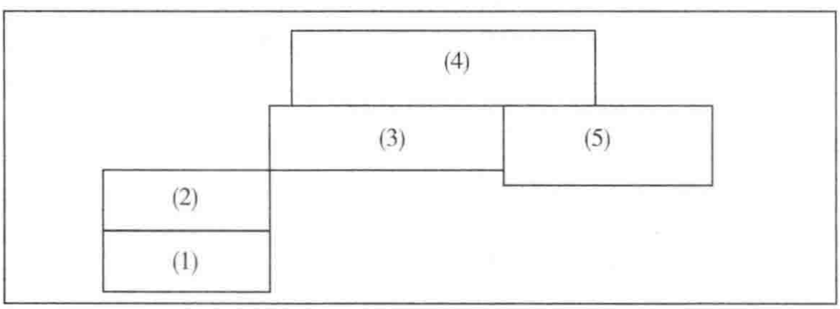

图 2-1　空间相邻矩阵图示说明

(2)距离矩阵：给定一个距离标准，如果空间单元 i 与 j 在这个距离标准之内，那么 $w_{ij}=1$，否则 $w_{ij}=0$。例如：k 最近邻(k-nearest)矩阵，对于空间单元 i 来说，与 i 的距离在某个给定标准之内的空间单元都是为 i 的邻居。

2.2 空间横截面模型设定

在经典线性回归模型中加入被解释变量的空间滞后项 Wy，就称之为空间滞后回归模型(spatial lag regression model，简称 SLM)或空间自回归模型(spatial autoregressive regression model，简称 SAR)，回归方程及其数据生成过程如下所示：

$$y=\rho Wy+\alpha\iota_n+X\beta+\varepsilon \tag{2.2}$$

$$y=(I_n-\rho W)^{-1}(\alpha\iota_n+X\beta)+(I_n-\rho W)^{-1}\varepsilon,\ \varepsilon\sim N(0,\sigma^2 I_n) \tag{2.3}$$

其中，y 是 n 维被解释变量列向量，X 是 $n\times k$ 阶解释变量矩阵，ι_n 是元素均为 1 的 n 维列向量，ρ 是空间自相关系数，是一标量，α，β 是模型的参数向量，ε 为随机扰动项，W 是 $n\times n$ 阶空间权重矩阵。

如果再在空间滞后回归模型中加入解释变量的空间滞后项 WX，则称之为空间 Durbin 模型(Spatial Durbin Model，简称 SDM)，回归方程及其数据生成过程如下所示：

$$y=\rho Wy+\alpha\iota_n+X\beta+WX\gamma+\varepsilon \tag{2.4}$$

$$y=(I_n-\rho W)^{-1}(\alpha\iota_n+X\beta+WX\gamma)+(I_n-\rho W)^{-1}\varepsilon,\ \varepsilon\sim N(0,\sigma^2 I_n) \tag{2.5}$$

如果在经典线性回归模型中考虑随机扰动的空间滞后项，则称之为空间误差模型(spatial error model，简称 SEM)，回归方程如下所示：

$$y=\alpha\iota_n+X\beta+u,\ u=\lambda Wu+\varepsilon,\ \varepsilon\sim N(0,\sigma^2 I_n) \tag{2.6}$$

既包含被解释变量的空间滞后项又包括随机扰动项的空间滞后项的模型称之为 SAC 模型：

$$y=\alpha\iota_n+\rho W_1 y+X\beta+u,\ u=\theta W_2 u+\varepsilon,\ \varepsilon\sim N(0,\sigma^2 I_n) \tag{2.7}$$

$$y=(I_n-\rho W_1)^{-1}(\alpha\iota_n+X\beta)+(I_n-\rho W_1)^{-1}(I_n-\theta W_2)^{-1}\varepsilon \tag{2.8}$$

空间移动平均过程 $u=(I_n-\theta W)\varepsilon$ 与 SLM 模型合并之后可以生成新的模型，空间自回归移动平均过程(spatial autoregressive moving average model，简称 SARMA)：

$$y=\alpha\iota_n+\rho W_1 y+X\beta+u,\ u=(I_n-\theta W_2)\varepsilon,\ \varepsilon\sim N(0,\sigma^2 I_n) \tag{2.9}$$

$$y=(I_n-\rho W_1)^{-1}(\alpha\iota_n+X\beta)+(I_n-\rho W_1)^{-1}(I_n-\theta W_2)\varepsilon$$

2.3　参数估计值的解释

空间回归模型能够揭示空间单元观测值之间的空间依存关系，任何一个空间单元解释变量观测值的改变，不仅会对自身造成影响(direct impact，直接影响)，还会对其他相邻空间单元造成影响(indirect impact，间接影响)，这是和经典计量经济学模型完全不同的地方。

线性回归模型能够通过因变量对解释变量求导数来解释参数的意义，在包含因变量或解释变量的空间滞后项的模型中，参数的意义变得丰富和复杂，解释起来困难一些，下面以 SDM 模型来加以说明。

由 $(I_n-\rho W)y=X\beta+WX\theta+\iota_n\alpha+\varepsilon$ 可得：

$$y=\sum_{r=1}^{k}S_r(W)x_r+V(W)\iota_n\alpha+V(W)\varepsilon \tag{2.10}$$

其中：$S_r(W)=V(W)(I_n\beta_r+W\theta_r)$，$V(W)=(I_n-\rho W)^{-1}=I_n+\rho W+\rho^2W^2+\rho^3W^3+\cdots$

为了说明 $S_r(W)$ 的作用，将式(2.10)展开成下面的形式：

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \sum_{r=1}^{k} \begin{pmatrix} S_r(W)_{11} & S_r(W)_{12} & \cdots & S_r(W)_{1n} \\ S_r(W)_{21} & S_r(W)_{22} & \cdots & S_r(W)_{21} \\ \vdots & \vdots & \ddots & \vdots \\ S_r(W)_{n1} & S_r(W)_{n2} & \cdots & S_r(W)_{nn} \end{pmatrix} \begin{pmatrix} x_{1r} \\ x_{2r} \\ \vdots \\ x_{nr} \end{pmatrix} + V(W)\iota_n\alpha + V(W)\varepsilon \tag{2.11}$$

方程组(2.11)的代表性单个方程可以表示如下：

$$y_i = \sum_{r=1}^{k} [S_r(W)_{i1}x_{1r} + S_r(W)_{i2}x_{2r} + \cdots + S_r(W)_{in}x_{nr}] + V(W)_i\iota_n\alpha + V(W)_i\varepsilon \tag{2.12}$$

其中：$S_r(W)_{ij}$ 表示矩阵 $S_r(W)$ 的第 i 行第 j 列上的元素，$V(W)_i$ 表示 $V(W)$ 的第 i 列。

由式(2.12)可以看出，y_i 对 x_{jr} 的偏导数一般情况下是不等于零的，而是由矩阵 $S_r(W)$ 的第 i 行第 j 列上的元素 $S_r(W)_{ij}$ 决定，即：

$$\frac{\partial y_i}{\partial x_{jr}} = S_r(W)_{ij} \tag{2.13}$$

一般来说，解释变量的变化带来的影响会因空间单元的不同而不一样，Pace and LeSage (2006)建议用一个概括性的度量标准来衡量这些不同的影响。

(1)平均直接影响(Average Direct Impact)：用 x_{ir} 表示解释变量 x_r 的第 i 个空间单元的观测值，x_{ir} 的变化对 y_i 的影响可以表示为：

$$\overline{M}(r)_{direct} = S_r(W)_{ii} = n^{-1}tr(S_r(W))。$$

(2)平均总影响(Average Total Impact)：不管是 $S_r(W)$ 第 i 行元素之和的变化对 y_i 的总影响，还是 $S_r(W)$ 第 j 列元素之和的变化对 y_i 的总影响，数值上都等于 $n^{-1}\iota'_n S_r(W)\iota_n$，称之为解释变量的变化对 y_i 的平均总影响，记为 $\overline{M}(r)_{total} = n^{-1}\iota'_n S_r(W)\iota_n$。

(3)平均间接影响(Average Indirect Impact)：平均总影响是 y_i 对所有 x_{jr} 的平均偏导数，平均直接影响是 y_i 对自身 x_{ir} 的平均偏导数，平均间接影响所有 y_i 对相邻 x_{jr} 的平均偏导数。所以平均间接影响 $\overline{M}(r)_{indirect} = \overline{M}(r)_{total} - \overline{M}(r)_{direct}$。

2.4 参数估计

2.4.1 极大似然估计

对于空间回归模型来说，最小二乘法会导致空间滞后模型参数的非一致性估

计，也会导致空间系数以及标准误差的非一致估计，相比较而言，基于极大似然估计方法的参数估计可以得到一致性结果(Lee，2004)。

1. SLM 模型和 SDM 模型

式(2.4)及式(2.5)表示的 SDM 模型可以转化为 SLM 模型，令 $Z=[\iota_n\ X\ WX]$，$\delta=[\alpha\ \beta\ \gamma]$，式(2.4)、式(2.5)就转换为式(2.14)、式(2.15)，这意味着 SLM 模型和 SDM 模型的似然函数形式是一样的。

$$y=\rho Wy+Z\delta+\varepsilon \tag{2.14}$$

$$y=(I_n-\rho W)^{-1}Z\delta+(I_n-\rho W)^{-1}\varepsilon,\ \varepsilon\sim N(0,\ \sigma^2 I_n) \tag{2.15}$$

由式(2.14)可知，如果 ρ 的真实值已知为 ρ^*，那么就有下式：

$$y-\rho^* Wy=Z\delta+\varepsilon \tag{2.16}$$

显然 δ 的 OLS 估计值 $\hat{\delta}=(Z'Z)^{-1}Z'(I_n-\rho^* W)y$，随机扰动项的方差 $\hat{\sigma}^2=n^{-1}e(\rho^*)'e(\rho^*)$，其中，$e(\rho^*)=y-\rho^* Wy-Z\hat{\delta}$。因此，首要的任务是得到 ρ 的极大似然估计。

SLM 模型(SDM 模型)的完全对数似然函数见式(1.17)(Anselin，1988)，ω 为空间权重矩阵 W 的 n 维特征值向量。

$$\ln L=-(n/2)\ln(\pi\sigma^2)+\ln|I_n-\rho W|-\frac{e'e}{2\sigma^2} \tag{2.17}$$

$$e=y-\rho Wy-Z\delta,\ \rho\in(\min(\omega)^{-1},\ \max(\omega)^{-1})$$

完全对数似然函数等价于只有未知参数 ρ 的紧凑型对数似然函数，Pace and Barry (1997)提出了这种转化方法式(2.18)，κ 是不依赖于 ρ 的常数项，$|I_n-\rho W|$ 是相应矩阵的行列式，$e(\rho)$，$\ln L(\rho)$ 都可以看成关于 ρ 的函数。

$$\ln L(\rho)=\kappa+\ln|I_n-\rho W|-(n/2)\ln(S(\rho)) \tag{2.18}$$

$$S(\rho)=e'(\rho)e(\rho)=e'_0e_0-2\rho e'_0e_d+\rho^2 e'_de_d$$

$$e(\rho)=e_0-\rho e_d,\ e_0=y-Z\delta_0,\ e_d=Wy-Z\delta_d$$

$$\delta_0=(Z'Z)^{-1}Z'y,\ \delta_d=(Z'Z)^{-1}Z'Wy$$

为了简化上述似然函数的优化过程，Pace and Barry (1997)建议在区间 $[\rho_{\min},\rho_{\max}]$ 上用关于 ρ 的 q 维向量来估计似然函数。

$$\begin{pmatrix}\ln L(\rho_1)\\ \ln L(\rho_2)\\ \vdots\\ \ln L(\rho_q)\end{pmatrix}=\kappa+\begin{pmatrix}\ln|I_n-\rho_1 W|\\ \ln|I_n-\rho_2 W|\\ \vdots\\ \ln|I_n-\rho_q W|\end{pmatrix}-(n/2)\begin{pmatrix}\ln(S(\rho_1))\\ \ln(S(\rho_2))\\ \vdots\\ \ln(S(\rho_q))\end{pmatrix} \tag{2.19}$$

一旦得到 ρ 的极大似然估计值 $\hat{\rho}$，我们就可以得到变量系数和随机扰动项方差的估计值以及相关的方差协方差矩阵。

$$\hat{\delta} = \delta_0 - \hat{\rho}\delta_d \tag{2.20}$$

$$\hat{\sigma}^2 = n^{-1}S(\hat{\rho}) \tag{2.21}$$

$$\hat{\Omega} = \hat{\sigma}^2 \left[(I_n - \hat{\rho}W)'(I_n - \hat{\rho}W) \right]^{-1} \tag{2.22}$$

2. SEM 模型

SEM 模型的表达式即数据生成过程如式(2.23)和式(2.24)所示,

$$y = X\beta + u, \tag{2.23}$$

$$u = \lambda Wu + \varepsilon,\ \varepsilon \sim N(0,\ \sigma^2 I_n)$$

$$y = X\beta + (I_n - \lambda W)^{-1}\varepsilon \tag{2.24}$$

完全对数似然函数如下所示:

$$\ln L = -(n/2)\ln(\pi\sigma^2) + \ln|I_n - \lambda W| - \frac{e'e}{2\sigma^2} \tag{2.25}$$

$$e = (I_n - \lambda W)(y - X\beta)$$

对于给定的 λ, 对数似然函数的优化表明(Ord, 1975; Anselin, 1988):

$\beta(\lambda) = (X(\lambda)'X(\lambda))^{-1}X(\lambda)'y(\lambda)$, 其中, $X(\lambda) = X - \lambda WX$, $y(\lambda) = y - \lambda Wy$,

$$\sigma^2(\lambda) = e(\lambda)'e(\lambda)/n,\ e(\lambda) = y(\lambda) - X(\lambda)\beta(\lambda)$$

因此, 我们可以得到关于 λ 的紧凑型对数似然函数:

$$\ln L(\lambda) = \kappa + \ln|I_n - \lambda W| - (n/2)\ln(S(\lambda)) \tag{2.26}$$

$$S(\lambda) = e'(\lambda)e(\lambda) \tag{2.27}$$

与 SLM 或 SDM 模型不一样的是, $S(\lambda)$ 不是空间参数的简单二次型, 它的计算需要随着 λ 的每一次选择, 都要对矩阵进行新的变换, 解决的办法是需要一些 $k \times k$ 阶或更小的 moment 矩阵。

$$A_{XX}(\lambda) = X'X - \lambda X'WX - \lambda X'W'X + \lambda^2 X'W'WX$$

$$A_{Xy}(\lambda) = X'y - \lambda X'Wy - \lambda X'W'y + \lambda^2 X'W'Wy$$

$$A_{yy}(\lambda) = y'y - \lambda y'Wy - \lambda y'W'y + \lambda^2 y'W'Wy$$

$$\beta(\lambda) = A_{XX}(\lambda)^{-1}A_{Xy}(\lambda)$$

$$S(\lambda) = A_{yy}(\lambda) - \beta(\lambda)'A_{XX}(\lambda)\beta(\lambda)$$

这些 moment 矩阵和紧凑型对数似然函数中的 λ 是同时变化的, 应用单变量优化方法得到 $\hat{\lambda}$ 之后, 将其代入 $\sigma^2(\lambda)$, $\beta(\lambda)$, $\Omega(\lambda)$ 中就可以得到相应的估计值。

$$\hat{\beta} = \beta(\hat{\lambda}) \tag{2.28}$$

$$\hat{\sigma}^2 = n^{-1}S(\hat{\lambda}) \tag{2.29}$$

$$\hat{\Omega} = \hat{\sigma}^2 \left[(I_n - \hat{\lambda}W)'(I_n - \hat{\lambda}W) \right]^{-1} \tag{2.30}$$

需要注意的是, SDM 模型嵌套于 SEM 模型中, SEM 模型的变化如下:

$$y = \alpha\iota_n + X\beta + (I_n - \lambda W)^{-1}\varepsilon$$
$$(I_n - \lambda W)y = \alpha(I_n - \lambda W)\iota_n + (I_n - \lambda W)X\beta + \varepsilon$$
$$y = \lambda Wy + \alpha(I_n - \lambda W)\iota_n + X\beta + WX(-\beta\lambda) + \varepsilon \tag{2.31}$$

如果 $-\beta\lambda$ 统计上是显著的，式(2.31)就是一 SDM 模型。

2.4.2 矩估计和工具变量法

1. 空间二阶段最小二乘法(spatial two stage least squares)

极大似然估计之外常用的估计方法就是矩估计，包括工具变量法(IV)、广义矩估计(GMM)等，这种方法不需要正态性假定，这是与极大似然估计截然不同的地方。

令 $Z=[Wy, X]$，$\gamma=[\rho, \beta]$，SLM 模型可以写成式(2.32)，这可以看成包含内生变量 Wy 和外生变量 X 的线性回归模型

$$y = Z\gamma + \varepsilon \tag{2.32}$$

可以用工具变量法来解决内生性问题，在 2SLS 方法里，第一阶段 Z 对 Q 回归并得到 Z 的拟合值 $\hat{Z}$，Q 是包含外生变量 X 的工具变量矩阵；

$$\hat{Z} = Q(Q'Q)^{-1}Q'Z \tag{2.33}$$

第二阶段用 $\hat{Z}$ 替代式(2.32)中的 Z，得到参数 γ 的空间 2SLS 估计。

$$\hat{\gamma}_{2SLS} = [\hat{Z}'\hat{Z}]^{-1}\hat{Z}'y = [Z'Q(Q'Q)^{-1}Q'Z]^{-1}Z'Q(Q'Q)^{-1}Q'y \tag{2.34}$$

对 $\hat{\gamma}_{2SLS}$ 的统计推断基于如下的渐进方差矩阵：

$$AsyVar[\hat{\gamma}_{2SLS}] = \hat{\sigma}^2[Z'Q(Q'Q)^{-1}Q'Z]^{-1} \tag{2.35}$$
$$\hat{\sigma}^2 = (y - Z\hat{\gamma}_{2SLS})'(y - Z\hat{\gamma}_{2SLS})/n$$

根据 SLM 模型的表达式，我们有 $Wy = W(I_n - \rho W)^{-1}X\beta + W(I_n - \rho W)^{-1}\varepsilon$，两边同时关于 X 取条件期望，可得：

$$E[Wy \mid X] = W(I_n - \rho W)^{-1}X\beta \tag{2.36}$$

再利用 $(I_n - \rho W)^{-1} = I_n + \rho W + \rho^2 W^2 + \rho^3 W^3 + \cdots$，进一步得到

$$E[Wy \mid X] = WX\beta + \rho W^2 X\beta + \rho^2 W^3 X\beta + \cdots \tag{2.37}$$

Kelejian and Robinson(1993)推导出 $\hat{\gamma}_{2SLS}$ 具有一致性，并基于式(2.37)建议工具变量从 $\{X, WX, W^2X, W^3X, \cdots\}$ 中选取。

2. 广义矩估计(generalized method of moments)

在 SEM 模型：$y = \alpha\iota_n + X\beta + u$，$u = \lambda Wu + \varepsilon$，$\varepsilon \sim N(0, \sigma^2 I_n)$ 中，Kelejian and Prucha (1999)建议用 GMM 方法来得到 λ 的一致性估计。可以找到以下三个矩条