

变系数空间计量模型的理论和应用

陈建宝 乔宁宁 著



科学出版社

变系数空间计量模型的 理论和应用

陈建宝 乔宁宁 著



科学出版社

北京

内 容 简 介

本书系统阐述了空间计量模型的重要性,针对复杂经济变量间普遍存在的非线性关系,提出了空间计量经济学理论发展的新挑战和新要求。在介绍相关预备知识之后,构建了半参数变系数空间滞后模型、半参数变系数空间误差回归模型和混合地理加权空间滞后回归模型(系数随地理位置变化)的估计方法,对估计量的大样本性质和小样本表现分别进行了数理论证和蒙特卡罗模拟研究,并将估计技术运用于现实经济问题分析中。这些模型估计量具有稳健性,可同时考察因变量的空间溢出效应和变量间的非线性特征,且有效地避免了非参数模型中的“维数灾难”问题。

本书的特色是理论和应用相结合,所述研究方法对于其他结构的半参数/非参数空间计量模型估计理论研究具有推广价值,其估计技术在经济、管理等学科中具有应用价值。本书可供相关领域的研究者和技术人员参考。

图书在版编目(CIP)数据

变系数空间计量模型的理论和应用/陈建宝,乔宁宁著. —北京:科学出版社,2018.3

ISBN 978-7-03-055984-5

I. ①变… II. ①陈… ②乔… III. ①空间模型 IV. ①O221.2

中国版本图书馆 CIP 数据核字 (2017) 第 310678 号

责任编辑:胡庆家/责任校对:彭珍珍
责任印制:张 伟/封面设计:蓝正设计

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>

北京建宏印刷有限公司印刷

科学出版社发行 各地新华书店经销

*

2018年3月第一版 开本:720×1000 B5

2018年3月第一次印刷 印张:13 1/2

字数:269 000

定价:98.00 元

(如有印装质量问题,我社负责调换)

作者简介

陈建宝，男，1965年生，澳大利亚科廷理工大学统计学博士。现任福建师范大学二级教授，博士生导师，统计学博士点和博士后流动站负责人，厦门大学宏观经济研究中心研究员，中国统计教育学会副会长，中国统计学会常务理事，《统计研究》编委。曾任厦门大学经济学院副院长、中国统计教材编审委员会委员。曾为美国芝加哥大学高级访问学者、澳大利亚西澳大学访问教授等。在国内外重要学术刊物上发表论文 140 多篇，出版专著 4 部；主持、参与完成国家和省部级重大、重点和面上项目 20 多项；获得国家级、省部级科研和教学奖励多项。主要研究方向为统计理论和方法、宏微观经济计量分析。

乔宁宁，男，1986年生，厦门大学经济学博士。现为中国人民银行太原中心支行国际收支处主任科员。在《经济学（季刊）》《数量经济技术经济研究》《统计研究》《旅游学刊》等学术刊物发表论文多篇。研究方向为空间计量模型和金融数据分析。

前 言

人类经济活动总是在一定的时间和空间维度上进行的,经济现象不仅表现在时间上的相关,其在空间上也存在某种程度的相关。例如,经济发达地区总是连成一片,相关或相同产业倾向于在同一地理空间聚集,而地理位置接近、个体间互相竞争和合作、模仿行为、溢出效应等,都成为产生这些空间交互效应的重要影响因素。

在实际经济研究过程中,经典的高斯-马尔可夫假设要求解释变量之间相互独立。如果所得数据间存在空间相关性,则彼此不能保持独立。同样,如果数据中存在空间异质性,也会违背高斯-马尔可夫假设中的误差项同方差假定。因此,为了处理这种空间相关性和空间异质性,空间计量经济学应运而生。自20世纪70年代末兴起以来,空间计量经济学获得了巨大发展,开始从边缘学科进入计量经济学的主流,不仅在区域经济问题、房地产经济学及经济地理学等传统经济学科中成为标准分析工具,而且在国际经济学、劳动经济学、资源和环境问题、城市化与人口发展等方面也得到了越来越广泛的运用。

然而,伴随着空间计量经济学的快速发展,越来越多的学者认识到,在现实经济运行中,变量之间的非线性关系比比皆是,学者们开始更多地考虑除线性以外模型的空间效应问题。与此同时,非线性特征的引入逐步使得传统参数空间计量模型的形式更为复杂,与其相关的各种估计理论也与线性模型存在较大差异,大多数情形下,很难将传统计量经济理论的研究成果直接推广到具有非线性特征的空间计量模型中,这使得空间计量模型在很多方面还处于研究的起步阶段,仍然面临诸多难点,而这也成为本书引入半参数/非参数空间计量模型进行研究和探索的主要出发点。事实上,相比传统的参数空间计量模型,非参数空间计量模型在处理非线性特征上具备较大的灵活性和稳健性。然而,这类模型对于高维数据却存在着不可避免的“维数灾难”问题,这在一定程度上限制了模型的应用空间,因此如何有效规避“维数灾难”问题,也成为本书在构建新模型方面需要考虑和解决的关键问题。基于这些考量,本书提出了两类新的变系数空间计量模型——半参数变系数空间滞后回归模型和半参数变系数空间误差回归模型,不仅能够有效刻画变量之间存在的非线性影响关系,在降维方面也具有较好的功效。

另外,在现实经济结构或社会结构中,空间相关性和空间异质性成为影响空间效应的两个重要层面。大多数研究主要针对其中某一方面展开,尤以涉及空间相关性的文献居多。然而,当空间相关性与空间异质性同时存在时,经典的计量经济学

估计方法可能不再简单适用,这时一个直接的困难便是如何处理单位间可能同时存在的空间相关性及空间异质性,两种空间效应的叠加无疑增加了模型估计的难度。为此,本书还提出了另外一类新的混合地理加权空间滞后回归模型。在某种意义上,这类模型综合了空间计量模型和混合地理加权回归模型的特点,不仅兼顾到空间观测点的位置对异质性产生的影响,反映了空间异质性的特征,而且还考虑了更为贴近实际的情形,即一部分系数随着空间位置变化,而其余部分为常系数,因而模型的适用性变得更为广泛。

非参数空间计量模型作为空间计量经济模型的重要分支,具有广阔的应用前景,而系统介绍这些内容的书籍依然较少,本书抛砖引玉,旨在对非参数空间计量模型估计理论方面的研究起到借鉴作用,同时对经济、管理等学科中的相关实证研究产生应用价值。本书内容主要包括对半参数变系数空间滞后回归模型、半参数变系数空间误差回归模型、混合地理加权空间滞后回归模型的研究,以及运用所构建的理论模型研究资源禀赋与地方公共品供给之间的关系等。具体来看,第1章为绪论部分;第2章为预备知识,综述空间计量模型和变系数回归模型的研究进展,并介绍一些本书涉及的已有估计方法和计算方法;第3章讨论半参数变系数空间滞后回归模型的估计及统计检验;第4章讨论半参数变系数空间误差回归模型的估计;第5章讨论混合地理加权空间滞后回归模型的估计;第6章为实证应用部分,主要运用本书构建的变系数空间计量模型研究我国资源禀赋、地方政府博弈与公共品供给之间的关系。

本书得到了国家社会科学基金规划项目(16BTJ018)、国家自然科学基金项目(71503220)、教育部人文社会科学重点研究基地重大项目(15JJD790029)、教育部人文社会科学研究规划基金项目(13YJA9100002)、福建省自然科学基金项目(2017J01396)和福建师范大学创新团队基金项目“概率与统计:理论和应用”(IRTL1704)的资助。对各级部门的大力支持,我们在此表示衷心的感谢!

由于时间、资料、知识、精力等多方面的局限性,书中疏漏甚至错误在所难免。恳请对该领域有兴趣的读者、专家不吝赐教。

作 者

2017年10月20日

目 录

前言

第 1 章	绪论	1
1.1	空间计量模型的重要性	1
1.2	非线性特征的广泛存在性	2
1.3	空间异质性的普遍性	3
1.4	空间计量模型面临的挑战	4
第 2 章	预备知识	5
2.1	空间计量模型	5
2.1.1	空间权重矩阵	5
2.1.2	截面数据空间计量模型	6
2.1.3	面板数据空间计量模型	7
2.1.4	非参数空间计量模型	9
2.2	变系数回归模型	10
2.2.1	常见的变系数回归模型	10
2.2.2	半参数变系数回归模型	14
2.2.3	地理加权回归模型	15
2.2.4	混合地理加权回归模型	18
2.3	主要涉及的估计方法和计算方法介绍	22
2.3.1	局部线性估计和局部多项式估计	22
2.3.2	截面似然估计法	26
2.3.3	常用的窗宽选择方法	27
2.3.4	三阶矩 χ^2 逼近法	30
2.3.5	Bootstrap 方法	31
2.4	本章小结	31
第 3 章	半参数变系数空间滞后回归模型的估计	33
3.1	引言	33
3.2	半参数变系数回归模型的空间相关性检验	34
3.2.1	半参数变系数回归模型	34
3.2.2	空间相关性检验	36
3.2.3	蒙特卡罗模拟结果	39

3.3	半参数变系数空间滞后回归模型的估计	44
3.3.1	模型设定	44
3.3.2	模型估计	46
3.3.3	估计的大样本性质	48
3.3.4	蒙特卡罗模拟结果	51
3.3.5	回归模型的比较分析	60
3.4	一类半参数空间滞后回归模型的估计	63
3.4.1	模型设定	63
3.4.2	蒙特卡罗模拟结果	64
3.5	变系数函数的稳定性检验	66
3.5.1	广义似然比统计量	66
3.5.2	蒙特卡罗模拟结果	70
3.6	引理和定理证明	71
3.7	本章小结	89
第 4 章	半参数变系数空间误差回归模型的估计	90
4.1	引言	90
4.2	半参数变系数空间误差回归模型的估计	91
4.2.1	模型设定	91
4.2.2	模型估计	92
4.2.3	估计的大样本性质	94
4.2.4	蒙特卡罗模拟结果	97
4.2.5	回归模型的比较分析	107
4.3	一类半参数空间误差回归模型的估计	109
4.3.1	模型设定	110
4.3.2	蒙特卡罗模拟结果	110
4.4	考虑空间相关结构时模型非参数部分的估计	113
4.4.1	模型估计	114
4.4.2	参数估计的大样本性质	116
4.4.3	非参数估计的大样本性质	117
4.5	引理和定理证明	117
4.6	本章小结	142
第 5 章	混合地理加权空间滞后回归模型的估计	144
5.1	引言	144
5.2	混合地理加权回归模型的空间相关性检验	145
5.2.1	混合地理加权回归模型	145

5.2.2 空间相关性检验	146
5.3 混合地理加权空间滞后回归模型的参数估计	149
5.3.1 模型设定	149
5.3.2 模型参数估计	150
5.4 蒙特卡罗模拟结果	152
5.4.1 空间相关性检验的模拟过程	152
5.4.2 混合地理加权空间滞后回归模型参数估计的模拟过程	155
5.5 本章小结	159
第 6 章 资源禀赋、地方政府博弈与公共品供给	160
6.1 引言	160
6.2 理论分析框架	162
6.2.1 基本假定与博弈过程	162
6.2.2 地方政府行为	165
6.3 实证框架	170
6.3.1 模型设定	171
6.3.2 模型估计	172
6.4 实证结果及分析	172
6.4.1 空间相关性检验	173
6.4.2 线性空间回归模型的估计结果	173
6.4.3 半参数变系数空间滞后回归模型的估计结果	176
6.4.4 半参数变系数空间误差回归模型的估计结果	180
6.5 本章小结	183
参考文献	185

第1章 绪 论

1.1 空间计量模型的重要性

空间计量经济学起源于区域科学和计量经济学共同发展,是计量经济学的一个重要分支。1974年 Paelinck 首次提出空间计量经济学的概念,即此拉开了空间计量经济学研究的序幕,随后在 Cliff, Ord, Anselin 等众多学者的努力下,空间计量经济学的框架体系逐步成型。Anselin (1988a) 对空间计量经济学进行了系统研究,针对空间计量经济学的概念、模型、估计、检验等方面提供了详尽的总结和扩展,随后在 2001 年又将空间计量经济学补充定义为:“在横截面和面板数据回归模型中处理空间交互作用(空间自相关)与空间结构(空间异质性)的计量经济学的分支”。Anselin 的工作为空间计量经济学的发展奠定了重要基础。

自 20 世纪 70 年代末以来,对空间计量经济学的理论探索得到了广泛的关注,研究领域几乎涉及经典计量经济学的绝大部分设定、估计以及检验方法。然而,空间维度作为与时间维度同等重要的一个经济联系维度,其本身仍具有同时间维度不同的特性。空间计量模型通过引入空间权重矩阵以及空间滞后因子,将研究对象的空间相关性引入模型之中,正是这一点拓宽了模型分析的广度,同时也增加了计量分析的复杂性。而空间滞后项的引入能够较好地挖掘潜在的空间相关性特征,减少原有模型不能被识别的信息量,这使得空间计量模型成为计量经济学界方兴未艾的热点议题。最初,大多数空间计量经济学的研究集中于区域经济问题、房地产问题、环境问题、城市化与人口发展问题等方面,得到了较为满意的研究成果。近些年来,伴随地理信息系统和空间数据分析软件,以及数据处理、信息处理等技术的发展,空间计量经济学的应用变得更为广泛。

事实上,社会经济中有大量研究数据都是按照一定的空间形式组织起来的,高度聚类是这些单位的典型特征,尤其是空间层面的地区聚类。在实际操作方面,这些聚类通常都被忽略或者当作一种干扰,而忽略这种关联性将极大地影响我们在研究中建立有意义的推论。空间计量经济学作为空间分析的重要组成部分,它的发展为减少这种代价提供了一种分析方法,空间信息的刻画也将有助于揭示社会过程之间的联系,这将极大地促进空间分析在众多领域中的应用,为人们研究社会经济问题开辟新的视野。可以说,无论从理论发展,还是从实际应用的角度,对空间计量经济学的研究都具有重要意义。

1.2 非线性特征的广泛存在性

大量的实际经济数据表明,经济变量之间大多呈现非线性影响关系,这些非线性关系不仅具有合理的经济理论基础,而且在宏观经济调控、微观市场结构分析等方面发挥着重要作用。例如,改革开放以后,我国不同时期的经济体制政策发生了很大变化,经济发展状况存有较大差异,这就导致了各个时期经济变量之间可能存在显著的结构变化,如果忽视这种变化,而直接采用线性模型分析变量之间的经济关系,那么,产生的后果很可能是经济意义被扭曲甚至是错误的。线性模型存在的固有的一些不足之处,使得传统线性时间序列计量经济学受到越来越多来自非线性时间序列计量经济学的挑战,继而涌现出大量学者投身于非线性模型的讨论热潮,并提出了多种非线性模型用来捕捉经济变量之间的各类非线性特征。与线性模型形式单一不同,非线性模型的形式呈现多样化,这里简要列举几类常见的非线性特征的表现形式。

其一,非线性特征的一种主要形式为非对称性。非对称性在宏观经济分析、金融资产收益率等方面均有体现。在宏观经济中,不仅经济周期本身具有明显的非对称性(刘金全,范剑青,2001),而且经济变量之间的影响关系也较为显著,如菲利普斯曲线(欧阳志刚,王世杰,2009;陈建宝,乔宁宁,2013)、货币政策反应函数(王立勇等,2010)等;在金融领域,股票市场和汇率市场同样存在着显著的非对称性(Hong, Lee, 2003;陈浪南,孙坚强,2010),如果建模时考虑到这些特点将能够更准确地反映和预测资产收益率的发展趋势。

其二,非线性特征的另一种典型形式为波动集聚性。在金融市场上,资产收益率的波动具有明显的集聚性特征(Engle, 1982; Bollerslev, 1986),而部分学者同时注意到了波动率中的杠杆效应,即负向冲击相对于正向冲击会引起更大的波动,继而非线性 ARCH(Auto-regressive Conditional Heteroskedasticity)类和 GARCH(Generalized Auto-regressive Conditional Heteroskedasticity)类等模型应运而生(赵振全等,2005; Christensen et al., 2012),并较好地刻画了变量中存在的非线性特征。

当然,除了非对称性和波动集聚性外,经济变量之间的非线性特征还存在多种表现形式(杨子晖,2010;段景辉,陈建宝,2011;章上峰等,2011),此处不再一一列出。

综上所述,在经济和金融分析中,非线性特征不仅具有普遍的存在性,而且在实证分析中得到了广泛验证。因此,如何进一步扩展和完善非线性建模分析,更好地将其应用于实际经济问题研究,具有重要的现实意义。

为了解决实际估计中遇到的非线性特征问题,计量学者考虑在传统参数计量

模型研究中引入非参数思想,以更好地解释变量内部结构未知情况。自20世纪40年代起,非参数技术及其相关研究逐步成为统计学和计量经济学中最为活跃的科研领域之一。最初,Bartlett在时间序列谱密度估计研究中,借鉴平滑技术,构建了非参数密度估计的基本思想。基于非参数密度估计思想,Nadaraya和Watson在20世纪60年代又提出了核回归思想。继而,Hastie和Tibshirani(1990),Green和Silverman(1994),Wand和Jones(1995),Fan和Gijbels(1996)分别从不同角度介绍了各种非参数模型的理论和应用。其间Stone(1977)关于非参加权函数大样本理论的发表为非参数模型统计推断方法的研究奠定了重要基础。后来,随着非参数模型理论研究的不断深入,计量学者陆续提出了局部多项式拟合、样条平滑等非参数技术思想,并逐步应用于时间序列数据模型、广义线性模型、面板数据模型及金融利率等实际模型之中。

1.3 空间异质性的普遍性

空间计量经济学作为计量经济学的一个分支,解决的是计量经济学中的空间效应问题,空间效应包括空间交互作用和空间结构两个方面(Anselin, 2001a)。其中空间交互作用更多地体现在空间依赖性上,而空间结构的变化则主要体现在空间异质性上。空间依赖性和空间异质性就是截面数据的依赖性和异质性在空间上的体现。具有依赖性的空间单位可能在一定程度上受到其他单位位置和距离的影响,这种位置和距离既包括地理空间意义上的,也包括广义的经济或社会网络空间意义上的;而空间异质性则主要指地理空间上的区域缺乏均质性,反映了经济实践中的空间单位之间经济行为关系的一种普遍存在的不稳定性。

对于空间异质性,一方面,在空间计量模型中,一些学者将其理解为模型设定的差异,其中一个重要表现形式为异方差性,这些可以采用经典计量经济学的基本方法进行处理。例如,异方差的White检验,还有通过面板数据模型的方差、协方差矩阵来处理空间异质性(Anselin, 1988a, 2001a)等。另一方面,针对空间异质性的特征,Brunsdon等(1996)和Fotheringham等(1997)提出了一种称为地理加权回归(Geographically Weighted Regression, GWR)的空间变系数模型,该模型的系数是观测个体的空间位置的函数,因而不同观测个体的系数是不同的。显然,空间异质性在第二种情况下更为广义,体现了局部异质的特点,而本书对空间异质性的理解也是基于第二种情况,它体现了因地理位置变化而产生的空间非平稳性。

为了更好地捕获空间数据中可能存在的异质性,地理加权回归模型被广泛应用于各种社会经济问题的研究中,其中不乏经济收敛性(LeSage, 1999; Eckey et al., 2007)、经济发展(苏方林, 2007; Partridge et al., 2008)、房价分析(Huang et al., 2010; 张琰, 梅长林, 2012)以及社会和人口发展(Wheeler, Waller, 2009; Shoff,

Yang, 2012) 等方面。相比通常的线性回归模型, 地理加权回归模型能够分析回归关系随着空间位置变化而变化的特征, 不言而喻, 这种模型的恰当运用和对经济问题的深刻反映, 为空间异质性的客观存在提供了重要支撑, 也为其广泛应用提供了强有力的保障。

1.4 空间计量模型面临的挑战

相比传统的线性空间计量模型, 非线性特征的引入使得空间计量模型的形式更为复杂, 与其相关的各种估计理论也与线性模型存在较大差异, 大多数情形下, 很难将传统计量经济理论的研究成果直接推广到具有非线性特征的空间计量模型中, 这使得空间计量模型在很多方面还处于研究的起步阶段 (Su, Jin, 2010; Su, 2012; 李坤明, 陈建宝, 2013), 仍然面临诸多难点。特别地, 联系到本书所要研究的变系数空间计量模型, 尤其是半参数变系数空间滞后回归模型和半参数变系数空间误差回归模型, 同样面临很多有待解决的瓶颈问题。然而, 对这类模型的探索又十分必要, 因为在实际问题分析中, 如果我们仅仅采用线性框架下的空间计量模型考察经济变量之间的关系, 极可能产生由模型形式误设而导致的估计偏差问题, 甚至得出错误的分析结论和误导性的政策建议。

另外, 在现实的经济结构或社会结构中, 空间相关性和空间异质性是影响空间效应的两个重要层面。大多数研究主要针对其中某一方面进行展开, 尤以涉及空间相关性的文献居多。然而, 当空间相关性与空间异质性同时存在时, 经典的计量经济学估计方法可能不再简单适用, 问题或将更为复杂。这时, 研究人员面临的一个直接困难便是如何处理截面单位可能同时存在的空间相关性及空间异质性, 两种空间效应的叠加无疑增加了计量模型估计的工作难度。

第2章 预备知识

2.1 空间计量模型

空间计量经济学主要是研究在计量经济模型中处理空间效应的一系列方法。由于空间变量的诸多特殊性质,很多情况下研究空间变量之间的关系需要在回归模型的基础上体现出变量的空间结构特征,这样构建的模型称为空间计量模型。回溯空间计量模型的整个发展历程,其间,很多学者对此理论做出了重要贡献,可谓百花齐放。这里将以空间数据的类型为出发点,针对学者们研究和应用较为广泛的截面数据空间计量模型和面板数据空间计量模型进行简要回顾,同时,给出非参数思想在空间计量模型方面的一些探索性研究。

2.1.1 空间权重矩阵

空间权重矩阵是空间相关性的定量化测度,是描述数据空间结构的基础。假设研究区域有 N 个空间单元,任何两个都存在一个空间关系,这样就有 $N \times N$ 对关系。于是需要 $N \times N$ 的矩阵存储这 N 个空间单元之间的关系。关于空间权重矩阵的构造方式很多,最简单的方法是基于空间单元的二元邻接关系得到相应的邻接矩阵,如果观测值 i 和 j 所在的空间单元在地理上相邻,即存在共同的边界,则 $w_{ij} = 1$, 否则 $w_{ij} = 0$; 主对角线上元素 w_{ii} 为零,即空间单元和其自身不存在空间自相关。

空间权重矩阵主要包括基于连通矩阵和基于距离两种计算方法。其一,基于连通矩阵的二元空间权重矩阵,一般包括 Rook 邻近计算方法和 Queen 邻近计算方法, Rook 邻近计算方法以共同边界来确定“邻居”,而 Queen 邻近计算方法除了共有边界邻居外,还包括具有共同顶点的邻居;其二,基于距离的二元空间权重矩阵则以权重矩阵 W 代表不同空间单元之间的距离,即 $w_{ij} = w_{ij}(d)$ 。其中, d 是两个空间单元的距离,在实际应用中,距离的定义常常不局限于欧氏距离,在有些情况还采用经济距离、社会距离、制度距离等。根据地理学第一定律,两个个体的空间关系紧密程度与其二者之间的距离成反比,故使用距离作为权重,用以描述空间个体之间关系的远近。因此,当使用距离矩阵时,权重是距离的倒数,但是根据相关研究,很多空间关系的强度随着距离的减弱程度,通常要强于简单的线性比例关系,因此在实际研究中,经常采用平方距离的倒数作为权重。另外,空间矩阵中的空间因子为处理上的方便通常会进行标准化,在估计中一般采用标准化后的

矩阵。

空间自相关性的度量方法通常分为全局空间自相关方法和局部空间自相关方法。其中,全局空间自相关方法基于全部区域,描述了所有区域单元的总体空间关系,主要测量指标包括 Moran's I 统计量、Geary C 统计量以及 G 统计量等;而关于局部空间相关性检验,其作用主要在于对空间相关性引起的空间差异性进行诊断,判断空间个体的属性取值相对应的热点区域或高发区域,从而对全局空间相关性分析方法进行补充。局部空间自相关测量指标主要包括局部 Moran's I 统计量、局部 Geary C 统计量、局部 G 统计量和 Moran 散点图等。

2.1.2 截面数据空间计量模型

在现实经济或社会结构中,由于空间依赖性广泛存在,空间相关性一直是空间计量经济学研究的重点领域。最早的研究起步于截面数据空间计量模型,包括空间滞后回归模型、空间误差回归模型和混合空间滞后回归模型等,研究重点主要集中于模型估计和统计检验两大方面。

截面数据空间计量模型的一般形式可写成

$$\begin{aligned} y_i &= \rho(W_1 Y)_i + u_i \\ u_i &= \lambda(W_2 u)_i + \varepsilon_i \end{aligned} \quad (2.1)$$

其中, $Y = (y_1, y_2, \dots, y_N)^T$ 是 N 维被解释变量观测值, x_i 表示 p 维解释变量观测值; $u = (u_1, u_2, \dots, u_N)^T$ 是随空间变化的误差项; ε_i 是白噪声 $\varepsilon = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_N)^T$ 中的第 i 个元素且满足 $\varepsilon_i \sim \text{i.i.d.}(0, \sigma^2)$; $\rho(W_1 Y)_i$, $\lambda(W_2 u)_i$ 为滞后项,其中,参数 ρ 反映因变量的空间邻接单元对于因变量的解释程度, λ 反映误差项的空间邻接项对于其本身的解释程度,为了保证模型的平稳性,一般需要满足 $|\rho| < 1$, $|\lambda| < 1$, W_1 和 W_2 是已知的空间权重矩阵。 β 为 p 维回归系数,其反映解释变量对于因变量变化的影响,对式 (2.1) 增加某些限定,可导出多种不同形式的计量模型。若 $W_2 = 0$, 则得到空间滞后模型;若 $W_1 = 0$, 则推出空间误差模型;若 $W_1 = 0$, $W_2 = 0$, 则模型变为普通的线性回归模型。

在模型估计上,Ord (1975) 首次提出了空间滞后回归模型的极大似然估计 (Maximum Likelihood Estimator, MLE) 方法,后有 Cliff 和 Ord (1981), Anselin (1988a), Smirnov 和 Anselin (2001) 等学者进行了相应的拓展研究。由于极大似然估计需要设定误差项服从正态的独立同分布,这在现实中往往很难符合, Lee (2004) 进一步放宽了模型误差项服从正态分布的假设,提出了空间滞后回归模型的拟极大似然 (Quasi-maximum Likelihood Estimation, QMLE) 估计,这为空间滞后回归模型在应用上的拓宽提供了有力的理论支持;后来,工具变量估计 (Haining, 1978; Bivand, 1984; Anselin, 1984; Drukker et al., 2010)、广义矩估计 (Kelejian, Robinson, 1993; Kelejian, Prucha, 1998, 1999, 2010; Lee, 2003, 2007a, 2007b;

Lee, Liu, 2010) 和贝叶斯估计 (Besag et al., 1991, 1995; Hepple, 1995; LeSage, 1997) 等方法逐步应用到截面数据空间计量模型的估计中, 成果显著。

在统计检验上, 空间自相关检验一直是空间计量经济学的一个重要研究方向, 是判断空间依赖性是否存在的关键依据。其中 Moran's I 统计量最早得到开发和运用 (Cliff, Ord, 1972, 1973, 1981), Anselin 和 Rey (1990) 进一步分析了在不同空间衔接结构下统计量的检验效果; Anselin 和 Kelejian (1997) 采用蒙特卡罗模拟方法, 研究了 Moran's I 检验在包含内生变量与采用两阶段最小二乘法 (Two Stage Least Square Method, 2SLS) 方法估计的回归模型中的有限样本表现。可以看出, Moran's I 检验是基于普通最小二乘法 (Ordinary Least Square, OLS) 和两阶段最小二乘法估计残差的空间相关性检验方法, 它的优点主要在于其结构简单且具有良好的有限样本表现。然而, Moran's I 统计量虽然能够检测模型中存在的空间相关性, 但并不能有效识别具体的空间相关性结构。基于此, Anselin 等 (2004) 提出了 LMERR, LMLAG 和稳健 (Robust) 的 ROLMERR, ROLMLAG 等检验统计量的判别准则, 用以推断空间滞后回归模型和空间误差回归模型是否存在空间相关性, 以决定哪种空间模型更加符合客观实际。如果在空间依赖性的检验中发现, LMLAG 比 LMERR 在统计上更加显著, 且 ROLMLAG 显著而 ROLMERR 不显著, 则可以断定适合的模型是空间滞后回归模型; 相反, 如果 LMERR 比 LMLAG 在统计上更加显著, 且 ROLMERR 显著而 ROLMLAG 不显著, 则可以断定空间误差回归模型是恰当的模型 (吴玉鸣, 2006)。除了上述检验方法外, Anselin (2001b) 提出了空间相关性的得分检验, 原假设模型设为一般的线性回归模型, 备择假设则为空间 ARMA(p, q) 模型或空间误差组合模型; Kelejian 和 Robinson (1998) 考虑了存在异方差情况下的空间滞后回归模型, 并讨论了该模型的空间自相关检验方法; Kelejian 和 Prucha (2001) 证明了: 当模型误差项服从独立同分布时, 空间滞后回归模型的空间相关性检验统计量 Moran's I 指标渐近服从正态分布、 χ^2 分布等标准分布。

2.1.3 面板数据空间计量模型

正如将普通的截面数据模型扩展到面板数据模型, 对于空间计量模型而言, 研究者同样很自然地将截面数据空间计量模型延伸到面板数据空间计量模型, 这样既能有效处理以往普通面板数据模型可能忽略的空间相关性, 又能解决截面数据空间计量模型可能存在的问题。

对于空间单元为 N 、时期为 T 的样本观测数据, 面板数据空间计量模型的一般形式如下:

$$\begin{aligned} y_{it} &= \rho(W_1 Y_t)_i + x_{it}^T \beta + b_i + \mu_t + u_{it} \\ u_{it} &= \lambda(W_2 u_t)_i + \varepsilon_{it} \end{aligned} \quad (2.2)$$

其中, $1 \leq i \leq N$, $1 \leq t \leq T$; $Y_t = (y_{1t}, y_{2t}, \dots, y_{Nt})^T$ 为 N 维被解释变量在第 t 期的观测值; $x_{it} = (x_{1it}, x_{2it}, \dots, x_{pit})^T$ 为 p 维解释变量中第 i 个个体在第 t 期的观测值; $\beta = (\beta_1, \beta_2, \dots, \beta_p)^T$ 为 p 维回归系数, ε_{it} 是白噪声 $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t}, \dots, \varepsilon_{Nt})^T$ 中第 i 个个体在第 t 期的取值, 并且 ε_t 满足 $E[\varepsilon_t] = 0$, $E[\varepsilon_t \varepsilon_t^T] = \sigma^2 I_N$, u_{it} 是误差项 $u_t = [u_{1t}, u_{2t}, \dots, u_{Nt}]^T$ 中第 i 个个体在第 t 期的取值; 模型中 b_i 和 μ_t 分别描述了样本的个体特性以及时间特性, 而空间权重矩阵 W_1 和 W_2 分别反映了个体变量和误差项的空间相关性, 并且满足对角线元素为 0, 非对角线上的元素反映了个体变量或者误差项对于临近区域的影响。因此, 若样本的个体特性或时间特性为固定的, 则该模型为固定效应空间计量模型, 否则为随机效应空间计量模型。类似地, 也可以得到面板数据空间滞后模型和面板数据误差滞后模型的具体形式, 这里不再重复说明。

Elhorst (2003) 提出了存在空间滞后相关和空间误差相关两种空间相关性结构的情况下, 涉及的几类面板数据空间计量模型, 主要包括固定效应模型、随机效应模型、固定系数模型以及随机系数模型。类似普通的面板数据模型, 研究者更多侧重于对含有固定效应和随机效应的面板数据空间计量模型的研究。其中, 对于空间滞后自回归固定效应模型, Anselin (2001a) 指出时间固定而截面个数趋于无穷大时, 只能够得到解释变量对应系数的一致估计量, 而无法得到截距项和空间自回归系数的一致估计量; Lee (2004, 2007a) 的研究表明当时间趋于无穷大时, 利用极大似然估计仍然能够得到截距项和空间自回归系数的一致估计量; 然而, 针对极大似然估计可能存在的问题, 部分学者指出采用矩估计方法或许是个更好的选择 (Kelejian, Prucha, 1999, 2002, 2004; Lee, 2003, 2007a, 2007b); 对于空间误差自回归固定效应模型, Elhorst (2003) 提出了消除固定效应的极大似然估计; Druska 和 Hoxby (2004) 构造了该模型的广义矩 (Generalized Method of Moments, GMM) 估计和可行的广义最小二乘 (Feasible Generalized Least Square, FGLS) 估计; Moscone 和 Tosetti (2011) 考虑了模型含有异方差的情况, 并提出了 GMM 估计方法; 对于随机效应模型, Elhorst (2003), Baltagi 和 Li (2004) 给出了空间误差自回归随机效应模型的设定形式及极大似然估计方法; 对于随机系数和固定系数模型, 诚如 Elhorst (2003) 所言, 当空间异质性不能被截距项完全捕捉时, 放松斜率项固定不变的假设, 引入斜率项在时间维度或空间维度上的变动, 就显得很有必要。基于这样的研究背景, Elhorst 对此类变系数的面板数据空间计量模型进行了系统阐述; 此外, Fingleton (2008) 对综合了空间滞后相关和空间误差相关的混合空间面板固定效应模型进行了 GMM 估计, Lee 和 Yu (2010) 采用拟极大似然估计方法同样对该模型进行了研究; Mutl 和 Pfaffermayr (2011) 进一步提出了混合空间面板随机效应模型的 GMM 估计方法, 以及固定效应模型的组内估计法。

上述研究主要基于面板数据空间计量模型的静态研究, 其实还有部分学者对