

含空间自回归误差的空间自回归模型 的经验似然推断

专业名称: 统计学

申 请 人: 唐 洁

指导老师: 秦永松 教授

论文答辩委员会

主席:_____

委员:_____

含空间自回归误差的空间自回归模型的经验似然推断

年级: 2020 级 专业: 统计学 研究方向: 经验似然

研究生: 唐洁 导师: 秦永松教授

摘 要

数据以时空的形式出现在生活方方面面, 如某时某地的无症状感染人数、房价、物种数, 对时空数据的研究已经发展成空间计量经济学, 已广泛用于流行病学、区域经济学、生物科学等等诸多领域. 空间截面数据模型在空间计量经济学中占据核心地位, 大部分空间面板数据模型的估计方法由空间截面数据估计方法推广而来, 因此, 研究其统计推断有重要的实用价值和理论意义.

Owen 提出的经验似然方法具有参数方法的有效性又具有非参数方法的稳定性, 其构造的置信区间具有域保持性, 变换不变性, 并且置信域的形状由数据自行决定等优点, 这使得该方法备受统计学者关注, 并在各种统计模型中得到应用. 然而经验似然应用于空间模型最大的问题在于估计方程是关于误差的线性-二次型形式. 针对该问题, Qin 根据拟似然方法得到的估计方程的特性引入鞅差序列, 将误差的线性-二次型转化为鞅差序列的线性形式, 问题得以解决, 由此经验似然应用于空间模型的研究发展了起来. 由于经验似然方法自身理论存在局限性, 在实际应用中常常遇到其他问题, 例如, 数据过少时经验似然比函数可能出现无解的情况, 数据维数发散时传统经验似然失效等问题, 因此, 本文研究的空间截面数据模型改进的经验似然推断具有一定的理论价值, 拓展了空间数据模型的统计推断方法.

本文的主要研究工作可以概括为以下两点:

一、本文在已有的理论研究基础上, 研究 SARAR 模型的调整经验似然推断, 解决凸包可能不含 0 的问题, 证明调整经验似然比统计量的渐近分布为卡方分布, 与普通的经验似然比统计量的渐近分布有相同的自由度. 我们利用这一理论结果构造模型参数的置信区域并设计数值试验, 模拟表明, 当样本量较少时, 该方法下置信区域的覆盖率更接近置信水平, 计算速度更快, 不需要 Bartlett 校正和 Bootstrap 方法那么复杂的程序.

二、本文在已有的理论研究基础上, 研究 SARAR 模型的高维经验似然推断, 解决维数爆炸的问题, 证明高维经验似然比统计量的渐近分布是自由度为 1 的卡方分布, 比普通经验似然比统计量的渐近分布的自由度小得多. 我们利用这一理论结果构造模型参数的置信区域并设计数值试验, 模拟表明, 当维数很高时, 该方法下置信区域的覆盖率更接近置信水平, 检验功效更优, 计算速度非常快.

关键词: SARAR 模型; 调整经验似然; 高维数据; 覆盖率

Empirical Likelihood for Spatial Autoregressive Models with Spatial Autoregressive Disturbances

Grade: 2020 Major: Statistics Research field: Empirical Likelihood

Graduate: Tang Jie Supervisor: Qin Yongsong

ABSTRACT

Data appears in all aspects of life in the form of time and space, such as the number of asymptomatic infections, housing prices, and species in a certain place at a certain time. The research on spatial panel data has developed into spatial econometrics, and has been widely used in many fields such as epidemiology, regional economics, and biological science. Spatial cross-sectional data model has been at the core of spatial econometrics, and its estimation methods can be extended to the spatial panel data model. Therefore, it is of great practical value and theoretical significance to study its statistical inference.

Since Owen proposed the empirical likelihood method, it has attracted the attention of many scholars with its effectiveness of parametric method and the stability of nonparametric method. The confidence interval constructed by empirical likelihood method has the advantages of domain retention, transformation invariance, and the shape of the confidence domain is determined by the data itself, which makes the method widely applied in various statistical models. However, the estimation equation obtained by the quasi likelihood method for the spatial cross-sectional data model is a linear quadratic form of error, which poses a great challenge to the application of empirical likelihood method to spatial models. To address this problem, Qin introduce a martingale sequence to transform the linear-quadratic form into a linear form to solve the problem, according to the characteristics of the estimation equation obtained by the quasi likelihood method, and this approach of transformation also opens a way to use empirical likelihood methods to other spatial models. However, the empirical likelihood method has many limitations in its own theory, and it often have other trouble in several practical application, such as the fact that the empirical likelihood ratio function may have no solution when the data size is small, and the traditional empirical likelihood fails when the data dimension is diverging. Therefore, innovative empirical likelihood methods of the spatial section data model studied in this paper has certain theoretical value and expands the statistical inference method of the spatial data model.

The main research work of this paper can be summarized as follows:

1、Based on the existing theoretical research, this paper studies the adjustment empirical likelihood inference for a SARAR model, solves the problem that convex hull may not contain zero, and proves that the asymptotic distribution of adjusted empirical likelihood ratio statistics is chi square

distribution, which has the same degree of freedom as the asymptotic distribution of usual empirical likelihood ratio statistics. We use this theoretical result to construct the confidence region of the model parameters and design numerical experiments. Simulation shows that when the sample size is small, the coverage of the confidence region under this method is closer to the confidence level, and the calculation speed is faster, without the complex programs such as Bartlett correction and Bootstrap method.

2、On the basis of existing theoretical research, this paper studies high-dimensional empirical likelihood inference for a SARAR model, solves the problem of dimension explosion, and proves that the asymptotic distribution of high-dimensional empirical likelihood ratio statistics is chi square distribution with degree of freedom of one, which is much less than the degree of freedom under usual empirical likelihood method. We use this theoretical result to construct the confidence region of the model parameters and design a numerical experiment. Our simulation shows that when the dimension is divergent, the coverage of the confidence region under this method is closer to the confidence level, the test power is better, and the calculation speed is very fast.

Key words: Spatial ARAR model; Adjusted Empirical Likelihood; High dimensional data; Confidence region

目 录

摘要	I
ABSTRACT	II
第一章 空间计量经济模型的经验似然研究进展	1
§1.1 空间权重矩阵	1
§1.2 空间数据模型的定义	3
§1.3 空间数据模型的研究概况	5
§1.4 经验似然方法的研究概况	6
§1.5 含空间自回归误差的空间自回归模型的普通经验似然推断	8
§1.6 本文的主要内容和创新点	12
第二章 含空间自回归误差的空间自回归模型的调整经验似然推断	14
§2.1 调整经验似然原理	14
§2.2 调整经验似然比统计量的极限分布	15
§2.3 模拟结果	16
§2.4 引理及定理的证明	21
第三章 含空间自回归误差的空间自回归模型的高维经验似然推断	25
§3.1 高维经验似然原理	25
§3.2 高维经验似然比统计量的极限分布	26
§3.3 模拟结果	26
§3.4 引理及定理的证明	32
结论及展望	39
参考文献	40
致 谢	45
声 明	46

第一章 空间计量经济模型的经验似然研究进展

众所周知, 各省之间的经济不是完全独立的, 经济数据涉及到一定的空间位置关系, 正如地理学第一定律^[1]所述, 所有事物都与其他事物相关联, 但较近的事物比较远的事物更关联. Cliff 和 Ord^[2]提出一种定量分析空间影响的模型, 近些年得到发展, 衍生出空间计量经济学学科, 广泛应用于区域经济、环境科学、犯罪学、经济地理、流行病学等多个研究领域. 总之, 空间数据几乎在社会各个领域出现, 并由广泛的应用前景, 因此其统计推断的研究显得尤其重要. 经验似然是 Owen^[3]提出的一种非参数统计推断方法, 许多研究表明此方法有显著的优点. 然而, 经验似然方法本身有许多限制, 比如似然比统计量存在的前提条件是 0 必须在估计函数集合的凸包内部, 高维情形下普通经验似然方法失效等, 如何让 0 一定在估计函数集合的凸包内部, 高维情形下经验似然方法仍然有效, 这些都成为经验似然领域新的研究问题.

本章将分别介绍空间权重矩阵的常见给定方式, 空间数据模型的定义及其研究概况, 经验似然方法的研究概况及其在空间截面数据模型的应用, 以及本文研究的内容和创新点.

§1.1 空间权重矩阵

许多经济数据都涉及一定的空间位置, 比如, 研究全国各省的国内生产总值、投资、贸易、研发等数据. 通常假设空间各省的变量相互独立, 这是不合理的, 各省经济有着广泛的联系, 而且距离越近的省份联系越紧密. 事实上, 各省之间的距离信息并不难获得, 比如是否相邻, 直线距离或者运输距离. 这种地理距离还可以推广至人际交往距离、网络距离、经济距离等等. 在一个空间系统里, 度量空间单元的空间依赖程度, 我们使用空间单元之间的空间距离. 假设一个空间系统有 n 个空间单元, 空间单元 i 与空间单元 j 之间的距离记为 w_{ij} , 则可定义空间权重矩阵如下:

$$W = \begin{pmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1} & w_{n2} & \cdots & w_{nn} \end{pmatrix}$$

其中一个空间系统中的同一空间单元的距离 w_{ii} 为 0 , 显然, 空间权重矩阵 W 主对角元素为 0 且为对称矩阵. 针对空间权重矩阵的设定, 目前有许多设定方法.

首先, 基于邻接关系定义的空间权重矩阵是最常见的, 即空间单元 i 与空间单元 j 相邻, 则 $w_{ij} = 1$, 反之, 则 $w_{ij} = 0$. 比照象棋中不同棋子的行走路线, 邻接关系可以分为: 两个空间单元有公共边、有公共点无公共边、有公共边或公共点, 该三种情况依次称为车邻接 (rook

contiguity)、象邻接 (bishop contiguity)、皇后邻接 (queen contiguity). 以表1-1中四个省举例.

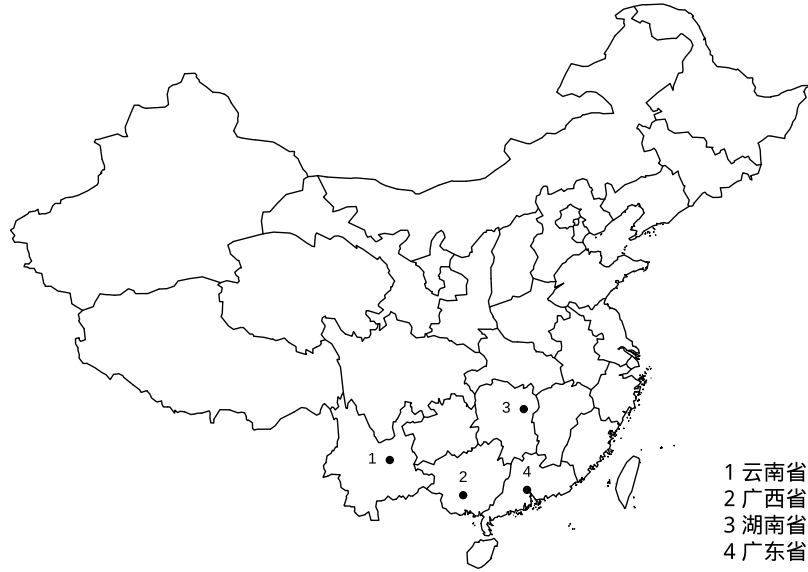


图 1-1 部分中国地图

图1-1给出四个省的空间位置, 其空间权重矩阵为 W_1 , 矩阵 W_1 的第一行表示, 云南省与广西省相邻, 与湖南省和广东省不相邻; 第二行表示, 广西省与其余三省均相邻; 以此类推. 在实践中, 有时对空间权重矩阵进行“行标准化”, 即将矩阵中的每个元素除以其所在行元素之和, 以保证每行元素之和为 1:

$$\tilde{w}_{ij} = \frac{w_{ij}}{\sum_j w_{ij}}$$

将 W_1 进行行标准化后得到 W_2 . 显然, 标准化之后的空间权重矩阵不一定是对称的.

$$W_1 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix} \Rightarrow W_2 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1/3 & 0 & 1/3 & 1/3 \\ 0 & 1/2 & 0 & 1/2 \\ 0 & 1/2 & 1/2 & 0 \end{pmatrix}$$

其次, 基于距离关系也可定义的空间权重矩阵, 即事先给定距离临界值 d , 记空间单元 i 与空间单元 j 的距离为 d_{ij} , 若 $d_{ij} < d$, 则 $w_{ij} = 1$, 否则 $w_{ij} = 0$. 还可以直接用距离之倒数作为空间权重:

$$w_{ij} = \frac{1}{d_{ij}}$$

在上式中, 距离关系可以是地理距离, 比如两个省行政中心的直线距离; 也可以是经济距离, 比如基于运输成本产生的费用; 甚至社交网络中的距离.

还有更多给定空间权重矩阵的方式, 可以参见翟卫欣和程承旗^[4], 英华和游万海^[5], Mawarni 和 Machdi^[6] 等.

§1.2 空间数据模型的定义

生活中的数据多以时间、空间的两个角度呈现,例如某时某地的新冠感染人数、生物种群数量、房价、降雨量、商品成交额,等等.下面以中国历年某些省份人均 GDP 为例,见表1-1,当固定时间 2021 年,就得到此刻下广西省、广东省、湖南省、云南省等空间单元的空间截面数据;当固定空间单元,如广西省,就得到此空间单元的时间序列数据.在 T 个

表 1-1 中国人均 GDP 历年数据 (元/人)

年份	广西省	广东省	湖南省	云南省	...
2021	80976.00	98285.00	69440.00	57686.00	...
2020	72000.00	88210.00	62900.00	51975.20	...
2019	70078.00	86956.00	60104.00	47944.00	...
2018	65534.00	81625.00	54763.00	37136.28	...
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots

数据来源: 国家统计局

时间点(按照先后次序)和 n 个空间得到的数据组成的时空数据集,我们称其为**空间面板数据**(又称时空数据).我们把拟合这三类数据的模型分别称为**空间截面数据模型**、**时间序列模型**和**空间面板数据模型**(又称**时空数据模型**,可继续分为动态或非动态),前两类模型都可以看成是第三类的特殊情形.这些模型统称为空间计量经济模型,或空间数据模型.

为了方便表示,我们引入一些记号.对于正整数 κ ,令 I_κ 记作 $\kappa \times \kappa$ 维单位矩阵, $\mathbf{1}_\kappa$ 记作 $\kappa \times 1$ 维由 1 组成的向量,以及 $\mathbf{0}_\kappa$ 记作 $\kappa \times 1$ 维由 0 组成的向量.令 $A_1 \otimes A_2$ 记作 A_1 和 A_2 两个矩阵的 Kronecker 积.令 $|\cdot|$, $\|\cdot\|$ 和 $\text{tr}(\cdot)$ 分别表示取绝对值, Frobenius 范数以及矩阵的迹.

普通的空间数据模型如下所示,该模型也称为带固定效应的动态含空间自回归误差的空间自回归面板数据模型:

$$\begin{cases} y_t = \lambda y_{t-1} + \rho_1 W_n y_t + x_t \beta + z \gamma + u_t, \\ u_t = \rho_2 M_n u_t + \varepsilon_t, \end{cases} \quad t = 1, 2, \dots, T, \quad (1-1)$$

其中, n 和 T 分别表示是空间单元数和时间单位数, $y_t = (y_{1t}, y_{2t}, \dots, y_{nt})'$ 是 $n \times 1$ 维被解释变量的观测值, $x_t = (x_{1t}, x_{2t}, \dots, x_{nt})'$ 是 $n \times p$ 维解释变量随时间变化的样本资料矩阵, $z = (z_1, z_2, \dots, z_n)'$ 是 $n \times q$ 维解释变量不随时间变化的样本资料矩阵,标量 λ ($|\lambda| < 1$) 是动态效应因子, ρ_j ($|\rho_j| < 1$), $j = 1, 2$ 是空间自相关系数, β 是 $p \times 1$ 维 x_t 的回归系数向量, γ 是 $q \times 1$ 维 z 的回归系数向量, W_n 是解释变量 y_t 的空间邻接权重矩阵, M_n 是扰动项 u_t 的空间邻接权重矩阵,它们都是预先给定的 $n \times n$ 空间权重矩阵,二者可以相等, $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t}, \dots, \varepsilon_{nt})'$ 是 $n \times 1$ 维误差向量, ε_{it} 为独立同分布的随机误差项,且满足 $E(\varepsilon_{it}) = 0$, $\text{Var}(\varepsilon_{it}^2) = \sigma^2$,其中 $0 < \sigma^2 < \infty$.

模型 (1-1) 可写成如下的矩阵形式：

$$\begin{pmatrix} A_n & 0 & \cdots & 0 \\ 0 & A_n & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_n \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_T \end{pmatrix} = \lambda \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_n \end{pmatrix} + \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_T \end{pmatrix} \beta + \begin{pmatrix} z \\ z \\ \vdots \\ z \end{pmatrix} \gamma + \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_T \end{pmatrix}$$

和

$$\begin{pmatrix} B_n & 0 & \cdots & 0 \\ 0 & B_n & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & B_n \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_T \end{pmatrix} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_T \end{pmatrix}$$

或者

$$\begin{cases} (I_T \otimes A_n)Y = \lambda Y_{-1} + X\beta + Z\gamma + u, \\ (I_T \otimes B_n)u = \varepsilon, \end{cases}$$

其中 $Y = (y'_1, y'_2, \dots, y'_T)'$, $Y_{-1} = (y'_0, y'_1, \dots, y'_{T-1})'$, $X = (x'_1, x'_2, \dots, x'_T)'$, $Z = \mathbf{1}_T \otimes z$, $u = (u'_1, u'_2, \dots, u'_T)'$, $\varepsilon = (\varepsilon'_1, \varepsilon'_2, \dots, \varepsilon'_T)'$, $A_n = A_n(\lambda) = I_n - \rho_1 W_n$ 以及 $B_n = B_n(\lambda) = I_n - \rho_2 M_n$.

当 t 为固定时间点且 $\lambda = 0$ 时, 模型 (1-1) 即为带固定效应的含空间自回归误差的空间自回归截面数据模型:

$$\begin{cases} y_t = \rho_1 W_n y_t + x_t \beta + z \gamma + u_t, \\ u_t = \rho_2 M_n u_t + \varepsilon_t. \end{cases}$$

当 t 为固定时间点且 $\lambda = 0, \gamma = 0$ 时, 模型 (1-1) 即为不带固定效应的含空间自回归误差的空间自回归截面数据模型:

$$\begin{cases} y_t = \rho_1 W_n y_t + x_t \beta + u_t, \\ u_t = \rho_2 M_n u_t + \varepsilon_t. \end{cases} \quad (1-2)$$

若 $\rho_1 = \rho_2 = \gamma = 0$, 模型 (1-1) 即为普通线性模型:

$$\begin{cases} y_t = x_t \beta + u_t, \\ u_t = \varepsilon_t. \end{cases}$$

当 t 为时间序列且 $\rho_1 = 0$ 时, 模型 (1-1) 即为带固定效应的动态含空间误差的面板数据模型:

$$\begin{cases} y_t = \lambda y_{t-1} + x_t \beta + z \gamma + u_t, & t = 1, 2, \dots, T, \\ u_t = \rho_2 M_n u_t + \varepsilon_t, \end{cases}$$

§1.3 空间数据模型的研究概况

Anselin^[7] 将空间计量经济学定义为: 在区域经济模型中处理由于空间因素导致的特殊性质的一系列方法. 该方法是对截面数据和面板数据中空间相互性 (地理单元之间存在的交互作用)、空间异质性 (不同地理单元之间存在的差异)、以及动态效应 (过去的空间单元对现在的空间单元的影响) 的定量研究.

Cliff 和 Ord^[2] 首次考虑到空间效应, 提出空间自回归模型, 引出了空间计量经济学的开端. Anselin^[7] 从计量经济学角度处理空间效应, 将空间效应视为计量经济学模型中普通问题的特例, 并系统整理当时研究空间自回归模型的方法. Cressie^[8] 对可能的空间数据类型进行划分, 空间计量经济学的研究逐渐由边缘走向主流. Kelejian 和 Prucha^[9] 改进了传统的 Moran I 检验, 并提出了二次线性中心极限定理. Kelejian 和 Prucha^[10] 提出使用广义矩估计方法研究空间自回归模型. Lee^[11] 证明了空间自回归模型的拟极大似然估计的渐近性质. Kelejian 和 Prucha^[12] 提出使用广义空间两段最小二乘法估计含空间自回归误差的空间自回归模型的参数. Kelejian 等^[13] 推荐使用工具变量的迭代版本估计空间自回归误差的空间自回归模型的参数. Kelejian 和 Prucha^[14] 研究了在空间模型中方差-协方差矩阵的异方差和自相关系数的相合估计. Arraiz 等^[15] 通过蒙特卡罗模拟表明, 在异方差情形下拟极大似然估计量不是相合估计. 同时证明在异方差情形下工具变量法导出的估计量为相合估计.

以上提到的都是学者们对于**空间截面数据模型**的研究, 并没有考虑时间维度, 既考虑空间又考虑时间的模型称为**空间面板数据模型**. 空间面板数据模型的理论研究始于 2000 年, 该模型考虑了空间相关性和时间依赖性, 将传统的时间序列方法、横截面数据方法以及普通面板数据模型进行了综合. Kapoor 等^[16] 使用广义矩估计和广义最小二乘估计分别研究了具有空间自相关误差项的线性回归空间面板数据模型, 并证明其参数估计的渐近正态性. Lee 和 Yu^[17] 提出使用极大似然估计方法研究具有空间滞后和空间扰动的面板模型. Lee 和 Yu^[18] 研究了具有固定效应的空间自回归面板数据模型的拟极大似然估计并证明了该估计的一致性与渐近正态性. 面板数据由于具有异质性, 考虑在模型中加入**个体效应项**来控制所有不随时间变化且反应空间特质的量, 个体效应又分为固定效应和随机效应. Fingleton 和 Legallo^[19] 用此估计方法研究了带空间移动平均误差项的空间面板数据模型. Lee 和 Yu^[20] 提出使用拟极大似然方法研究带固定效应的空间面板数据模型, 并提出正交转换的间接估计方法. 对于含有个体固定效应的空间误差面板数据模型, 文利霞^[21] 研究了该模型参数的

拟极大似然估计及其渐近性质. 很多时候, 个体效应并不能完全解释空间异质性, 于是考虑将模型系数视为变化的, 这样系数随着观测个体或时间的变化而变化的模型, 我们称为**变系数空间面板数据模型**. 该模型以 Anselin^[7] 提出的空间看似无关回归模型为雏形, 类似截面到面板的设定方式提出. 邓明^[22] 提出使用极大似然估计变系数空间自回归面板数据模型的参数, 并且在一定条件下证明了估计量的一致性和渐近正态性.

前面所提及的模型均为静态空间面板数据模型, 并没有考虑过去的空间对现在的空间也有影响, Anselin^[23] 提出加入动态元素于其中得到**动态空间面板数据模型**. Lee 和 Yu^[24] 使用极大似然方法和拟极大似然方法研究一种具有时间和个体固定效应的空间动态面板数据模型, 得到了当个体数和时间都很大时, 这些方法得到的估计具有一致性和渐近正态性. Yu 和 Lee^[25] 使用拟极大似然估计方法研究具有空间自相关误差项的单位根空间动态面板数据模型, 并证明其拟极大似然估计具有渐近正态性. Baltagi^[26] 等提出一种四阶段空间广义矩估计方法研究具有空间移动平均随机效应的时空动态面板数据模型, Su 和 Yang^[27] 使用拟极大似然估计方法研究含空间误差的动态面板数据模型, 其中讨论了观测值初值为内生或外生的情形. 目前关于内生性空间动态面板数据模型并不多.

§1.4 经验似然方法的研究概况

目前, 对空间计量模型的估计, 使用比较多的方法是极大似然法 (如文献 [2, 8, 17, 24])、拟极大似然法 (如文献 [11, 18, 20–22, 24, 25, 27])、两阶段最小二乘法 (如文献 [12, 16]) 或广义矩估计 (如文献 [10, 16, 19, 26]), 然而, 在渐近分布的渐近协方差未知时, 使用这些参数方法构造空间计量模型中参数的置信域是不容易的, 更重要的是, 估计的渐近协方差可能会影响基于正态近似的模型参数置信域的精度.

Owen^[3, 28] 提出的经验似然方法是一种非参数统计推断方法, 该方法构造置信区间除了有域保持性, 变换不变性及置信域的形状由数据自行决定等优点外, 还有 Bartlett 纠偏性及无需构造枢轴统计量等优点克服了参数方法的缺点, Qin 和 Lawless^[29] 在 Owen 基础上给出了普通情况下的经验似然估计方法. 起初, 经验似然方法都应用在非空间模型上, 比如 Owen^[30] 提出使用经验似然方法研究线性模型, Kolaczyk^[31] 使用经验似然方法研究广义线性模型, Qin^[32] 使用经验似然方法研究部分线性模型等, Chen 和 Keilegom^[33] 发表了一篇使用经验似然方法研究回归模型的综述, 更多的研究可以参考文献 [34–37].

最近, Jin 和 Lee^[38] 以及 Qin^[39] 分别独立发现误差序列的线性-二次型可以转化为鞅差序列的线性形式, 从而构造含空间误差的空间自回归模型的估计函数, 并使用经验似然方法构造该模型的经验似然比统计量, 证明其经验似然比统计量的极限分布是渐近卡方分布, 经验似然方法在空间模型上才得以推广. 利用误差序列的线性-二次型转化为鞅差序列的线性形式的思想, 基于广义矩估计方法得到估计方程后, Qin 和 Lei^[40] 以及 Qin^[41] 分别使用经验

似然方法研究不含空间误差的空间自回归截面数据模型及含空间自回归误差的空间自回归截面数据模型；基于极大似然方法得到估计方程后, Li 等^[42] 使用经验似然方法研究含空间自回归误差的非参数回归截面数据模型, Rong 等^[43] 使用经验似然方法研究含空间自回归误差的空间自回归面板数据模型, 曾庆樊等^[44] 使用经验似然方法研究时变系数空间面板数据模型. 理论上, 经验似然方法能在空间计量模型上得到应用, 然而实际中仍存在问题仍有许多问题值得研究, 例如凸包问题和维数爆炸问题.

凸包问题会导致经验似然比无法计算的问题. 经验似然比统计量存在的前提条件是 0 必须在估计函数集合的凸包内部, 此时经验似然比才存在, Tsao^[45] 在小样本下证明了有较高的概率使得 0 不在估计函数集的凸包内部, Owen^[46] 将凸包问题作为经验似然未来之一的挑战问题提出, 此后, 有许多学者对此问题进行讨论, 提出各种方法, 这其中比较典型的是 Bartolucci^[47] 提出的惩罚经验似然, Chen 等^[48] 提出的调整经验似然, Emerson 和 Owen^[49] 提出的平衡添加经验似然方法. 在本文第三章中, 我们使用文献 [48] 中 Chen 等提出的调整经验似然思想, 在 SARAR 模型下, 讨论调整经验似然比统计量的构造及其相应的性质, 并与普通经验似然进行模拟比较.

维数爆炸问题会导致传统统计方法失效. 处理这类问题的非参数方法主要依赖大样本理论, 高维情形下会出现样本量相比数据维数显得较少的情况, 从而大样本理论处理高维数据方法失效. 另外, 传统的经验似然方法在处理高维数据时不能满足稳健性要求; 高维导致样本量变少使得理论上的渐近性难以实现; 维数的增加导致数据的计算量迅速上升等等. 在经验似然领域中, 处理高维数据的常见手段分两种, 一类以修正原来的经验似然比统计量为基础, 另一类采用降维的思想. 第一类中, Shi^[50] 指出, 当协变量维数以合理的速度趋于无穷大时, 仍可利用经验似然方法构造高维线性模型参数的置信域, 不过修正的经验似然统计量的渐近分布为正态分布而非卡方. Hjort 等^[51] 从理论上拓展了经验似然方法的适用范围, 即当 $p = o_p(n^{1/3}) \rightarrow \infty$ 时, 在一定条件下, 经验似然比统计量渐近分布为正态分布. Chen^[52] 研究了均值模型下数据维数对经验似然的影响, 并证明当 $p = o_p(n^{1/2}) \rightarrow \infty$ 时, 比文献 [51] 更少的约束条件下, 经验似然比统计量渐近分布为正态分布. Li 等^[53] 研究了高维变系数部分线性模型的经验似然推断, 提出纠偏经验似然方法, 并证明其统计量渐近分布为正态. 更多文献可见 [54–56] 等等. 这类方法虽然使得统计量不发散, 但都以 $p < n$ 为前提在应用中有所局限.

第二类中, 在保证重要信息不损失的情形下对高维进行降维, 其中著名的方法就是惩罚经验似然. 惩罚的思想是在进行参数估计的同时, 利用惩罚函数将较小的系数估计值压缩为零, 而将系数估计值较大的保留, 在估计出系数的同时选择出重要变量, 这可以同时实现变量选择和系数估计两个目标. 惩罚变量选择普遍采用“损失函数 + 惩罚函数”的方式, 惩罚函数的选择有很多, 比如熵惩罚、桥回归^[57]、LASSO 惩罚^[58]、岭回归、硬门限惩罚^[59] 等等,

一个关键问题是如何选择惩罚函数. Antoniadis 和 Fan^[60] 与 Fan 和 Li^[61] 对如何选择惩罚函数提出深刻的见解, 他们在文献 [61] 中提出 SCAD 惩罚方法, 该方法得到的估计量具有无偏性、稀疏性、连续性. 类似地, 惩罚经验似然也使用“经验似然比函数 + 惩罚函数”方式. Tang 和 Leng^[62] 首次将惩罚经验似然方法于高维分析多变量的均值向量和线性模型的高维回归系数, 证明惩罚经验似然具有与普通经验似然的类似性质, 例如, 依靠数据确定置信区域的形状和取向, 无须估计共协方差, 相同的渐近分布等, 研究高维数据的非参数方法开辟新的道路. Leng 和 Tang^[63] 将高维惩罚经验似然方法应用于普通估计方程的参数估计和变量选择, 并证明惩罚经验似然方法具有 oracle 特征. Lahiri 和 Mukhopadhyay^[64] 推广高维中一种惩罚经验似然方法可适用于 $p > n$ 的情形, 成功应用于总体均值模型. 至此, 高维经验似然方法备受学者们关注, 比如 Fang 等^[56] 研究了半参数模型的高维惩罚经验似然, Yan 和 Chen^[65] 研究了高维广义线性模型的惩罚拟似然 SCAD 估计, 更多可参见文献 [64–66] 等等. 在本文第四章中, 我们基于降维的思想给出一个新的经验似然方法, 在高维 SARAR 模型下, 讨论新的经验似然比统计量的构造及其相应的性质, 并与普通经验似然进行模拟比较.

直至目前, 经验似然方法在空间模型中的应用都是不含个体效应的, 因为鞅差变换方法使用的前提是估计方程为模型误差序列的线性-二次型的形式, 带固定效应的空间模型经验似然推断将留给后来的研究.

§1.5 含空间自回归误差的空间自回归模型的普通经验似然推断

本文将研究不带固定效应的空间截面数据模型 (1-2), 并简写为 SARAR 模型, 用新的符号记为:

$$Y_n = \rho_1 W_n Y_n + X_n \beta + u_{(n)}, u_{(n)} = \rho_2 M_n u_{(n)} + \epsilon_{(n)}, \quad (1-3)$$

其中, n 是空间单元数量, $X_n = (x_1, x_2, \dots, x_n)'$ 是 $n \times p$ 维解释变量的样本资料矩阵, β 是 $p \times 1$ 维 X_n 的回归系数向量, $Y_n = (y_1, y_2, \dots, y_n)'$ 是 $n \times 1$ 维响应变量, $\rho_j, j = 1, 2$ 是空间自回归系数且 $|\rho_j| < 1, j = 1, 2$, W_n 是解释变量 Y_n 的空间邻接权重矩阵, M_n 是扰动项 $u_{(n)}$ 的空间邻接权重矩阵, 它们都是已知的 $n \times n$ 空间邻接权重矩阵 (非随机), 二者可以相等, $\epsilon_{(n)}$ 是 $n \times 1$ 维空间误差向量, 且满足

$$E\epsilon_{(n)} = 0, \text{Var}(\epsilon_{(n)}) = \sigma^2 I_n.$$

SARAR 模型对存在空间依赖性的数据有较好的解释作用, 无论是滞后项存在空间依赖性还是扰动项存在空间依赖性, 空间依赖性由 ρ_j 刻画, ρ_1 度量空间滞后项 $W_n Y_n$ 对解释变量 Y_n 的影响, 也就是 Y_n 之间的空间依赖性, 即相邻空间之间可能存在扩散、溢出等效应, ρ_2 度量空间误差滞后项 $M_n u_{(n)}$ 对扰动项 $u_{(n)}$ 的影响, 也就是不包含在 X_n 中且对 Y_n 有影响的

遗漏变量的空间依赖性.

记 $A_n(\rho_1) = I_n - \rho_1 W_n$, $B_n(\rho_2) = I_n - \rho_2 M_n$, 并且假设 $A_n(\rho_1)$ 和 $B_n(\rho_2)$ 是非奇异矩阵, 于是由模型 (1-3) 可知:

$$Y_n = A_n^{-1}(\rho_1)X_n\beta + A_n^{-1}(\rho_1)B_n^{-1}(\rho_2)\epsilon_{(n)}.$$

此时, 假设 $\epsilon_{(n)}$ 服从标准正态分布的, 则 Y_n 服从期望为 $A_n^{-1}(\rho_1)X_n\beta$, 方差为 $A_n^{-1}(\rho_1)B_n^{-1}(\rho_2)\sigma^2 I_n[A_n^{-1}(\rho_1)B_n^{-1}(\rho_2)]'$ 的正态分布, 于是, Y_n 的拟似然函数为:

$$F = (2\pi)^{-\frac{n}{2}} |B_n(\rho_2)| |A_n(\rho_1)| (\sigma^2)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2\sigma^2} \epsilon'_{(n)} \epsilon_{(n)}\right\},$$

其中 $\epsilon_{(n)} = B_n(\rho_2)\{A_n(\rho_1)Y_n - X_n\beta\}$, 进而, Y_n 的拟对数似然函数为:

$$L = -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log \sigma^2 + \log |A_n(\rho_1)| + \log |B_n(\rho_2)| - \frac{1}{2\sigma^2} \epsilon'_{(n)} \epsilon_{(n)},$$

令 $G_n = B_n(\rho_2)W_n A_n^{-1}(\rho_1)B_n^{-1}(\rho_2)$, $H_n = M_n B_n^{-1}(\rho_2)$, $\tilde{G}_n = \frac{1}{2}(G_n + G'_n)$ 及 $\tilde{H}_n = \frac{1}{2}(H_n + H'_n)$, 对对数似然函数 L 求偏导数可得:

$$\begin{aligned} \partial L / \partial \beta &= \frac{1}{\sigma^2} X'_n B'_n(\rho_2) \epsilon_{(n)}, \\ \partial L / \partial \rho_1 &= \frac{1}{\sigma^2} \{B_n(\rho_2)W_n A_n^{-1}(\rho_1)X_n\beta\}' \epsilon_{(n)} + \frac{1}{\sigma^2} \{\epsilon'_{(n)} \tilde{G}_n \epsilon_{(n)} - \sigma^2 \text{tr}(\tilde{G}_n)\}, \\ \partial L / \partial \rho_2 &= \frac{1}{\sigma^2} \{\epsilon'_{(n)} \tilde{H}_n \epsilon_{(n)} - \sigma^2 \text{tr}(\tilde{H}_n)\}, \\ \partial L / \partial \sigma^2 &= \frac{1}{2\sigma^4} \{\epsilon'_{(n)} \epsilon_{(n)} - n\sigma^2\}. \end{aligned}$$

令上述偏导数等于 0, 可以获得以下估计方程:

$$X'_n B'_n(\rho_2) \epsilon_{(n)} = 0, \quad (1-4)$$

$$\{B_n(\rho_2)W_n A_n^{-1}(\rho_1)X_n\beta\}' \epsilon_{(n)} + \{\epsilon'_{(n)} \tilde{G}_n \epsilon_{(n)} - \sigma^2 \text{tr}(\tilde{G}_n)\} = 0, \quad (1-5)$$

$$\epsilon'_{(n)} \tilde{H}_n \epsilon_{(n)} - \sigma^2 \text{tr}(\tilde{H}_n) = 0, \quad (1-6)$$

$$\epsilon'_{(n)} \epsilon_{(n)} - n\sigma^2 = 0. \quad (1-7)$$

记 \tilde{g}_{ij} , \tilde{h}_{ij} , b_i 和 s_i 分别表示矩阵 \tilde{G}_n 第 i 行第 j 列的元素, 矩阵 \tilde{H}_n 第 i 行第 j 列的元素, 矩阵 $X'_n B'_n(\rho_2)$ 第 i 列向量和向量 $B_n(\rho_2)W_n A_n^{-1}(\rho_1)X_n\beta$ 的第 i 个元素, ϵ_i 表示 $\epsilon_{(n)}$ 的第 i 个元素, 并且规定, 当求和符号的上标等于 0 时我们令该和为 0. 为了处理 (1-5) 和 (1-6) 中的二次型形式, 需要引入文献 [9] 中介绍的鞅差序列, 定义 σ -域: $\mathcal{F}_0 = \{\emptyset, \Omega\}$, $\mathcal{F}_i =$

$\sigma(\epsilon_1, \epsilon_2, \dots, \epsilon_i), 1 \leq i \leq n$. 令

$$\tilde{Y}_{in} = \tilde{g}_{ii}(\epsilon_i^2 - \sigma^2) + 2\epsilon_i \sum_{j=1}^{i-1} \tilde{g}_{ij}\epsilon_j, \quad \tilde{Z}_{in} = \tilde{h}_{ii}(\epsilon_i^2 - \sigma^2) + 2\epsilon_i \sum_{j=1}^{i-1} \tilde{h}_{ij}\epsilon_j. \quad (1-8)$$

则 $\mathcal{F}_{i-1} \subseteq \mathcal{F}_i$, \tilde{Y}_{in} 是 \mathcal{F}_i -可测的, 并且 $E(\tilde{Y}_{in}|\mathcal{F}_{i-1}) = 0$. 因此, $\{\tilde{Y}_{in}, \mathcal{F}_i, 1 \leq i \leq n\}$ 和 $\{\tilde{Z}_{in}, \mathcal{F}_i, 1 \leq i \leq n\}$ 构成两个鞅差序列, 且

$$\epsilon'_{(n)} \tilde{G}_n \epsilon_{(n)} - \sigma^2 \text{tr}(\tilde{G}_n) = \sum_{i=1}^n \tilde{Y}_{in}, \quad \epsilon'_{(n)} \tilde{H}_n \epsilon_{(n)} - \sigma^2 \text{tr}(\tilde{H}_n) = \sum_{i=1}^n \tilde{Z}_{in}. \quad (1-9)$$

根据 (1-4) 到 (1-9), 可知得分函数为:

$$\omega_i(\theta) = \begin{pmatrix} b_i \epsilon_i \\ \tilde{g}_{ii}(\epsilon_i^2 - \sigma^2) + 2\epsilon_i \sum_{j=1}^{i-1} \tilde{g}_{ij}\epsilon_j + s_i \epsilon_i \\ \tilde{h}_{ii}(\epsilon_i^2 - \sigma^2) + 2\epsilon_i \sum_{j=1}^{i-1} \tilde{h}_{ij}\epsilon_j \\ \epsilon_i^2 - \sigma^2 \end{pmatrix}_{(p+3) \times 1},$$

由 Owen^[28], 得到关于 $\theta = (\beta', \rho_1, \rho_2, \sigma^2)' \in R^{p+3}$ 的经验似然比统计量:

$$L_n(\theta) = \sup_{p_i, 1 \leq i \leq n} \prod_{i=1}^n (np_i),$$

此处, $\{p_i\}$ 满足:

$$p_i \geq 0, 1 \leq i \leq n, \sum_{i=1}^n p_i = 1, \sum_{i=1}^n p_i \omega_i(\theta) = 0.$$

由拉格朗日乘子可以算得:

$$p_i = \frac{1}{n} \cdot \frac{1}{1 + \lambda'(\theta) \omega_i(\theta)}, 1 \leq i \leq n,$$

及

$$\ell_n(\theta) \triangleq -2 \log L_n(\theta) = 2 \sum_{i=1}^n \log\{1 + \lambda'(\theta) \omega_i(\theta)\},$$

其中 $\lambda(\theta) \in R^{p+3}$ 为以下方程的解:

$$\frac{1}{n} \sum_{i=1}^n \frac{\omega_i(\theta)}{1 + \lambda'(\theta) \omega_i(\theta)} = 0. \quad (1-10)$$

在本文中, 记 $\mu_j = E(\epsilon_1^j), j = 3, 4$, 用 $\text{Vec}(\text{diag} A)$ 表示由矩阵 A 的对角线上的元素构成的列向量. 为了获得调整经验似然比统计量 $\ell_n(\theta)$ 的渐近分布, 需要如下假设条件:

A1. $\{\epsilon_i, 1 \leq i \leq n\}$ 是均值为 0, 方差有限的独立同分布随机变量序列, 且存在 $\eta_1 > 0$, 使 $E|\epsilon_1|^{4+\eta_1} < \infty$.

A2. $W_n, M_n, A_n^{-1}(\rho_1), B_n^{-1}(\rho_2)$ 及 $\{x_i\}$ 满足如下条件：

(i) 矩阵 $W_n, M_n, A_n^{-1}(\rho_1)$, 和 $B_n^{-1}(\rho_2)$ 的行元素的绝对值之和与列元素的绝对值之和均一致有界；

(ii) $\{x_i\}, i = 1, 2, \dots, n$ 一致有界.

A3. 存在常数 $c_j > 0, j = 1, 2$ 使得 $0 < c_1 \leq \lambda_{\min}(n^{-1}\Sigma_{p+3}) \leq \lambda_{\max}(n^{-1}\Sigma_{p+3}) \leq c_2 < \infty$, 其中 $\lambda_{\min}(A)$ 和 $\lambda_{\max}(A)$ 分别表示矩阵 A 的最小和最大的特征值, 且

$$\Sigma_{p+3} = \Sigma'_{p+3} = Cov \left\{ \sum_{i=1}^n \omega_i(\theta) \right\} = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} & \Sigma_{13} & \Sigma_{14} \\ \Sigma_{21} & \Sigma_{22} & \Sigma_{23} & \Sigma_{24} \\ \Sigma_{31} & \Sigma_{32} & \Sigma_{33} & \Sigma_{34} \\ \Sigma_{41} & \Sigma_{42} & \Sigma_{43} & \Sigma_{44} \end{pmatrix}, \quad (1-11)$$

其中,

$$\begin{aligned} \Sigma_{11} &= \sigma^2 \{B_n(\rho_2)X_n\}' B_n(\rho_2)X_n, \\ \Sigma_{12} &= \sigma^2 \{B_n(\rho_2)X_n\}' B_n(\rho_2)W_n A_n^{-1}(\rho_1)X_n \beta + \mu_3 \{B_n(\rho_2)X_n\}' Vec(diag \tilde{G}_n), \\ \Sigma_{13} &= \mu_3 \{B_n(\rho_2)X_n\}' Vec(diag \tilde{H}_n), \quad \Sigma_{14} = \mu_3 \{B_n(\rho_2)X_n\}' \mathbf{1}_n, \\ \Sigma_{22} &= 2\sigma^4 tr(\tilde{G}_n^2) + \sigma^2 \{B_n(\rho_2)W_n A_n^{-1}(\rho_1)X_n \beta\}' B_n(\rho_2)W_n A_n^{-1}(\rho_1)X_n \beta \\ &\quad + (\mu_4 - 3\sigma^4) \|Vec(diag \tilde{G}_n)\|^2 + 2\mu_3 \{B_n(\rho_2)W_n A_n^{-1}(\rho_1)X_n \beta\}' Vec(diag \tilde{G}_n), \\ \Sigma_{23} &= 2\sigma^4 tr(\tilde{G}_n \tilde{H}_n) + (\mu_4 - 3\sigma^4) Vec'(diag \tilde{G}_n) Vec(diag \tilde{H}_n) \\ &\quad + \mu_3 \{B_n(\rho_2)W_n A_n^{-1}(\rho_1)X_n \beta\}' Vec(diag \tilde{H}_n), \\ \Sigma_{24} &= (\mu_4 - \sigma^4) tr(\tilde{G}_n) + \mu_3 \{B_n(\rho_2)W_n A_n^{-1}(\rho_1)X_n \beta\}' \mathbf{1}_n, \\ \Sigma_{33} &= 2\sigma^4 tr(\tilde{H}_n^2) (\mu_4 - 3\sigma^4) \|Vec(diag \tilde{H}_n)\|^2, \\ \Sigma_{34} &= (\mu_4 - \sigma^4) tr(\tilde{H}_n), \quad \Sigma_{44} = n(\mu_4 - \sigma^4). \end{aligned}$$

注 1 (1-11) 的证明见文献 [39] 的引理 3. 条件 A1 至 A3 是空间模型的常见假设, 如 A1 和 A2 在文献 [11] 假设 1、4、5 和 6 中使用, A3 在文献 [12] 和文献 [9] 证明定理 1 中用到.

由 Qin^[39] 可得普通经验似然统计量的极限理论：

定理 1.1 在假设条件 (A1)-(A3) 及模型 (1-3) 下, 当 $\theta = \theta_0, n \rightarrow \infty$, 有

$$\ell_n(\theta_0) \xrightarrow{d} \chi_{p+3}^2,$$

其中, χ_{p+3}^2 表示自由度为 $p+3$ 的卡方分布.

取定 $\alpha, 0 < \alpha < 1$, 设 $z_\alpha(p+3)$ 满足 $P(\chi_{p+3}^2 > z_\alpha(p+3)) = \alpha$, 由定理1.1 可确定基于普通经验似然方法下 θ_0 的置信水平为 $1 - \alpha$ 的渐近置信域为:

$$\{\theta : \ell_n(\theta) \leq z_\alpha(p+3)\}.$$

§1.6 本文的主要内容和创新点

本文研究含空间自相关误差的空间自回归模型的调整经验似然以及高维经验似然推断问题, 并证明其相应的经验似然比统计量服从卡方分布.

研究动机: 空间数据模型应用十分广泛, 现有的研究空间数据模型主要是广义矩估计, 极大似然估计和拟极大似然估计, 针对经验似然估计的研究相对较少, 而经验似然方法的应用比较广泛, 那么对于空间数据模型的经验似然研究是十分有必要的.

本文的主要内容:

第一章为空间计量经济模型的经验似然研究进展, 简要地介绍空间权重矩阵的常见设定、空间数据模型的定义、该模型在传统统计方法上的研究概况以及在经验似然方法上的研究概况, 最后给出本文研究的主要内容和创新点.

第二章主要介绍普通经验似然方法的原理、本文重点研究的空间截面模型的定义以及普通经验似然方法在该模型的推广.

第三章主要研究含空间自相关误差的空间自回归模型的调整经验似然推断问题. 我们利用调整经验似然方法, 构造出含空间自相关误差的空间自回归模型的调整经验似然比统计量, 证明调整经验似然统计量的极限分布为卡方分布, 并模拟比较调整经验似然与普通经验似然方法的优劣, 模拟结果表明, 调整经验似然比普通经验似然置信域的覆盖精度更高, 计算速度更快. 在样本容量较大时, 无论误差项是否服从正态分布, 普通经验似然方法和调整经验似然方法均可, 在样本容量较小时, 我们更推荐使用调整经验似然方法.

第四章主要研究含空间自相关误差的空间自回归模型的高维经验似然推断问题. 我们利用高维经验似然方法, 构造出含空间自相关误差的空间自回归模型的高维经验似然比统计量, 证明高维经验似然统计量的极限分布为卡方分布, 并模拟比较高维经验似然与普通经验似然方法的优劣, 模拟结果表明, 当维数逐渐变大的时候, 普通经验似然已经失效, 而高维经验似然方法仍有非常接近名义水平的覆盖精度, 且计算速度非常快.

本文的创新点:

1. 本文首次考虑含空间自相关误差的空间自回归模型的调整/高维经验似然的推断问题, 并得到调整/高维经验似然统计量的极限分布;

2. 本文设计了数值模拟对主要的结论进行验证, 同时与普通的经验似然方法相比较, 我们提出的调整/高维经验似然方法下的覆盖率更接近名义水平, 计算速度更快;

3. 本文在普通经验似然方法中加入了伪观测值, 从而提出调整经验似然, 解决普通经验似然凸包不含 0 时, 经验似然比没有数值解的问题.

4. 本文将普通经验似然中高维约束巧妙地变换为一维约束, 由此提出高维经验似然方法, 解决维数灾难问题.

第二章 含空间自回归误差的空间自回归模型的调整经验似然推断

§2.1 调整经验似然原理

我们首先给出估计方程的定义：假设 $X \in R^d$ 是总体， X_1, X_2, \dots, X_n 来自 X 的 n 个独立同分布的观测值，有 r 个已知函数 $g_j(x, \theta), 1 \leq j \leq r$ 满足

$$Eg_j(X, \theta) = 0, 1 \leq j \leq r,$$

其中 $\theta \in \Theta \in R^p$ ，我们称 $g_j(x, \theta), 1 \leq j \leq r$ 为 r 个 p 维估计方程。

Owen^[28] 提出的 (普通) 经验似然方法，假设 0 在 $g(X_i, \theta), 1 \leq i \leq n$ 构成的凸包里并定义得分函数为：

$$g(x, \theta) = (g_1(x, \theta), g_2(x, \theta), \dots, g_r(x, \theta))^T, x \in R^d, \theta \in \Theta,$$

以及 (普通) 经验似然比统计量为：

$$\ell_E(\theta) = \sum_{i=1}^n \log\{1 + t^T(\theta)g(X_i, \theta)\},$$

其中 $t(\theta)$ 是下面方程的解：

$$\sum_{i=1}^n \frac{g(X_i, \theta)}{1 + t^T(\theta)g(X_i, \theta)} = 0.$$

根据 Qin 和 Lawless^[29] 的推导，我们可以得到如下关于该经验似然比统计量 $\ell_E(\theta)$ 的渐近理论，见定理2.1.

定理 2.1 假设 $E\|g(X, \theta_0)\|^3 < \infty$ 并且 $Cov(g(X, \theta_0))$ 正定，其中 $\|a\|$ 是 R^r 的 L_2 范数以及 θ_0 是参数 θ 的真值，那么对于固定的 p 和 r ，随着 $n \rightarrow \infty$ ，有

$$2\ell_E(\theta_0) \xrightarrow{d} \chi_r^2,$$

其中 χ_r^2 自由度为 r 的卡方分布。

为了解决凸包问题，Chen^[48] 提出在 $\{g_i(X, \theta), 1 \leq i \leq n\}$ 构成的凸包中加入一个伪观测值， $g_{n+1} = g_{n+1}(\theta) = -a_n \bar{g}_n$ ， $\bar{g}_n = \bar{g}_n(\theta) = n^{-1} \sum_{i=1}^n g_i(X, \theta)$ ， $a_n > 0$ 。 a_n 称为调整参数。在新的估计方程组成的凸包中 $\{g_i(X, \theta), 1 \leq i \leq n+1\}$ 必然包含 0。

定义调整经验似然的得分函数：

$$g(x, \theta) = (g_1(x, \theta), g_2(x, \theta), \dots, g_{n+1}(x, \theta))^T, x \in R^d, \theta \in \Theta,$$

以及调整经验似然比统计量为：

$$\ell_E^*(\theta) = \sum_{i=1}^{n+1} \log\{1 + \hat{t}^T(\theta)g(X_i, \theta)\},$$

其中 $\hat{t}(\theta)$ 是下面方程的解：

$$\sum_{i=1}^{n+1} \frac{g(X_i, \theta)}{1 + \hat{t}^T(\theta)g(X_i, \theta)} = 0.$$

显然，如果 $\bar{g}_n = 0$ ，那么 $\ell_E^*(\theta) = \ell_E(\theta) = 0$ 。如果给定 θ 值使得 $\bar{g}_n(\theta) \approx 0$ ，那么 $\ell_E^*(\theta) \approx \ell_E(\theta)$ 。基于定理2.1不难得到调整经验似然比统计量 $\ell_E^*(\theta)$ 的渐近分布，见定理2.2。

定理 2.2 假设定理2.1的条件都满足，则对于固定的 p 和 r ，当 $a_n = o_p(n^{2/3})$ ， $n \rightarrow \infty$ 时，有

$$2\tilde{\ell}_E(\theta_0) \xrightarrow{d} \chi_{r_0}^2,$$

其中 $\chi_{r_0}^2$ 表示自由度为 r_0 的卡方分布。

§2.2 调整经验似然比统计量的极限分布

由上述理论可知，关于 $\theta = (\beta', \rho_1, \rho_2, \sigma^2)' \in R^{p+3}$ 的调整经验似然比统计量：

$$L_n^*(\theta) = \sup_{p_i, 1 \leq i \leq n+1} \prod_{i=1}^{n+1} \{(n+1)p_i\}, \quad (2-1)$$

此处， $\{p_i\}$ 满足：

$$p_i \geq 0, 1 \leq i \leq n+1, \sum_{i=1}^{n+1} p_i = 1, \sum_{i=1}^{n+1} p_i \omega_i(\theta) = 0.$$

其中， $\omega_{n+1}(\theta) = -a_n \bar{\omega}_n$ ， $\bar{\omega}_n \triangleq \bar{\omega}_n(\theta) = n^{-1} \sum_{i=1}^n \omega_i(\theta)$ ， $a_n > 0$ 。由拉格朗日乘子可以算得：

$$p_i^* = \frac{1}{n+1} \cdot \frac{1}{1 + \lambda'(\theta)\omega_i(\theta)}, 1 \leq i \leq n+1,$$

及

$$\ell_n^*(\theta) \triangleq -2 \log L_n^*(\theta) = 2 \sum_{i=1}^{n+1} \log\{1 + \lambda'(\theta)\omega_i(\theta)\}, \quad (2-2)$$

其中 $\lambda(\theta) \in R^{p+3}$ 为以下方程的解：

$$\frac{1}{n+1} \sum_{i=1}^{n+1} \frac{\omega_i(\theta)}{1 + \lambda'(\theta)\omega_i(\theta)} = 0. \quad (2-3)$$

调整经验似然的主要结论如下：

定理 2.3 在假设条件 (A1)-(A3) 及模型 (1-3) 下, 当 $\theta = \theta_0, n \rightarrow \infty$, 且 $a_n = o_p(n^{2/3})$ 时, 有

$$\ell_n^*(\theta_0) \xrightarrow{d} \chi_{p+3}^2,$$

其中, χ_{p+3}^2 表示自由度为 $p+3$ 的卡方分布.

取定 $\alpha, 0 < \alpha < 1$, 设 $z_\alpha(p+3)$ 满足 $P(\chi_{p+3}^2 > z_\alpha(p+3)) = \alpha$, 由定理 2.3 可确定基于调整经验似然方法下 θ_0 的置信水平为 $1 - \alpha$ 的渐近置信域为:

$$\{\theta : \ell_n^*(\theta) \leq z_\alpha(p+3)\}.$$

考虑检验问题: $H_0 : \theta = \theta_0, H_1 : \theta \neq \theta_0$ 由定理 2.3 知显著水平为 α 的渐近否定域为:

$$W \triangleq \{\theta : \ell_n^*(\theta) > z_\alpha(p+3)\}.$$

定理 2.4 条件同定理 1, 当 $\theta \neq \theta_0, n \rightarrow \infty$ 且 $\|E\{\bar{\omega}_n(\theta)\}\| > 0$ 时, 有

$$n^{-1/3}\ell_n^*(\theta) \xrightarrow{p} \infty, n^{-1/3}\ell_n(\theta) \xrightarrow{p} \infty,$$

其中, $\bar{\omega}_n(\theta) = n^{-1} \sum_{i=1}^n \omega_i(\theta)$.

由定理 2.4 可知, 定理 2.3 确定的检验规则在 $\theta \neq \theta_0$ 时的渐近功效为: $P(W|H_1) = P_{H_1}(\ell_n^*(\theta) > z_\alpha(p+3)) = P_{H_1}(n^{-1/3}\ell_n^*(\theta) > n^{-1/3}z_\alpha(p+3)) \rightarrow 1$.

§2.3 模拟结果

我们通过模拟比较 EL 和 AEL 优劣, 给定置信水平 $1 - \alpha = 0.95$, 分别给出 $\ell_n(\theta_0) \leq z_{0.05}(p+3)$ 和 $\ell_n^*(\theta_0) \leq z_{0.05}(p+3)$ 在 5000 次模拟中出现的比例, 其中 θ_0 是 θ 的真实值. 在模拟中, 使用如下模型: $Y_n = \rho_1 W_n Y_n + X_n \beta + u(n), u(n) = \rho_2 M_n u(n) + \epsilon(n)$, 其中 $x_i = \frac{i}{n+1}$, $1 \leq i \leq n, X_n = (x_1, x_2, \dots, x_n)'$, $\beta = 3.5, a_n = \max\{1, \log(n)/2\}$, (ρ_1, ρ_2) 分别取为 $(-0.85, -0.15)$ 、 $(-0.85, 0.15)$ 、 $(0.85, -0.15)$ 和 $(0.85, 0.15)$, ϵ_i 's 分别来自 $N(0, 1)$, $N(0, 0.75)$, $t(5)$ 和 $\chi_4^2 - 4$.

空间权重矩阵 $W_n = (w_{ij})$ 中 w_{ij} 的度量采用皇后邻接 (文献 [2] 第 18 页). 首先考虑空间单元的 7 种理想情况: 规则正方形网格 $n = m \times m, m = 3, 4, 5, 6, 7, 10, 13, 16, 20$ 分别表示 W_n 为 $grid_9$ 、 $grid_{16}$ 、 $grid_{25}$ 、 $grid_{36}$ 、 $grid_{49}$ 、 $grid_{100}$ 、 $grid_{169}$ 、 $grid_{256}$ 和 $grid_{400}$. 其次, 考虑一个实例, 美国俄亥俄州哥伦布 49 个相邻规划街区的权重矩阵, 记为 W_{49} , 该矩阵出现在文献 [2] 第 188 页. 最后, 考虑 $W_n = I_5 \otimes W_{49}$, 其中 \otimes 是 kronecker 乘积, 这相当于将五个独立的地区合并在一起, 每个地区都有类似的相邻结构. 模拟中 W_n 将进行标准化使得行和为 1, 即 w_{ij} 由 $w_{ij} / \sum_{j=1}^n w_{ij}$ 替换, 同时取 $M_n = W_n$.

表 2-1 EL 和 AEL 置信域中的覆盖率于 $\epsilon_i \sim N(0, 1)$

(ρ_1, ρ_2)	$W_n = M_n$	EL	AEL	(ρ_1, ρ_2)	$W_n = M_n$	EL	AEL
(-0.85, -0.15)	<i>grid</i> ₉	0.1994	1	(-0.85, 0.15)	<i>grid</i> ₉	0.1988	1
	<i>grid</i> ₁₆	0.5152	0.7800		<i>grid</i> ₁₆	0.5214	0.7942
	<i>grid</i> ₂₅	0.6892	0.7736		<i>grid</i> ₂₅	0.6826	0.7752
	<i>grid</i> ₃₆	0.7776	0.8270		<i>grid</i> ₃₆	0.7760	0.8240
	<i>grid</i> ₄₉	0.8334	0.8648		<i>grid</i> ₄₉	0.8248	0.8672
	<i>grid</i> ₁₀₀	0.9052	0.9184		<i>grid</i> ₁₀₀	0.9004	0.9150
	<i>grid</i> ₁₆₉	0.9252	0.9326		<i>grid</i> ₁₆₉	0.9302	0.9362
	<i>grid</i> ₂₅₆	0.9326	0.9374		<i>grid</i> ₂₅₆	0.9354	0.9414
	<i>grid</i> ₄₀₀	0.9382	0.9416		<i>grid</i> ₄₀₀	0.9378	0.9404
	<i>W</i> ₄₉	0.8286	0.8630		<i>W</i> ₄₉	0.8322	0.8688
	<i>I</i> ₅ \otimes <i>W</i> ₄₉	0.9298	0.9346		<i>I</i> ₅ \otimes <i>W</i> ₄₉	0.9340	0.9390
(0.85, -0.15)	<i>grid</i> ₉	0.1852	1	(0.85, 0.15)	<i>grid</i> ₉	0.1762	1
	<i>grid</i> ₁₆	0.4840	0.7584		<i>grid</i> ₁₆	0.4692	0.7544
	<i>grid</i> ₂₅	0.6676	0.7538		<i>grid</i> ₂₅	0.6438	0.7360
	<i>grid</i> ₃₆	0.7706	0.8232		<i>grid</i> ₃₆	0.7534	0.8046
	<i>grid</i> ₄₉	0.8116	0.8510		<i>grid</i> ₄₉	0.8120	0.8452
	<i>grid</i> ₁₀₀	0.8896	0.9052		<i>grid</i> ₁₀₀	0.8964	0.9116
	<i>grid</i> ₁₆₉	0.9204	0.9300		<i>grid</i> ₁₆₉	0.9176	0.9252
	<i>grid</i> ₂₅₆	0.9300	0.9346		<i>grid</i> ₂₅₆	0.9326	0.9372
	<i>grid</i> ₄₀₀	0.9368	0.9398		<i>grid</i> ₄₀₀	0.9396	0.9426
	<i>W</i> ₄₉	0.8320	0.8614		<i>W</i> ₄₉	0.8182	0.8520
	<i>I</i> ₅ \otimes <i>W</i> ₄₉	0.9304	0.9352		<i>I</i> ₅ \otimes <i>W</i> ₄₉	0.9314	0.9360

表 2-2 EL 和 AEL 置信域中的覆盖率于 $\epsilon_i \sim N(0, 0.75)$

(ρ_1, ρ_2)	$W_n = M_n$	EL	AEL	(ρ_1, ρ_2)	$W_n = M_n$	EL	AEL
(-0.85, -0.15)	<i>grid</i> ₉	0.2026	1	(-0.85, 0.15)	<i>grid</i> ₉	0.1888	1
	<i>grid</i> ₁₆	0.5320	0.7926		<i>grid</i> ₁₆	0.4924	0.7770
	<i>grid</i> ₂₅	0.6804	0.7720		<i>grid</i> ₂₅	0.6754	0.7646
	<i>grid</i> ₃₆	0.7822	0.8330		<i>grid</i> ₃₆	0.7724	0.8266
	<i>grid</i> ₄₉	0.8360	0.8652		<i>grid</i> ₄₉	0.8300	0.8666
	<i>grid</i> ₁₀₀	0.8992	0.9120		<i>grid</i> ₁₀₀	0.9038	0.9200
	<i>grid</i> ₁₆₉	0.9246	0.9332		<i>grid</i> ₁₆₉	0.9216	0.9292
	<i>grid</i> ₂₅₆	0.9362	0.9406		<i>grid</i> ₂₅₆	0.9348	0.9400
	<i>grid</i> ₄₀₀	0.9426	0.9448		<i>grid</i> ₄₀₀	0.9454	0.9476
	<i>W</i> ₄₉	0.8274	0.8600		<i>W</i> ₄₉	0.8336	0.8660
	<i>I</i> ₅ \otimes <i>W</i> ₄₉	0.9302	0.9366		<i>I</i> ₅ \otimes <i>W</i> ₄₉	0.9288	0.9338
(0.85, -0.15)	<i>grid</i> ₉	0.1976	1.0000	(0.85, 0.15)	<i>grid</i> ₉	0.1784	1.0000
	<i>grid</i> ₁₆	0.4828	0.7762		<i>grid</i> ₁₆	0.4674	0.7552
	<i>grid</i> ₂₅	0.6592	0.7512		<i>grid</i> ₂₅	0.6488	0.7482
	<i>grid</i> ₃₆	0.7710	0.8252		<i>grid</i> ₃₆	0.7532	0.8064
	<i>grid</i> ₄₉	0.8174	0.8526		<i>grid</i> ₄₉	0.8138	0.8490
	<i>grid</i> ₁₀₀	0.8978	0.9092		<i>grid</i> ₁₀₀	0.8840	0.8988
	<i>grid</i> ₁₆₉	0.9244	0.9322		<i>grid</i> ₁₆₉	0.9268	0.9326
	<i>grid</i> ₂₅₆	0.9270	0.9344		<i>grid</i> ₂₅₆	0.9290	0.9330
	<i>grid</i> ₄₀₀	0.9440	0.9468		<i>grid</i> ₄₀₀	0.9350	0.9394
	<i>W</i> ₄₉	0.8300	0.8614		<i>W</i> ₄₉	0.8342	0.8634
	<i>I</i> ₅ \otimes <i>W</i> ₄₉	0.9354	0.9410		<i>I</i> ₅ \otimes <i>W</i> ₄₉	0.9274	0.9322

表 2-3 EL 和 AEL 置信域中的覆盖率于 $\epsilon_i \sim t(5)$

(ρ_1, ρ_2)	$W_n = M_n$	EL	AEL	(ρ_1, ρ_2)	$W_n = M_n$	EL	AEL
(-0.85, -0.15)	$grid_9$	0.1508	1	(-0.85, 0.15)	$grid_9$	0.1522	1
	$grid_{16}$	0.4160	0.6848		$grid_{16}$	0.4062	0.6734
	$grid_{25}$	0.5634	0.6610		$grid_{25}$	0.5722	0.6718
	$grid_{36}$	0.6854	0.7350		$grid_{36}$	0.6704	0.7302
	$grid_{49}$	0.7406	0.7782		$grid_{49}$	0.7396	0.7776
	$grid_{100}$	0.8350	0.8544		$grid_{100}$	0.8308	0.8484
	$grid_{169}$	0.8746	0.8866		$grid_{169}$	0.8778	0.8860
	$grid_{256}$	0.9026	0.9086		$grid_{256}$	0.903	0.9098
	$grid_{400}$	0.9112	0.914		$grid_{400}$	0.9126	0.9166
	W_{49}	0.7430	0.7826		W_{49}	0.7328	0.7764
	$I_5 \otimes W_{49}$	0.8910	0.8998		$I_5 \otimes W_{49}$	0.9010	0.9068
(0.85, -0.15)	$grid_9$	0.1392	1	(0.85, 0.15)	$grid_9$	0.1300	1
	$grid_{16}$	0.3656	0.6410		$grid_{16}$	0.3598	0.6336
	$grid_{25}$	0.5476	0.6436		$grid_{25}$	0.5344	0.6324
	$grid_{36}$	0.6486	0.7034		$grid_{36}$	0.6274	0.6868
	$grid_{49}$	0.7176	0.7568		$grid_{49}$	0.7206	0.7580
	$grid_{100}$	0.8254	0.8426		$grid_{100}$	0.8258	0.8452
	$grid_{169}$	0.8620	0.8726		$grid_{169}$	0.8804	0.8878
	$grid_{256}$	0.8922	0.8990		$grid_{256}$	0.8932	0.9028
	$grid_{400}$	0.9114	0.9154		$grid_{400}$	0.9098	0.9138
	W_{49}	0.7306	0.7750		W_{49}	0.7206	0.7590
	$I_5 \otimes W_{49}$	0.8876	0.8940		$I_5 \otimes W_{49}$	0.8928	0.8998

表 2-4 EL 和 AEL 置信域中的覆盖率于 $\epsilon_i + 4 \sim \chi_4^2$

(ρ_1, ρ_2)	$W_n = M_n$	EL	AEL	(ρ_1, ρ_2)	$W_n = M_n$	EL	AEL
(-0.85, -0.15)	$grid_9$	0.1874	1	(-0.85, 0.15)	$grid_9$	0.1766	1
	$grid_{16}$	0.4288	0.7236		$grid_{16}$	0.4222	0.7102
	$grid_{25}$	0.5888	0.6928		$grid_{25}$	0.5708	0.6688
	$grid_{36}$	0.6880	0.7422		$grid_{36}$	0.6792	0.7356
	$grid_{49}$	0.7292	0.7696		$grid_{49}$	0.7310	0.7740
	$grid_{100}$	0.8526	0.8690		$grid_{100}$	0.8408	0.8548
	$grid_{169}$	0.8866	0.8948		$grid_{169}$	0.8894	0.8974
	$grid_{256}$	0.9096	0.9152		$grid_{256}$	0.908	0.9140
	$grid_{400}$	0.9272	0.9314		$grid_{400}$	0.914	0.9168
	W_{49}	0.7274	0.7700		W_{49}	0.7440	0.7796
	$I_5 \otimes W_{49}$	0.9026	0.9088		$I_5 \otimes W_{49}$	0.8992	0.9038
(0.85, -0.15)	$grid_9$	0.1622	1	(0.85, 0.15)	$grid_9$	0.1540	1
	$grid_{16}$	0.4030	0.6758		$grid_{16}$	0.3662	0.6566
	$grid_{25}$	0.5588	0.6468		$grid_{25}$	0.5514	0.6486
	$grid_{36}$	0.6638	0.7226		$grid_{36}$	0.6488	0.7094
	$grid_{49}$	0.7434	0.7822		$grid_{49}$	0.7124	0.7556
	$grid_{100}$	0.8246	0.8420		$grid_{100}$	0.8324	0.8480
	$grid_{169}$	0.8746	0.8848		$grid_{169}$	0.8744	0.8842
	$grid_{256}$	0.9052	0.9110		$grid_{256}$	0.9024	0.9090
	$grid_{400}$	0.9214	0.9252		$grid_{400}$	0.9164	0.9200
	W_{49}	0.7230	0.7644		W_{49}	0.7358	0.7742
	$I_5 \otimes W_{49}$	0.8964	0.9032		$I_5 \otimes W_{49}$	0.9026	0.9086

覆盖率方面，由表2-1至表2-4可知，当误差项 ϵ_i 服从正态分布时，EL 和 AEL 方法在置信域中表现良好，随着空间样本单元数量增大，二者的覆盖率接近名义水平 0.95；当误差项 ϵ_i 服从 t 正态分布、卡方分布时也有不错的表现。同时可看出，空间样本单元数量较大时，AEL 所得的置信域的覆盖率为 EL 所得的置信域的覆盖率很接近；空间样本单元数量较少时，AEL 所得的置信域的覆盖率要远远优于 EL 所得的置信域的覆盖率。总之，无论误差项服从什么分布，无论是大样本还是小样本，AEL 置信域要宽于 EL 的置信域（文献 [24] 第 435 页），AEL 所得的置信域的覆盖率要优于 EL。

表 2-5 计算 EL 和 AEL 覆盖率所用时长对比于 $\epsilon_i \sim N(0, 1)$

$(\rho_1, \rho_2) = (0.85, 0.15)$ $W_n = M_n$	模拟 100 次用时 (s)		模拟 1000 次用时 (s)		模拟 5000 次用时 (min)	
	EL	AEL	EL	AEL	EL	AEL
<i>grid</i> ₉	3.034911	1.501408	30.98868	14.89316	2.604628	1.264644
<i>grid</i> ₁₆	2.157349	1.701035	22.03516	16.80333	1.842554	1.439872
<i>grid</i> ₂₅	1.939484	1.699198	19.49585	17.88881	1.761386	1.617301
<i>grid</i> ₃₆	1.920298	2.158245	19.43822	18.54867	1.715465	1.664107
<i>grid</i> ₄₉	2.032650	1.975585	19.32017	19.46934	1.629343	1.589021
<i>grid</i> ₁₀₀	2.406406	2.364031	20.84890	20.38607	1.742501	1.743671
<i>grid</i> ₁₆₉	2.192560	2.246201	20.11281	20.63526	1.800101	1.740035
<i>grid</i> ₂₅₆	2.408131	2.271219	22.96711	22.07802	1.953248	1.965933
<i>grid</i> ₄₀₀	2.569816	2.574532	27.46002	27.25366	2.322573	2.249785
总时间	20.6616	18.49145	202.6669	177.9563	17.3718	15.27437

计算效率方面，EL 的估计函数集的凸包内部有可能不包含 0，EL 方程 (1-10) 不一定有解，则经验似然比统计量 $L_n(\theta)$ 不一定存在，AEL 的估计函数集的凸包内部必包含 0，AEL 方程 (3-2) 一定有解，则调整经验似然比统计量 $L_n^*(\theta)$ 一定存在。我们推荐使用修正的 Newton-Raphson 算法 [2] 求方程的解，该算法的收敛性证明见文献 [34]，表2-6 给出两种经验似然方法的运行总时间，可以看出模拟 100 次时，AEL 比 EL 快 2 秒左右，模拟 1000 次时，AEL 比 EL 快 25 秒左右，模拟 5000 次时，AEL 比 EL 快 120 秒左右。小样本情形，两者运行时间的差异较大，可能的原因是，在求解 EL 方程的解时，如果初始点远离真实解，其收敛性不能保证，直到迭代次数（设置迭代 300 次）用完被迫退出循环，而 AEL 方程的解很快收敛，迭代次数还没用完就已退出循环。该模拟在运行环境为：处理器为 1.4GHz, Intel Core i5，内存 8 GB 上运行得出。

表 2-6 模拟中不同假设下的 SARAR 模型

Model	β	ρ_1	ρ_2	σ^2
H_0	3.5	0.85	0.15	1
A_1	3.5	0.85	0.15	1.001
A_2	3.5	0.85	0.15	1.1
A_3	3.5	0.85	0.15	2
A_4	3.5	0.85	0.15	3

检验功效方面，考虑假设检验问题，原假设 $H_0 : \theta = \theta_0$ 对备择假设 $H_1 : \theta \neq \theta_0$ ，其

中备择假设检验 H_1 考虑表2-6 中 A_1 、 A_2 、 A_3 和 A_4 四种模型设定. A_1 、 A_2 、 A_3 和 A_4 分别代表了非真值参数与真值参数从近到远的距离, 即: $\|\theta - \theta_0\| < n^{-1}$ 、 $\|\theta - \theta_0\| < n^{-1/3}$ 、 $\|\theta - \theta_0\| > n^{-1/12}$ 和 $\|\theta - \theta_0\| > n^{1/9}$.

由表2-7 可见四种对立假设被拒绝的比例情况, 随着参数与真值参数的距离逐渐变大, EL 和 AEL 被拒绝的比例都逐渐增大; 当参数与真值参数距离很大时, EL 和 AEL 被拒绝的比例都很接近 1, 甚至等于 1.

表 2-7 EL 和 AEL 拒绝域中的拒绝比例于 $\epsilon_i \sim N(0, \sigma^2)$

$W_n = M_n$	EL					AEL				
	H_0	A_1	A_2	A_3	A_4	H_0	A_1	A_2	A_3	A_4
$grid_9$	0.8308	0.8306	0.8504	0.9634	0.9928	0.0000	0.0000	0.0000	0.0000	0.0000
$grid_{16}$	0.5308	0.5312	0.5730	0.9134	0.9912	0.2478	0.2484	0.3034	0.7666	0.9562
$grid_{25}$	0.3422	0.3428	0.3970	0.8840	0.9940	0.2452	0.2458	0.3082	0.8382	0.9908
$grid_{36}$	0.2530	0.2534	0.3132	0.8980	0.9976	0.2054	0.2052	0.2586	0.8690	0.9962
$grid_{49}$	0.1894	0.1890	0.2380	0.9316	0.9994	0.1584	0.1592	0.2028	0.9172	0.9988
$grid_{100}$	0.1090	0.1094	0.1680	0.9938	1	0.0958	0.0968	0.1510	0.9920	1
$grid_{169}$	0.0764	0.0770	0.1374	1	1	0.0678	0.0682	0.1302	1	1
$grid_{400}$	0.0622	0.0620	0.1858	1	1	0.0602	0.0600	0.1776	1	1

表 2-8 拒绝域中 EL 和 AEL 的拒绝比例对比于 $\epsilon_i \sim N(0, \sigma^2)$

$W_n = M_n$	H_0		A_1		A_2		A_3		A_4	
	EL	AEL	EL	AEL	EL	AEL	EL	AEL	EL	AEL
$grid_9$	0.8308	0.0000	0.8306	0.0000	0.8504	0.0000	0.9634	0.0000	0.9928	0.0000
$grid_{16}$	0.5308	0.2478	0.5312	0.2484	0.5730	0.3034	0.9134	0.7666	0.9912	0.9562
$grid_{25}$	0.3422	0.2452	0.3428	0.2458	0.3970	0.3082	0.8840	0.8382	0.9940	0.9908
$grid_{36}$	0.2530	0.2054	0.2534	0.2052	0.3132	0.2586	0.8980	0.8690	0.9976	0.9962
$grid_{49}$	0.1894	0.1584	0.1890	0.1592	0.2380	0.2028	0.9316	0.9172	0.9994	0.9988
$grid_{100}$	0.1090	0.0958	0.1094	0.0968	0.1680	0.1510	0.9938	0.9920	1	1
$grid_{169}$	0.0764	0.0678	0.0770	0.0682	0.1374	0.1302	1	1	1	1
$grid_{400}$	0.0622	0.0602	0.0620	0.0600	0.1858	0.1776	1	1	1	1

由表2-8 进行 EL 与 AEL 的两两对比. 原假设 H_0 下, AEL 被拒绝比例比 EL 更接近 α 为 0.05 的名义水平, AEL 被拒绝比例优于 EL; 当样本较大且 $\|\theta - \theta_0\| < n^{-1/3}$ 时, 两种方法被拒绝的比例相近, 但 AEL 被拒绝比例更优于 EL 被拒绝比例; 当样本较大且 $\|\theta - \theta_0\| > n^{-1/3}$ 时, 两种方法下被拒绝的比例都趋近于 1; 当样本数量很少时, 比如 9 个样本, EL 在原假设下被拒绝比例很高, AEL 在对立假设下被接受比例很高. 总之, AEL 犯第一类错误的比例优于 EL, AEL 犯第二类错误的比例略劣于 EL.

总结, 调整经验似然比一般经验似然置信域的覆盖精度更高, 计算速度更快, 检验功效略小. 大样本时, 无论误差项是否服从正态分布, 推荐 EL 和 AEL 方法 (文献 [20]); 小样本时, 模拟结果推荐 AEL 方法.

§2.4 引理及定理的证明

引理 2.1 设条件 (A1)-(A3) 满足, 则 $\theta = \theta_0$, $n \rightarrow \infty$ 时, 有

$$\begin{aligned} \max_{1 \leq i \leq n} \|\omega_i(\theta)\| &= o_p(n^{1/2}) \text{ a.s.}, \\ \Sigma_{p+3}^{-1/2} \sum_{i=1}^n \omega_i(\theta) &\xrightarrow{d} N(0, I_{p+3}), \\ n^{-1} \sum_{i=1}^n \omega_i(\theta) \omega_i'(\theta) &= n^{-1} \Sigma_{p+3} + o_p(1), \\ \sum_{i=1}^n \|\omega_i(\theta)\|^3 &= O_p(n), \end{aligned}$$

其中 Σ_{p+3} 已在 (1-11) 中给出.

引理 2.1 的证明: 见文献 [39] 的引理 3.

定理 2.3 的证明: 令记 $\lambda = \lambda(\theta)$, λ 是满足方程 (3-2) 的解, 只要 $a_n = o_p(n)$, 令 $\omega^* = \max_{1 \leq i \leq n} \|\omega_i(\theta)\|$, 由引理 2.1 可知,

$$\bar{\omega}_n = O_p(n^{-1/2}), \quad \omega^* = o_p(n^{1/2})$$

令 $\omega_i = \omega_i(\theta)$, $\rho = \|\lambda\|$, $\hat{\lambda} = \lambda/\rho$, 有 $\|\hat{\lambda}\| = 1$, $\lambda = \rho\hat{\lambda}$, $\lambda' = \rho\hat{\lambda}'$.

于是, 由 (3-2) 可得,

$$\begin{aligned} 0 &= \frac{\hat{\lambda}'}{n} \sum_{i=1}^{n+1} \frac{\omega_i}{1 + \lambda' \omega_i} \\ &= \frac{\hat{\lambda}'}{n} \sum_{i=1}^{n+1} \omega_i - \frac{\rho}{n} \sum_{i=1}^{n+1} \frac{(\hat{\lambda}' \omega_i)^2}{1 + \rho \hat{\lambda}' \omega_i} \\ &= \frac{\hat{\lambda}'}{n} \left(\sum_{i=1}^n \omega_i + \omega_{n+1} \right) - \frac{\rho}{n} \left(\sum_{i=1}^n \frac{(\hat{\lambda}' \omega_i)^2}{1 + \rho \hat{\lambda}' \omega_i} + \frac{(\hat{\lambda}' \omega_{n+1})^2}{1 + \rho \hat{\lambda}' \omega_{n+1}} \right) \\ &= \hat{\lambda}' \bar{\omega}_n (1 - a_n/n) - \frac{\rho}{n} \left(\sum_{i=1}^n \frac{(\hat{\lambda}' \omega_i)^2}{1 + \rho \hat{\lambda}' \omega_i} + \frac{(\hat{\lambda}' \omega_{n+1})^2}{1 + \rho \hat{\lambda}' \omega_{n+1}} \right) \\ &\leq \hat{\lambda}' \bar{\omega}_n (1 - a_n/n) - \frac{\rho}{n} \sum_{i=1}^n \frac{(\hat{\lambda}' \omega_i)^2}{1 + \rho \hat{\lambda}' \omega_i} \\ &\leq \hat{\lambda}' \bar{\omega}_n (1 - a_n/n) - \frac{\rho}{n} \sum_{i=1}^n \frac{(\hat{\lambda}' \omega_i)^2}{1 + \rho \omega^*} \\ &= \hat{\lambda}' \bar{\omega}_n - \frac{\rho}{n(1 + \rho \omega^*)} \sum_{i=1}^n (\hat{\lambda}' \omega_i)^2 + O_p(n^{-3/2} a_n) \\ &= \hat{\lambda}' \bar{\omega}_n - \frac{\rho}{n(1 + \rho \omega^*)} \sum_{i=1}^n (\hat{\lambda}' \omega_i)^2 + o_p(n^{-1/2}). \end{aligned}$$

从而，

$$\frac{\rho}{1 + \rho\omega^*} \cdot \frac{1}{n} \sum_{i=1}^n (\hat{\lambda}'\omega_i)^2 = \hat{\lambda}'\bar{\omega}_n + o_p(n^{-1/2}).$$

令 $V_0 = n^{-1} \sum_{i=1}^n \omega_i(\theta)\omega_i'(\theta)$ ，结合条件 A3 可知，

$$\frac{1}{n} \sum_{i=1}^n (\hat{\lambda}'\omega_i)^2 = \hat{\lambda}' \cdot \frac{1}{n} \sum_{i=1}^n \omega_i\omega_i' \cdot \hat{\lambda} = \hat{\lambda}'V_0\hat{\lambda} \geq \lambda_{\min}(V_0),$$

$$|\hat{\lambda}'\bar{\omega}_n| \leq |\hat{\lambda}'\Sigma_{p+3}^{1/2}\Sigma_{p+3}^{-1/2}\bar{\omega}_n| \leq \|\Sigma_{p+3}^{1/2}\hat{\lambda}\| \cdot \|\Sigma_{p+3}^{-1/2}\bar{\omega}_n\| \leq \lambda_{\max}(\Sigma_{p+3}^{1/2})\|\hat{\lambda}\| \cdot \|\Sigma_{p+3}^{-1/2}\bar{\omega}_n\|,$$

即，

$$\frac{\rho}{1 + \rho\omega^*} \lambda_{\min}(V_0) \leq \lambda_{\max}(\Sigma_{p+3}^{1/2})\|\hat{\lambda}\| \cdot \|\Sigma_{p+3}^{-1/2}\bar{\omega}_n\| + o_p(n^{-1/2}).$$

由引理2.1可知 $\|\Sigma_{p+3}^{-1/2}\bar{\omega}_n\| = O_p(n^{-1})$ ，结合条件 A3 得，

$$\frac{\rho}{1 + \rho\omega^*} = O_p(n^{-1/2}),$$

再由引理2.1 可知，

$$\rho = O_p(n^{-1/2}),$$

即，

$$\lambda = O_p(n^{-1/2}). \quad (2-4)$$

令 $\gamma_i = \lambda'\omega_i$ ，由柯西-施瓦茨不等式可知， $\max_{1 \leq i \leq n} |\gamma_i| \leq \|\lambda\| \cdot \omega^*$ ， $|\gamma_{n+1}| = |\lambda'\omega_{n+1}| = |a_n\lambda'\bar{\omega}_n| \leq a_n\|\lambda\| \cdot \|\bar{\omega}_n\|$ 由引理2.1 可得，

$$\max_{1 \leq i \leq (n+1)} |\gamma_i| = o_p(1), \quad (2-5)$$

由 (3-2) 可得，

$$\begin{aligned} 0 &= \frac{1}{n} \sum_{i=1}^{n+1} \frac{\omega_i}{1 + \lambda'\omega_i} \\ &= \frac{1}{n} \sum_{i=1}^{n+1} \omega_i - \frac{1}{n} \sum_{i=1}^{n+1} \frac{\omega_i \lambda'\omega_i}{1 + \lambda'\omega_i} \\ &= \frac{1}{n} \sum_{i=1}^{n+1} \omega_i - \left\{ \frac{1}{n} \sum_{i=1}^{n+1} \omega_i \omega_i' \right\} \lambda + \frac{1}{n} \sum_{i=1}^{n+1} \frac{\omega_i (\lambda'\omega_i)^2}{1 + \lambda'\omega_i} \\ &= \bar{\omega}_n + \frac{1}{n} \omega_{n+1} - V_0 \lambda - \frac{1}{n} \omega_{n+1} \gamma_{n+1} + \frac{1}{n} \sum_{i=1}^n \frac{\omega_i \gamma_i^2}{1 + \lambda'\omega_i} + \frac{1}{n} \cdot \frac{\omega_{n+1} \gamma_{n+1}^2}{1 + \lambda'\omega_{n+1}} \end{aligned}$$

结合引理2.1 和条件 A3, 上式可得,

$$\lambda = V_0^{-1}\bar{\omega}_n + o_p(n^{-1/2}), \quad (2-6)$$

其中, $\|n^{-1} \sum_{i=1}^n \omega_i \gamma_i^2\| \leq n^{-1} \sum_{i=1}^n \|\omega_i\|^3 \cdot \|\lambda\|^2 = O_p(n^{-1}) = o_p(n^{-1/2})$.

由 (2-5) 可以泰勒展开 $\log(1 + \gamma_i) = \gamma_i - \gamma_i^2/2 + \nu_i$, 其中, 存在 $B > 0$, 使得当 $n \rightarrow \infty$ 时,

$$P(|\nu_i| \leq B|\gamma_i|^3, 1 \leq i \leq n+1) \rightarrow 1$$

因此, 由 (3-1), (2-6) 和泰勒展开, 可以得到

$$\begin{aligned} \ell_n^*(\theta_0) &= 2 \sum_{j=1}^{n+1} \log(1 + \gamma_j) = 2 \sum_{j=1}^{n+1} \gamma_j - \sum_{j=1}^{n+1} \gamma_j^2 + 2 \sum_{j=1}^{n+1} \nu_j \\ &= 2 \sum_{j=1}^n \gamma_j - \sum_{j=1}^n \gamma_j^2 + 2\gamma_{n+1} - \gamma_{n+1}^2 + 2 \sum_{j=1}^{n+1} \nu_j \\ &= 2n\lambda'\bar{\omega}_n - n\lambda'V_0\lambda + 2 \sum_{j=1}^{n+1} \nu_j + o_p(1) \\ &= 2n(V_0^{-1}\bar{\omega}_n)'\bar{\omega}_n + 2n \cdot o_p(n^{-1/2})\bar{\omega}_n - n\bar{\omega}_n'V_0^{-1}\bar{\omega}_n \\ &\quad - 2n \cdot o_p(n^{-1/2})\bar{\omega}_n - n \cdot o_p(n^{-1/2})V_0o_p(n^{-1/2}) + 2 \sum_{j=1}^{n+1} \nu_j + o_p(1) \\ &= n\bar{\omega}_n'V_0^{-1}\bar{\omega}_n + 2 \sum_{j=1}^{n+1} \nu_j + o_p(1) \\ &= \{n\Sigma_{p+3}^{-1/2}\bar{\omega}_n\}'\{n\Sigma_{p+3}^{-1/2}V_0\Sigma_{p+3}^{-1/2}\}^{-1}\{n\Sigma_{p+3}^{-1/2}\bar{\omega}_n\} + 2 \sum_{j=1}^{n+1} \nu_j + o_p(1). \end{aligned}$$

结合引理2.1 和条件 A3, 可知,

$$\{n\Sigma_{p+3}^{-1/2}\bar{\omega}_n\}'\{n\Sigma_{p+3}^{-1/2}V_0\Sigma_{p+3}^{-1/2}\}^{-1}\{n\Sigma_{p+3}^{-1/2}\bar{\omega}_n\} \xrightarrow{d} \chi_{p+3}^2.$$

其中,

$$\left| \sum_{j=1}^{n+1} \nu_j \right| \leq \left| \sum_{j=1}^n \nu_j \right| + |\nu_{n+1}| \leq B\|\lambda\|^3 \sum_{j=1}^n \|\omega_j\|^3 + B\|\lambda\|^3 \|\omega_{n+1}\|^3 = O_p(n^{-1/2}) + o_p(1) = o_p(1).$$

定理2.3 证毕.

定理2.4的证明: 令 $\bar{\omega}_n = \bar{\omega}_n(\theta)$, 当 $n \rightarrow \infty$ 时, 由大数定律可知, $\|\bar{\omega}_n'\bar{\omega}_n\| \xrightarrow{p} \delta^2 > 0$, 类似文献 [39] 中 (4.2) 的证明可得 $\omega^* \triangleq \max_{1 \leq i \leq n} \|\omega_i(\theta)\| = o_p(n^{1/2})$. 令 $\tilde{\lambda} = n^{-2/3}\bar{\omega}_n M$, M 为正数, $\tilde{\gamma}_i = \tilde{\lambda}'\omega_i(\theta)$, 有 $\max_{1 \leq i \leq n} |\tilde{\gamma}_i| \leq \|\tilde{\lambda}\| \cdot \omega^* = o_p(1)$, $|\tilde{\gamma}_{n+1}| = |\tilde{\lambda}'\omega_{n+1}(\theta)| = o_p(1)$, 得

$$\max_{1 \leq i \leq (n+1)} |\tilde{\gamma}_i| = o_p(1), \quad (2-7)$$

由 (2-7) 可以泰勒展开 $\log(1 + \tilde{\gamma}_i) = \tilde{\gamma}_i - \tilde{\gamma}_i^2/2 + \tilde{\nu}_i$, 其中, 存在 $B > 0$, 使得当 $n \rightarrow \infty$ 时,

$$P(|\tilde{\nu}_i| \leq B|\tilde{\gamma}_i|^3, 1 \leq i \leq n+1) \rightarrow 1$$

因此, 由 (2-7) 和泰勒展开, 可以得到

$$\begin{aligned} \sum_{j=1}^{n+1} \log(1 + \tilde{\gamma}_j) &= \sum_{j=1}^{n+1} \tilde{\gamma}_j - \frac{1}{2} \sum_{j=1}^{n+1} \tilde{\gamma}_j^2 + \sum_{j=1}^{n+1} \tilde{\nu}_j \\ &= \sum_{j=1}^n \tilde{\gamma}_j - \frac{1}{2} \sum_{j=1}^n \tilde{\gamma}_j^2 + \sum_{j=1}^{n+1} \tilde{\nu}_j + \tilde{\gamma}_{n+1} - \frac{1}{2} \tilde{\gamma}_{n+1}^2 \\ &= n^{1/3} \delta^2 M - \frac{1}{2} \sum_{j=1}^n \tilde{\gamma}_j^2 + \sum_{j=1}^{n+1} \tilde{\nu}_j + o_p(1). \end{aligned}$$

其中,

$$\left| \sum_{j=1}^n \tilde{\gamma}_j^2 \right| \leq |n \tilde{\lambda}' V_0 \tilde{\lambda}| \leq |n \tilde{\lambda}' \lambda_{\max}(V_0) \tilde{\lambda}| = O_p(n^{-4/3}) = o_p(1),$$

$$\left| \sum_{j=1}^{n+1} \tilde{\nu}_j \right| \leq B \|\tilde{\lambda}\|^3 \sum_{j=1}^n \|\omega_j(\theta)\|^3 + B \|\tilde{\lambda}\|^3 \|\omega_{n+1}(\theta)\|^3 = O_p(n^{-5/2}) + O_p(n^{-2}) = o_p(1),$$

当 $\theta \neq \theta_0$ 时, $\sum_{j=1}^n \|\omega_j(\theta)\|^3 = O_p(n)$ 的证明与文献 [39] 中 (4.5) 的证明类似. 由 (2-1) 和最大化问题的对偶性, 可得

$$\begin{aligned} W^*(\theta) &\triangleq \log L_n^*(\theta) = \sup_{p_i, 1 \leq i \leq n+1} \left\{ \sum_{i=1}^{n+1} \log(n+1)p_i \right\} \\ &= - \sup_{\lambda} \left\{ \sum_{i=1}^{n+1} \log(1 + \lambda' \omega_i) \right\} \\ &\leq - \sum_{i=1}^{n+1} \log(1 + \tilde{\lambda}' \omega_i) \\ &= -n^{1/3} \delta^2 M + o_p(1) \end{aligned}$$

因为 M 是个任意大的数, 于是, $-2n^{-1/3} W^*(\theta) \rightarrow \infty$ 即 $n^{-1/3} \ell_n^*(\theta) \xrightarrow{p} \infty$. 类似可证, $n^{-1/3} \ell_n(\theta) \xrightarrow{p} \infty$.

定理2.4 证毕.

第三章 含空间自回归误差的空间自回归模型的高维经验似然推断

§3.1 高维经验似然原理

我们借鉴分块经验似然思想. 令 $s = [r/l]$, 其中 $l \leq r$, l 是一个正数以及 $[a]$ 表示对 a 取整. 令

$$\tilde{g}_i(X, \theta) = \frac{1}{s} \sum_{j=(i-1)s+1}^{is} g_j(X, \theta), 1 \leq i \leq l,$$

并且当 $r > sl$ 时,

$$\tilde{g}_{l+1}(X, \theta) = \frac{1}{r - sl} \sum_{j=sl+1}^r g_j(X, \theta),$$

不然, $\tilde{g}_{l+1}(X, \theta)$ 为 0.

令 $l_0 = l + 1$, 当 $r > sl$ 时, 否则 $l_0 = l$, 得到分块得分函数定义:

$$\tilde{g}(x, \theta) = (\tilde{g}_1(x, \theta), \tilde{g}_2(x, \theta), \dots, \tilde{g}_{l_0}(x, \theta))^T, x \in R^d, \theta \in \Theta,$$

以及该分块经验似然统计量为:

$$\tilde{\ell}_E(\theta) = \sum_{i=1}^n \log\{1 + \tilde{t}^T(\theta)\tilde{g}(X_i, \theta)\},$$

其中 $\tilde{t}(\theta)$ 是下列方程的解:

$$\sum_{i=1}^n \frac{\tilde{g}(X_i, \theta)}{1 + \tilde{t}^T(\theta)\tilde{g}(X_i, \theta)} = 0.$$

事实上, 令 $\mathbf{1}_s$ 表示由数字 1 为元素组成的 s 维向量并且当 $r = sl$ 时定义

$$H_{l \times r} = \begin{pmatrix} s^{-1}\mathbf{1}_s^T & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & s^{-1}\mathbf{1}_s^T & \mathbf{0} & \dots & \mathbf{0} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & s^{-1}\mathbf{1}_s^T \end{pmatrix},$$

以及当 $r > sl$ 时可定义类似矩阵 H . 于是, 可得

$$\tilde{g}(x, \theta) = Hg(x, \theta).$$

基于定理2.1很容易获得分块经验似然比统计量 $\tilde{\ell}_E(\theta)$ 的渐近分布, 见定理3.1.

定理 3.1 假设定理2.1的条件都满足, 则对于固定的 p 和 r , 当 $n \rightarrow \infty$, 有

$$2\tilde{\ell}_E(\theta_0) \xrightarrow{d} \chi_{l_0}^2,$$

其中 $\chi_{l_0}^2$ 表示自由度为 l_0 的卡方分布.

§3.2 高维经验似然比统计量的极限分布

根据分块似然思想得到关于 θ 的高维经验似然比统计量:

$$\tilde{L}_n(\theta) = \sup \left\{ \prod_{i=1}^n (np_i) : p_i \geq 0, \sum_{i=1}^n p_i = 1, \sum_{i=1}^n p_i \tilde{\omega}_i(\theta) = 0 \right\},$$

其中 $\tilde{\omega}_i(\theta) = \omega_i^T(\theta) \mathbf{1}_{p+3}$. 根据经验似然框架, 可得:

$$\tilde{\ell}_n(\theta) \triangleq -2 \log \tilde{L}_n(\theta) = 2 \sum_{i=1}^n \log \{1 + \lambda(\theta) \tilde{\omega}_i(\theta)\}, \quad (3-1)$$

其中 $\lambda(\theta) \in R^1$ 为以下方程的解:

$$\frac{1}{n} \sum_{i=1}^n \frac{\tilde{\omega}_i(\theta)}{1 + \lambda(\theta) \tilde{\omega}_i(\theta)} = 0. \quad (3-2)$$

高维经验似然的主要结论如下:

定理 3.2 在假设条件 (A1)-(A3) 及模型 (1-3) 下, 当 $\theta = \theta_0, n \rightarrow \infty$, 且 $p \rightarrow \infty$,

$$\tilde{\ell}_n(\theta_0) \xrightarrow{d} \chi_1^2,$$

其中 χ_1^2 表示自由度为 1 的卡方分布.

取定 $\alpha, 0 < \alpha < 1$, 设 $z_\alpha(1)$ 满足 $P(\chi_1^2 \geq z_\alpha(1)) = \alpha$, 由定理 3.2 可确定基于高维经验似然方法下 θ_0 的置信水平为 $1 - \alpha$ 的渐近置信域为:

$$\{\theta : \tilde{\ell}_n(\theta_0) \leq z_\alpha(1)\}.$$

§3.3 模拟结果

我们通过模拟比较 EL 和 HEL 优劣, 给定置信水平 $1 - \alpha = 0.95$, 由式 (1-12) 和 (2-4) 分别得到一般经验似然以及高经验似然方法关于 θ 的置信域, 我们针对假设检验问题 $H_0 : \theta = \theta_0, H_1 : \theta \neq \theta_0$, 设计实验.

在模拟中, 我们使用如下模型 $Y_n = \rho_1 W_n Y_n + X_n \beta_0 + u_{(n)}, u_{(n)} = \rho_2 M_n u_{(n)} + \epsilon_{(n)}$, 其中 $(\rho_1, \rho_2) = (0.85, 0.15)$, $\beta_0 = \mathbf{1}_{p+3}$, $M_n = W_n$, 以及 $\{X_i\} \sim N(\mathbf{0}, \Sigma_P), i = 1, 2, \dots, n$ 且 $\Sigma_P = I_{\{i=j, 1 \leq i, j \leq p\}}$. 为了 p 与 n 同时增长, 我们考虑 $p = [3n^{index}]$, $index = 0, 0.1, 0.2, 0.3, 0.4, 0.5$, 其中 $[x]$ 是对 x 取整. 空间权重矩阵 $W_n = (w_{ij})$, 我们采用皇后邻接方式并考虑空间单元的 5 种理想情况: 规则正方形网格 $n = m \times m, m = 10, 15, 18, 20, 25, 30$.

实验 1

记 $\theta'_0 = (\beta'_0, \rho_1, \rho_2, \sigma_0^2)$ 为 θ 的真实值, 误差项 $\epsilon'_i s$ 依次服从 $N(0, 1)$, $N(0, 0.75)$, $t(5)$ 以及 $\chi^2_4 - 4$, 则 σ_0^2 依次为 1, 0.75, 2.5 和 8. 我们给出 $\tilde{\ell}_n(\theta_0) \leq z_{0.05}(1)$ 和 $\ell_n(\theta_0) \leq z_{0.05}(p+3)$ 在 2000 次模拟在接受域中出现的比例, 模拟结果见表 3-1.

实验 2

记 $\theta' = (\beta', \rho_1, \rho_2, \sigma^2)$, 其中 $(\beta, \sigma^2)' = (\beta_0, \sigma_0^2)' + \Delta \mathbf{1}'_{p+1}$, Δ 依次取 0, 0.1, 1, 2. 令 $\epsilon_i \sim N(0, 1)$, 我们给出 $\tilde{\ell}_n(\theta) \geq z_{0.05}(1)$ 和 $\ell_n(\theta) \geq z_{0.05}(p+3)$ 在 2000 次模拟在拒绝域中出现的比例, 模拟结果见表 3-2.

实验分析

一类错误方面, 我们设计了真值情形下置信域的覆盖率, 由表3-1可知, 当 $index \leq 0.2$ 时, 无论误差服从哪种分布, EL 和 HEL 方法在置信域中表现良好, 随着空间样本单元数量增大, 二者的覆盖率接近 0.95; 当 $index > 0.2$ 时, EL 在 $N(0, 1)$ 、 $N(0, 0.75)$ 、 $t(5)$ 以及 χ^2 分布下落在接受域中的范围依次为 [0.022, 0.905]、[0.020, 0.909]、[0.011, 0.840] 以及 [0.010, 0.849], 这些值都远远小于 0.95, 与此同时, HEL 在 $N(0, 1)$ 、 $N(0, 0.75)$ 、 $t(5)$ 以及 χ^2 分布下落在接受域中的范围依次为 [0.938, 0.952]、[0.937, 0.956]、[0.928, 0.948] 以及 [0.919, 0.952], 这些值非常接近 0.95. 也就是说, 在高维下, EL 犯第一类错误的几率非常大, HEL 犯第一类错误的几率接近名义水平.

二类错误方面, 我们设计了非真值情形下拒绝域的覆盖率, 对于 $H_0 : \theta = \theta_0$ 和 $H_1 : \theta \neq \theta_0$ 的假设检验问题, 我们考虑四种模型设定, $\Delta = 0, 0.1, 1, 2$ 代表了非真值参数与真值参数从近到远的不同距离, 其中 $\Delta = 0$ 代表了 $\|\theta - \theta_0\| = 0$. 由表3-2可知, 四种参数情形被拒绝的比例情况, 随着参数与真值参数的距离逐渐变大, EL 和 HEL 被拒绝的比例都逐渐增大; 当参数与真值参数距离很大时, EL 和 HEL 被拒绝的比例都很接近 1, 甚至等于 1. 当 $index \geq 0.3$ 时, 即使在 $\Delta = 0.1$, θ_0 与 θ 比较近时, EL 都会大概率判入拒绝域, 而 HEL 是随着参数与真值之间距离程度判断是否落入拒绝域, 距离越远拒绝越彻底, 距离比较近拒绝没有那么彻底, 可见在犯第二类错误方面, HEL 比 EL 更优.

为了说明高维下, 高维经验似然比统计量与理论相符合, 我们作了 EL 和 HEL 的分位数图, 参数与实验一相同, 误差服从标准正态. 图3-1为 $\ell_n(\theta_0)$ 的样本分位数与其理论分布 $\chi^2(p+3)$ 分位数图, 图3-2为 $\tilde{\ell}_n(\theta_0)$ 与其理论分布 $\chi^2(1)$ 分位数图. 由图3-1和图3-2直观地可以看到, 在低维情形下 ($index \leq 0.2$), 随着样本容量的增大, $\ell_n(\theta_0)$ 的样本分位数与其理论分布 $\chi^2(p+3)$ 分位数拟合越好, 在高维情形下 ($index \geq 0.4$), $\ell_n(\theta_0)$ 的样本分位数与其理论分布 $\chi^2(p+3)$ 分位数拟合很差, 而不论在高维或低维情形下, 随着样本容量的增大, $\tilde{\ell}_n(\theta_0)$ 与其理论分布 $\chi^2(1)$ 分位数都拟合得很好.

总结, 在高维下, HEL 比 EL 犯一类错误更少, 检验功效更优, 统计量样本分位数与理论分位数拟合更好, 因此, 我们推荐 HEL 方法.

表 3-1 Coverage probabilities of the EL and HEL confidence regions

n	p	$\epsilon_i \sim N(0, 1)$		$\epsilon_i \sim N(0, 0.75)$		$\epsilon_i \sim t(5)$		$\epsilon_i \sim \chi^2(4) - 4$	
		EL	HEL	EL	HEL	EL	HEL	EL	HEL
100	3	0.861	0.948	0.872	0.940	0.778	0.904	0.775	0.910
	5	0.824	0.944	0.824	0.941	0.720	0.931	0.730	0.920
	8	0.740	0.950	0.718	0.957	0.591	0.918	0.624	0.919
	12	0.548	0.943	0.545	0.937	0.392	0.928	0.452	0.919
	19	0.250	0.940	0.266	0.938	0.132	0.930	0.182	0.920
	30	0.022	0.941	0.020	0.940	0.011	0.928	0.010	0.929
225	3	0.920	0.946	0.917	0.948	0.866	0.919	0.877	0.921
	5	0.910	0.945	0.907	0.952	0.857	0.928	0.861	0.927
	9	0.863	0.941	0.878	0.949	0.786	0.926	0.788	0.921
	15	0.774	0.949	0.783	0.953	0.614	0.942	0.665	0.936
	26	0.508	0.944	0.508	0.944	0.317	0.938	0.384	0.932
	45	0.096	0.950	0.084	0.954	0.025	0.941	0.047	0.943
324	3	0.929	0.953	0.936	0.948	0.898	0.933	0.904	0.943
	5	0.918	0.947	0.925	0.950	0.871	0.930	0.883	0.936
	10	0.889	0.947	0.902	0.951	0.814	0.931	0.827	0.928
	17	0.832	0.947	0.827	0.953	0.701	0.943	0.716	0.943
	30	0.599	0.950	0.605	0.947	0.412	0.942	0.450	0.936
	54	0.141	0.952	0.136	0.954	0.046	0.944	0.075	0.953
400	3	0.931	0.950	0.929	0.950	0.895	0.929	0.907	0.941
	5	0.929	0.949	0.925	0.936	0.891	0.937	0.897	0.940
	10	0.905	0.952	0.911	0.956	0.856	0.936	0.850	0.945
	18	0.856	0.942	0.850	0.949	0.723	0.945	0.740	0.941
	33	0.655	0.938	0.655	0.952	0.459	0.948	0.476	0.950
	60	0.163	0.944	0.186	0.946	0.061	0.943	0.088	0.939
625	3	0.948	0.948	0.938	0.941	0.922	0.942	0.935	0.945
	6	0.935	0.945	0.924	0.951	0.906	0.943	0.918	0.942
	11	0.931	0.952	0.926	0.953	0.879	0.938	0.885	0.941
	21	0.889	0.952	0.890	0.942	0.811	0.945	0.812	0.938
	39	0.739	0.946	0.751	0.947	0.557	0.949	0.590	0.938
	75	0.262	0.948	0.267	0.944	0.085	0.945	0.130	0.944
900	3	0.951	0.948	0.935	0.948	0.921	0.936	0.931	0.945
	6	0.942	0.946	0.941	0.955	0.915	0.946	0.922	0.949
	12	0.928	0.945	0.936	0.940	0.906	0.945	0.902	0.946
	23	0.905	0.948	0.909	0.948	0.840	0.937	0.849	0.942
	46	0.792	0.952	0.798	0.956	0.624	0.945	0.625	0.952
	90	0.353	0.946	0.337	0.951	0.118	0.942	0.172	0.949

表 3-2 Coverage probabilities of the EL and HEL rejection regions

n	p	$\Delta = 0$		$\Delta = 0.1$		$\Delta = 1$		$\Delta = 2$	
		EL	HEL	EL	HEL	EL	HEL	EL	HEL
100	3	0.146	0.058	0.205	0.091	0.994	0.083	1.000	0.976
	5	0.180	0.062	0.256	0.091	1.000	0.427	1.000	0.999
	8	0.282	0.066	0.325	0.082	1.000	0.732	1.000	1.000
	12	0.453	0.053	0.475	0.082	1.000	0.851	1.000	1.000
	19	0.755	0.058	0.726	0.065	1.000	0.891	1.000	1.000
	30	0.979	0.046	0.968	0.091	1.000	0.945	1.000	1.000
225	3	0.079	0.050	0.140	0.094	1.000	0.358	1.000	1.000
	5	0.097	0.044	0.129	0.089	1.000	0.907	1.000	1.000
	9	0.114	0.063	0.181	0.099	1.000	0.994	1.000	1.000
	15	0.220	0.060	0.263	0.100	1.000	1.000	1.000	1.000
	26	0.513	0.057	0.525	0.100	1.000	1.000	1.000	1.000
	45	0.925	0.045	0.911	0.108	1.000	1.000	1.000	1.000
324	3	0.066	0.041	0.132	0.128	1.000	0.577	1.000	1.000
	5	0.078	0.058	0.129	0.113	1.000	0.989	1.000	1.000
	10	0.114	0.048	0.168	0.116	1.000	1.000	1.000	1.000
	17	0.183	0.052	0.240	0.132	1.000	1.000	1.000	1.000
	30	0.392	0.052	0.526	0.136	1.000	1.000	1.000	1.000
	54	0.857	0.052	0.953	0.136	1.000	1.000	1.000	1.000
400	3	0.072	0.061	0.146	0.141	1.000	0.735	1.000	1.000
	5	0.062	0.061	0.160	0.166	1.000	0.999	1.000	1.000
	10	0.092	0.047	0.143	0.117	1.000	1.000	1.000	1.000
	18	0.137	0.046	0.271	0.141	1.000	1.000	1.000	1.000
	33	0.342	0.048	0.586	0.142	1.000	1.000	1.000	1.000
	60	0.807	0.050	0.984	0.177	1.000	1.000	1.000	1.000
625	3	0.059	0.056	0.152	0.191	1.000	0.926	1.000	1.000
	6	0.060	0.058	0.134	0.184	1.000	1.000	1.000	1.000
	11	0.076	0.052	0.151	0.168	1.000	1.000	1.000	1.000
	21	0.120	0.054	0.351	0.170	1.000	1.000	1.000	1.000
	39	0.253	0.050	0.857	0.194	1.000	1.000	1.000	1.000
	75	0.740	0.056	1.000	0.234	1.000	1.000	1.000	1.000
900	3	0.049	0.058	0.212	0.226	1.000	0.989	1.000	1.000
	6	0.050	0.055	0.168	0.221	1.000	1.000	1.000	1.000
	12	0.061	0.055	0.216	0.224	1.000	1.000	1.000	1.000
	23	0.101	0.046	0.552	0.235	1.000	1.000	1.000	1.000
	46	0.221	0.059	0.994	0.272	1.000	1.000	1.000	1.000
	90	0.654	0.049	1.000	0.334	1.000	1.000	1.000	1.000

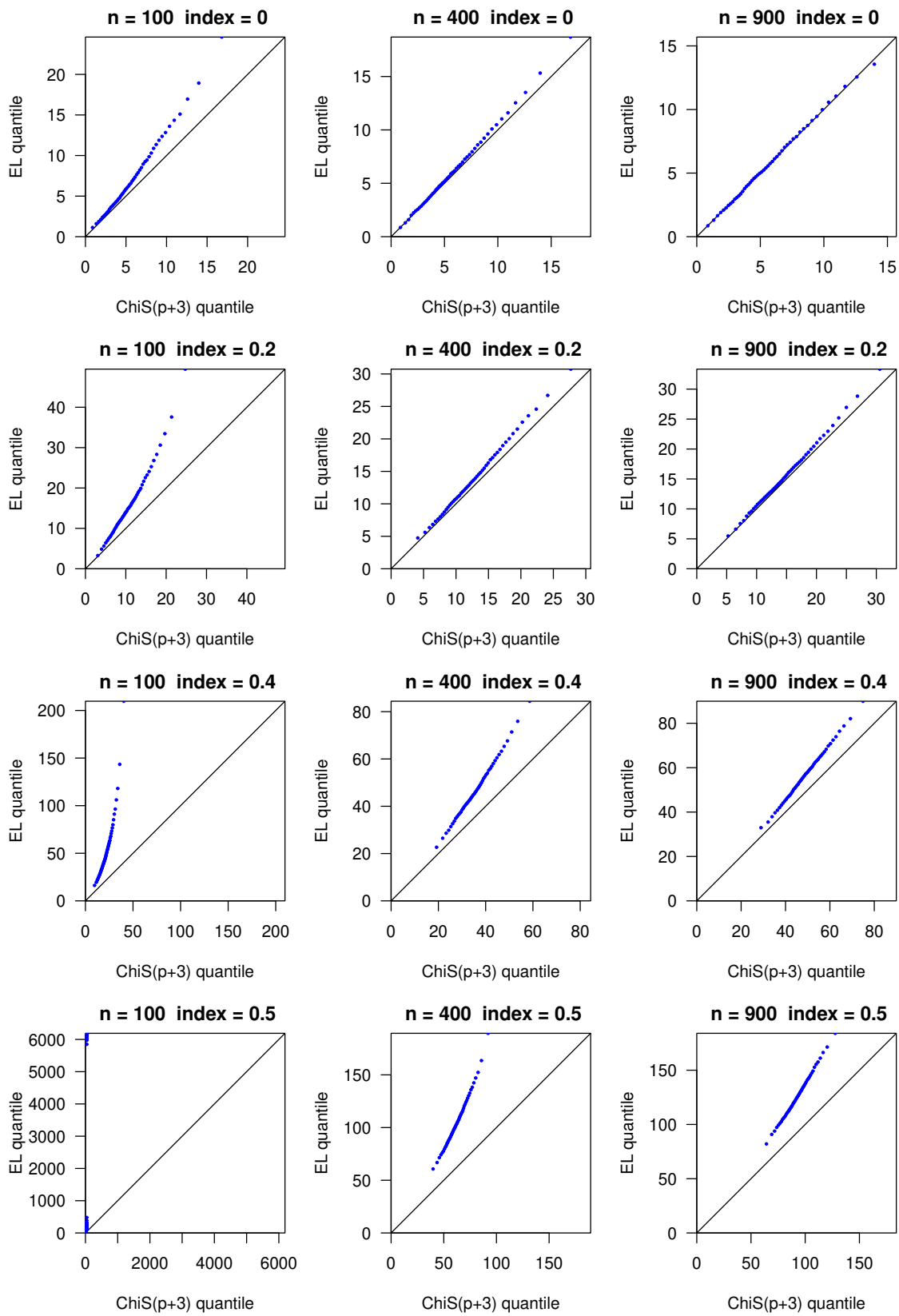


图 3-1 $\ell_n(\theta_0)$ 与 $\chi^2(p+3)$ 分位数图

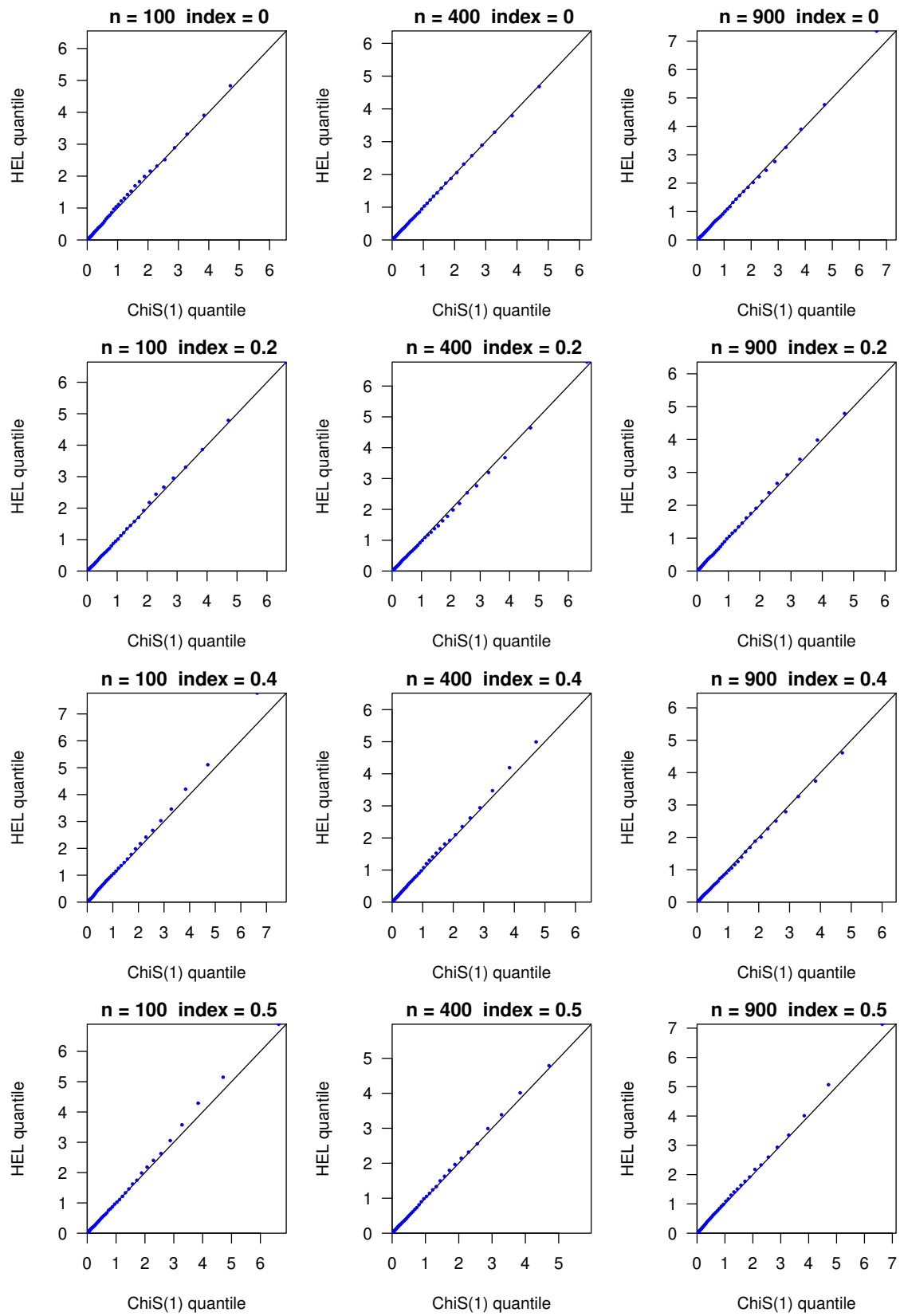


图 3-2 $\tilde{\ell}_n(\theta_0)$ 与 $\chi^2(1)$ 分位数图

§3.4 引理及定理的证明

为了证明主要结果，我们给出一些引理. 令

$$\tilde{Q}_n = \sum_{i=1}^n \sum_{j=1}^n a_{nij} \epsilon_{ni} \epsilon_{nj} + \sum_{i=1}^n b_{ni} \epsilon_{ni},$$

其中 ϵ_{ni} 是实值随机变量, 记 a_{nij} 和 b_{ni} 分别为二次型和线性形式的实值系数. 引理3.1需要以下假设条件:

(C1) $\{\epsilon_{ni}, 1 \leq i \leq n\}$ 是独立随机变量满足均值为 0 以及存在 $\eta_1 > 0$ 使得 $\sup_{1 \leq i \leq n, n \geq 1} E|\epsilon_{ni}|^{4+\eta_1} < \infty$ 成立;

(C2) 对任意 $1 \leq i, j \leq n, n \geq 1$, 有 $a_{nij} = a_{nji}$, $\sup_{1 \leq j \leq n, n \geq 1} \sum_{i=1}^n |a_{nij}| < \infty$, 以及存在 $\eta_2 > 0$, 使得 $\sup_{n \geq 1} n^{-1} \sum_{i=1}^n |b_{ni}|^{2+\eta_2} < \infty$ 成立.

给定条件 (C1) 和 (C2), 则根据 Kelejian 和 Prucha [9] 中的定理 1 可知, \tilde{Q}_n 的均值和方差如下:

$$\mu_{\tilde{Q}_n} = \sum_{i=1}^n a_{nii} \sigma_{ni}^2,$$

$$\sigma_{\tilde{Q}_n}^2 = 2 \sum_{i=1}^n \sum_{j=1}^n a_{nij}^2 \sigma_{ni}^2 \sigma_{nj}^2 + \sum_{i=1}^n b_{ni}^2 \sigma_{ni}^2 + \sum_{i=1}^n \{a_{nii}^2 (\mu_{ni}^{(4)} - 3\sigma_{ni}^4) + 2b_{ni} a_{nii} \mu_{ni}^{(3)}\}, \quad (3-3)$$

其中 $\sigma_{ni}^2 = E(\epsilon_{ni}^2)$ and $\mu_{ni}^{(s)} = E(\epsilon_{ni}^s)$ for $s = 3, 4$.

引理 3.1 若假设条件 (C1) 和 (C2) 成立, 且存在常数 $c > 0$, 满足 $n^{-1} \sigma_{\tilde{Q}_n}^2 \geq c$, 则有

$$\frac{\tilde{Q}_n - \mu_{\tilde{Q}_n}}{\sigma_{\tilde{Q}_n}} \xrightarrow{d} N(0, 1).$$

引理3.1的证明: 见 Kelejian 和 Prucha [9] 中的定理 1.

引理 3.2 令 $\xi_1, \xi_2, \dots, \xi_n$ 是一个平稳随机变量序列, 且对给常数 $s > 0$, 有 $E|\xi_1|^s < \infty$, 则

$$\max_{1 \leq i \leq n} |\xi_i| = o(n^{1/s}), \quad a.s.$$

引理3.2的证明: 由 Owen [28] 中引理 3 的证明可得.

引理 3.3 若假设条件 (A1)-(A3) 满足, 那么当 $n \rightarrow \infty$ 时, 有

$$Z_n = \max_{1 \leq i \leq n} |\tilde{\omega}_i(\theta)| = o_p(n^{1/2}) \quad a.s., \quad (3-4)$$

$$\sigma_{\tilde{Q}}^{-1} \sum_{i=1}^n \tilde{\omega}_i(\theta) \xrightarrow{d} N(0, 1), \quad (3-5)$$

$$n^{-1} \sum_{i=1}^n \omega_i(\theta) \omega_i^\tau(\theta) = n^{-1} \Sigma_{p+3} + o_p(1), \quad (3-6)$$

$$\sum_{i=1}^n |\tilde{\omega}_i(\theta)|^3 = O_p(n). \quad (3-7)$$

其中 Σ_{p+3} 见式 (1-11) 以及 $\sigma_{\tilde{Q}}^2 = \mathbf{1}_{p+3}^\tau \Sigma_{p+3} \mathbf{1}_{p+3}$.

引理3.3的证明：首先式 (3-6) 的证明见文献 [39] 中的引理 3. 接下来我们证明式 (3-4). 因为

$$\begin{aligned} Z_n &\leq \max_{1 \leq i \leq n} |\mathbf{1}_p^\tau b_i \epsilon_i| + \max_{1 \leq i \leq n} \left| \tilde{g}_{ii}(\epsilon_i^2 - \sigma^2) + 2\epsilon_i \sum_{j=1}^{i-1} \tilde{g}_{ij} \epsilon_j + s_i \epsilon_i \right| \\ &\quad + \max_{1 \leq i \leq n} \left| \tilde{h}_{ii}(\epsilon_i^2 - \sigma^2) + 2\epsilon_i \sum_{j=1}^{i-1} \tilde{h}_{ij} \epsilon_j \right| + \max_{1 \leq i \leq n} |\epsilon_i^2 - \sigma^2| \\ &\leq \max_{1 \leq i \leq n} |\mathbf{1}_p^\tau b_i \epsilon_i| + \max_{1 \leq i \leq n} \left| \tilde{g}_{ii}(\epsilon_i^2 - \sigma^2) \right| + \max_{1 \leq i \leq n} \left| 2\epsilon_i \sum_{j=1}^{i-1} \tilde{g}_{ij} \epsilon_j \right| + \max_{1 \leq i \leq n} |s_i \epsilon_i| \\ &\quad + \max_{1 \leq i \leq n} \left| \tilde{h}_{ii}(\epsilon_i^2 - \sigma^2) \right| + \max_{1 \leq i \leq n} \left| 2\epsilon_i \sum_{j=1}^{i-1} \tilde{h}_{ij} \epsilon_j \right| + \max_{1 \leq i \leq n} |\epsilon_i^2 - \sigma^2|, \end{aligned}$$

由条件 (A1) 和 (A2) 以及引理3.2可得,

$$\begin{aligned} \max_{1 \leq i \leq n} |\mathbf{1}_p^\tau b_i \epsilon_i| &\leq \|\mathbf{1}_p\| \max_{1 \leq i \leq n} \|b_i\| \max_{1 \leq i \leq n} |\epsilon_i| = o_p(n^{1/4}), \\ \max_{1 \leq i \leq n} \left| \tilde{g}_{ii}(\epsilon_i^2 - \sigma^2) \right| &= \max_{1 \leq i \leq n} |\tilde{g}_{ii}| o_p(n^{1/2}) = o_p(n^{1/2}), \\ \max_{1 \leq i \leq n} \left| \epsilon_i \sum_{j=1}^{i-1} \tilde{g}_{ij} \epsilon_j \right| &\leq \left(\max_{1 \leq i \leq n} |\epsilon_i| \right)^2 \cdot \max_{1 \leq i \leq n} \left(\sum_{j=1}^{i-1} |\tilde{g}_{ij}| \right) = o_p(n^{1/2}), \\ \max_{1 \leq i \leq n} |s_i \epsilon_i| &= \max_{1 \leq i \leq n} |s_i| o_p(n^{1/4}) = o_p(n^{1/4}), \\ \max_{1 \leq i \leq n} |\epsilon_i^2 - \sigma^2| &= o_p(n^{1/2}), \end{aligned}$$

同理可得,

$$\max_{1 \leq i \leq n} \left| \tilde{h}_{ii}(\epsilon_i^2 - \sigma^2) \right| = o_p(n^{1/2}), \quad \max_{1 \leq i \leq n} \left| \epsilon_i \sum_{j=1}^{i-1} \tilde{h}_{ij} \epsilon_j \right| = o_p(n^{1/2}),$$

因此, $Z_n = o_p(n^{1/2})$, 即式 (3-4) 的证明完成.

我们继续证明式 (3-5). 因为

$$\begin{aligned}\tilde{\omega}_i(\theta) &= \mathbf{1}_p^\tau b_i \epsilon_i + \tilde{g}_{ii}(\epsilon_i^2 - \sigma^2) + 2\epsilon_i \sum_{j=1}^{i-1} \tilde{g}_{ij} \epsilon_j + s_i \epsilon_i + \tilde{h}_{ii}(\epsilon_i^2 - \sigma^2) + 2\epsilon_i \sum_{j=1}^{i-1} \tilde{h}_{ij} \epsilon_j + (\epsilon_i^2 - \sigma^2) \\ &= (\tilde{g}_{ii} + \tilde{h}_{ii} + 1)(\epsilon_i^2 - \sigma^2) + 2\epsilon_i \sum_{j=1}^{i-1} (\tilde{g}_{ij} + \tilde{h}_{ij}) \epsilon_j + (\mathbf{1}_p^\tau b_i + s_i) \epsilon_i,\end{aligned}$$

可得,

$$\sum_{i=1}^n \tilde{\omega}_i = \sum_{i=1}^n (\tilde{g}_{ii} + \tilde{h}_{ii} + 1)(\epsilon_i^2 - \sigma^2) + 2 \sum_{i=1}^n \sum_{j=1}^{i-1} (\tilde{g}_{ij} + \tilde{h}_{ij}) \epsilon_i \epsilon_j + \sum_{i=1}^n (\mathbf{1}_p^\tau b_i + s_i) \epsilon_i.$$

令

$$Q_n = \sum_{i=1}^n \sum_{j=1}^n u_{ij} \epsilon_i \epsilon_j + \sum_{j=1}^n v_i \epsilon_i,$$

其中

$$u_{ii} = \tilde{g}_{ii} + \tilde{h}_{ii} + 1, u_{ij} = \tilde{g}_{ij} + \tilde{h}_{ij} (i \neq j), v_i = \mathbf{1}_p^\tau b_i + s_i.$$

于是可得,

$$\tilde{Q}_n = Q_n - E(Q_n) = \sum_{i=1}^n \tilde{\omega}_i = \sum_{i=1}^n \left\{ u_{ij}(\epsilon_i^2 - \sigma^2) + 2 \sum_{j=1}^{i-1} u_{ij} \epsilon_i \epsilon_j + v_i \epsilon_i \right\}.$$

我们为了获得 \tilde{Q}_n 的渐近分布, 首先检验条件 (C2) 是否成立. 由条件 A2(i) 可知,

$$\sum_{i=1}^n |u_{ij}| \leq \sum_{i=1}^n |\tilde{g}_{ij}| + \sum_{i=1}^n |\tilde{h}_{ij}| + 1 \leq C.$$

进而由 C_r 不等式可知,

$$n^{-1} \sum_{i=1}^n |v_i|^3 \leq C n^{-1} \sum_{i=1}^n |\mathbf{1}_p^\tau b_i|^3 + C n^{-1} \sum_{i=1}^n |s_i|^3, \quad (3-8)$$

其中

$$n^{-1} \sum_{i=1}^n |\mathbf{1}_p^\tau b_i|^3 \leq C \max_{1 \leq i \leq n} \|x_i\|^3 \max_{1 \leq i \leq n} \left(\sum_{k=1}^n |b_{ik}| \right)^3 \leq C, \quad (3-9)$$

此处 b_{ik} 是 $B_n(\rho_2)$ 在 (i, k) 位置上的元素, 以及由条件 (A2) 可得,

$$n^{-1} \sum_{i=1}^n |s_i|^3 \leq C. \quad (3-10)$$

由不等式 (3-8)-(3-10), 我们得出 $n^{-1} \sum_{i=1}^n |v_i|^3 \leq C$. 因此, 引理3.1中条件 (C2) 满足.

接下来我们导出 \tilde{Q}_n 的方差. 首先, 我们有

$$\begin{aligned} \sum_{i=1}^n \sum_{j=1}^n u_{ij}^2 &= \sum_{i=1}^n \left\{ (\tilde{g}_{ii} + \tilde{h}_{ii} + 1)^2 + \sum_{i \neq j} (\tilde{g}_{ij} + \tilde{h}_{ij})^2 \right\} \\ &= 2tr(\tilde{G}_n) + 2tr(\tilde{H}_n) + 2tr(\tilde{G}_n \tilde{H}_n) + n + tr(\tilde{G}_n^2) + tr(\tilde{H}_n^2), \end{aligned}$$

$$\begin{aligned} \sum_{i=1}^n u_{ii}^2 &= \sum_{i=1}^n (\tilde{g}_{ii} + \tilde{h}_{ii} + 1)^2 \\ &= \|Vec(diag(\tilde{G}_n))\|^2 + \|Vec(diag(\tilde{H}_n))\|^2 + 2tr(\tilde{G}_n) + 2tr(\tilde{H}_n) \\ &\quad + 2Vec^\tau(diag(\tilde{G}_n))Vec(diag(\tilde{H}_n)) + n, \end{aligned}$$

$$\begin{aligned} \sum_{i=1}^n v_i^2 &= \sum_{i=1}^n (\mathbf{1}_p^\tau b_i + s_i)^2 = \mathbf{1}_p^\tau \left(\sum_{i=1}^n b_i b_i^\tau \right) \mathbf{1}_p + \sum_{i=1}^n s_i^2 + \mathbf{1}_p^\tau \left(\sum_{i=1}^n b_i s_i \right) \\ &= \mathbf{1}_p^\tau \{B_n(\rho_2)X_n\}^\tau B_n(\rho_2)X_n \mathbf{1}_p \\ &\quad + \{B_n(\rho_2)W_n A_n^{-1}(\rho_1)X_n \beta\}^\tau B_n(\rho_2)W_n A_n^{-1}(\rho_1)X_n \beta \\ &\quad + 2\mathbf{1}_p^\tau \{B_n(\rho_2)X_n\}^\tau B_n(\rho_2)W_n A_n^{-1}(\rho_1)X_n \beta, \end{aligned}$$

以及

$$\begin{aligned} \sum_{i=1}^n u_{ii} v_i &= \sum_{i=1}^n (\tilde{g}_{ii} + \tilde{h}_{ii} + 1)(\mathbf{1}_p^\tau b_i + s_i) \\ &= \mathbf{1}_p^\tau \sum_{i=1}^n b_i \tilde{g}_{ii} + \sum_{i=1}^n \tilde{g}_{ii} s_i + \mathbf{1}_p^\tau \sum_{i=1}^n b_i \tilde{h}_{ii} + \sum_{i=1}^n \tilde{h}_{ii} s_i + \mathbf{1}_p^\tau \sum_{i=1}^n b_i + \sum_{i=1}^n s_i \\ &= \mathbf{1}_p^\tau \{B_n(\rho_2)X_n\}^\tau Vec(diag \tilde{G}_n) + \{B_n(\rho_2)W_n A_n^{-1}(\rho_1)X_n \beta\}^\tau Vec(diag \tilde{G}_n) \\ &\quad + \mathbf{1}_p^\tau \{B_n(\rho_2)X_n\}^\tau Vec(diag \tilde{H}_n) + \{B_n(\rho_2)W_n A_n^{-1}(\rho_1)X_n \beta\}^\tau Vec(diag \tilde{H}_n) \\ &\quad + \mathbf{1}_p^\tau \{B_n(\rho_2)X_n\}^\tau \mathbf{1}_n + \{B_n(\rho_2)W_n A_n^{-1}(\rho_1)X_n \beta\}^\tau \mathbf{1}_n. \end{aligned}$$

由引理3.1中式 (3-3) 可知 \tilde{Q}_n 的方差为:

$$\begin{aligned} \sigma_{\tilde{Q}}^2 &= 2 \sum_{i=1}^n \sum_{j=1}^n u_{ij}^2 \sigma^4 + \sum_{i=1}^n v_i^2 \sigma^2 + \sum_{i=1}^n \{u_{ii}^2 (\mu_4 - 3\sigma^4) + 2u_{ii} v_i \mu_3\} \\ &= 2\sigma^4 \left\{ 2tr(\tilde{G}_n) + 2tr(\tilde{H}_n) + 2tr(\tilde{G}_n \tilde{H}_n) + n + tr(\tilde{G}_n^2) + tr(\tilde{H}_n^2) \right\} \\ &\quad + \sigma^2 \left[\{B_n(\rho_2)W_n A_n^{-1}(\rho_1)X_n \beta\}^\tau B_n(\rho_2)W_n A_n^{-1}(\rho_1)X_n \beta \right. \\ &\quad \left. + \mathbf{1}_p^\tau \{B_n(\rho_2)X_n\}^\tau B_n(\rho_2)X_n \mathbf{1}_p + 2\mathbf{1}_p^\tau \{B_n(\rho_2)X_n\}^\tau B_n(\rho_2)W_n A_n^{-1}(\rho_1)X_n \beta \right] \\ &\quad + (\mu_4 - 3\sigma^4) \left\{ \|Vec(diag(\tilde{G}_n))\|^2 + \|Vec(diag(\tilde{H}_n))\|^2 + 2tr(\tilde{G}_n) + 2tr(\tilde{H}_n) \right\} \end{aligned}$$

$$\begin{aligned}
 & + 2Vec^\tau(diag(\tilde{G}_n))Vec(diag(\tilde{H}_n)) + n \Big\} \\
 & + 2\mu_3 \left[\mathbf{1}_p^\tau \{B_n(\rho_2)X_n\}^\tau Vec(diag\tilde{G}_n) + \{B_n(\rho_2)W_nA_n^{-1}(\rho_1)X_n\beta\}^\tau Vec(diag\tilde{G}_n) \right. \\
 & \quad + \mathbf{1}_p^\tau \{B_n(\rho_2)X_n\}^\tau Vec(diag\tilde{H}_n) + \{B_n(\rho_2)W_nA_n^{-1}(\rho_1)X_n\beta\}^\tau Vec(diag\tilde{H}_n) \\
 & \quad \left. + \mathbf{1}_p^\tau \{B_n(\rho_2)X_n\}^\tau \mathbf{1}_n + \{B_n(\rho_2)W_nA_n^{-1}(\rho_1)X_n\beta\}^\tau \mathbf{1}_n \right] \\
 & = \mathbf{1}_{p+3}^\tau \Sigma_{p+3} \mathbf{1}_{p+3}
 \end{aligned}$$

其中 Σ_{p+3} 见式 (1-11). 由条件 (A3), 我们可知 $n^{-1}\sigma_{\tilde{Q}} \geq c_1 > 0$, 进而由引理3.1, 可得

$$\frac{\tilde{Q}_n - E(\tilde{Q}_n)}{\sigma_{\tilde{Q}}} \xrightarrow{d} N(0, 1).$$

注意到 $E(\tilde{Q}_n) = 0$, 因此我们可得式 (3-5).

最后, 我们证明式 (3-7). 由条件 (A1) 和 (A2), 可知

$$\begin{aligned}
 \sum_{i=1}^n E|\tilde{\omega}_i|^3 & \leq \sum_{i=1}^n E|\mathbf{1}_p^\tau b_i \epsilon_i|^3 + \sum_{i=1}^n E|\tilde{g}_{ii}(\epsilon_i^2 - \sigma^2) + 2\epsilon_i \sum_{j=1}^{i-1} \tilde{g}_{ij}\epsilon_j + s_i \epsilon_i|^3 \\
 & \quad + \sum_{i=1}^n E|\tilde{h}_{ii}(\epsilon_i^2 - \sigma^2) + 2\epsilon_i \sum_{j=1}^{i-1} \tilde{h}_{ij}\epsilon_j|^3 + \sum_{i=1}^n E|(\epsilon_i^2 - \sigma^2)|^3 \\
 & = O(n).
 \end{aligned}$$

其余部分证明同文献 [39] 式 (4.5) 的证明. 由 Markov 不等式, 我们得到 $\sum_{i=1}^n |\tilde{\omega}_i|^3 = O_p(n)$, 式 (3-7) 的证明完成.

我们下面证明主要结果.

定理3.2的证明:

令 $\theta = \theta_0$ 以及 $\lambda = \lambda(\theta)$, 其中 λ 是方程 (3-2) 的解. 首先, 我们证明 $\lambda = O_p(n^{-1/2})$. 由方程 (3-2), 我们可得

$$0 = \bar{\omega}_n(\theta) - \frac{1}{n} \sum_{i=1}^n \frac{\lambda \tilde{\omega}_i^2(\theta)}{1 + \lambda \tilde{\omega}_i(\theta)}.$$

因为,

$$|\bar{\omega}_n(\theta)| \geq \frac{\lambda S_0}{1 + |\lambda| Z_n},$$

其中 Z_n 见式 (3-4), $\bar{\omega}_n(\theta) = n^{-1} \sum_{i=1}^n \tilde{\omega}_i(\theta)$, $S_0 = n^{-1} \sum_{i=1}^n \tilde{\omega}_i^2(\theta)$, 结合引理 3.3 和条件 (A3), 我们有

$$\frac{\lambda}{1 + |\lambda| Z_n} = O_p(n^{-1/2}).$$

因此, 由引理 3.3 可知,

$$\lambda = O_p(n^{-1/2}). \quad (3-11)$$

其次, 我们将得到 λ 的展开式. 令 $\gamma_i = \lambda \tilde{\omega}_i(\theta)$, 由式 (3-11), 得

$$\max_{1 \leq i \leq n} |\gamma_i| = o_p(1). \quad (3-12)$$

再由式 (3-2) 可得,

$$\begin{aligned} 0 &= \frac{1}{n} \sum_{i=1}^n \frac{\tilde{\omega}_i(\theta)}{1 + \lambda \tilde{\omega}_i(\theta)} \\ &= \frac{1}{n} \sum_{i=1}^n \tilde{\omega}_i(\theta) - \frac{1}{n} \sum_{i=1}^n \lambda \tilde{\omega}_i^2(\theta) + \frac{1}{n} \sum_{i=1}^n \frac{\lambda^2 \tilde{\omega}_i^3(\theta)}{1 + \lambda \tilde{\omega}_i(\theta)} \\ &= \bar{\omega}_n(\theta) - S_0 \lambda + \frac{\lambda^2}{n} \sum_{i=1}^n \frac{\tilde{\omega}_i^3(\theta)}{1 + \lambda \tilde{\omega}_i(\theta)}. \end{aligned}$$

结合式 (3-7), (3-11) 和 (3-12), 可得

$$\lambda = S_0^{-1} \bar{\omega}_n(\theta) + \zeta, \quad (3-13)$$

其中 $|\zeta|$ 的上界为

$$n^{-1} |\lambda|^2 \sum_{i=1}^n |\tilde{\omega}_i|^3 = O_p(n^{-1}). \quad (3-14)$$

因此, λ 的展开式完成.

最后, 我们展开 $\tilde{\ell}_n(\theta)$. 由式 (3-12), 我们可得 $\log(1 + \gamma_i) = \gamma_i - \gamma_i^2/2 + u_i$, 其中存在 $B > 0$, 使得

$$P(|u_i| \leq B |\gamma_i|^3, 1 \leq i \leq n) \rightarrow 1, \text{ as } n \rightarrow \infty.$$

因此, 由式 (3-1), (3-13) 和泰勒展开式, 得

$$\begin{aligned} \tilde{\ell}_n(\theta) &= 2 \sum_{j=1}^n \log(1 + \gamma_j) = 2 \sum_{j=1}^n \gamma_j - \sum_{j=1}^n \gamma_j^2 + 2 \sum_{j=1}^n u_j \\ &= 2n\lambda \bar{\omega}_n(\theta) - n\lambda^2 S_0 + 2 \sum_{j=1}^n u_j \\ &= nS_0^{-1} \bar{\omega}_n^2 - nS_0 \zeta^2 + 2 \sum_{j=1}^n u_j. \end{aligned} \quad (3-15)$$

由式 (3-6), 得

$$S_0 = \mathbf{1}^\tau \left\{ n^{-1} \sum_{i=1}^n \omega_i(\theta) \omega_i^\tau(\theta) \right\} \mathbf{1} = n^{-1} \mathbf{1}^\tau \Sigma_{p+3} \mathbf{1} + o_p(1), \quad (3-16)$$

其中 $\mathbf{1} = \mathbf{1}_{p+3}$. 将式 (3-16) 代入 (3-15), 得

$$nS_0^{-1}\bar{\omega}_n^2 = \left(\frac{n\bar{\omega}_n}{\sqrt{\mathbf{1}^\tau \Sigma_{p+3} \mathbf{1}}} \right)^2 + o_p(1).$$

再由式 (3-5), 得

$$\frac{n\bar{\omega}_n}{\sqrt{\mathbf{1}^\tau \Sigma_{p+3} \mathbf{1}}} \xrightarrow{d} N(0, 1),$$

由 Cramer-Wold 定理可知 $\tilde{\ell}_n(\theta) \xrightarrow{d} \chi_1^2$. 另一方面, 由引理 3.3, 我们可知 $nS_0\zeta^2 = O_p(n^{-1}) = o_p(1)$ 以及

$$\left| \sum_{j=1}^n u_j \right| \leq B|\lambda|^3 \sum_{j=1}^n |\tilde{\omega}_j|^3 = O_p(n^{-1/2}) = o_p(1).$$

因此, 定理 3.2 的证明完成.

结论及展望

本文主要利用两种经验似然方法研究 SARAR 模型的凸包问题和高维问题. 我们得到该模型在不同方法下的经验似然比统计量, 在一定的假设条件下, 证明了其对应的统计量的渐近分布为渐近卡方分布, 并构造模型参数的置信区域. 最后, 我们通过数值模拟验证了本文的主要结论, 同时比较本文改进的经验似然比统计量与普通经验似然比统计量的功效. 从模拟结果知道, 本文提出的经验似然方法比普通经验似然方法有更接近名义水平的覆盖率以及更快的计算速度.

本文虽然得到了一些较好的主要结论和模拟结果, 但仍然有许多问题值得今后进一步探讨, 主要有以下两个方面:

1. 从本文的主要结论看, 本文的研究是基于空间截面模型情形下进行的, 因此可以考虑更复杂的情形 (如空间面板数据模型等), 但这样的推广有一定难度, 将留待今后做深入研究;
2. 从本文的降维思想看, 本文仅仅将样本分为一组, 我们还可以考虑多组的情形.
3. 从本文的模拟结果看, 本文的模拟仅进行了有限样本的数值模拟, 我们还可以考虑其它的模拟方法.

参考文献

- [1] Tobler, R. A Computer movie simulating urban growth in the Detroit Region [J]. *Economic Geography*, 1970, **46(2)**, 234-240.
- [2] Cliff, A. D., and Ord, J. K. *Spatial autocorrelation* [M]. London: Pion Ltd, 1973.
- [3] Owen, A. B. Empirical likelihood ratio confidence intervals for a single function [J]. *Biometrika*, 1988, **75(2)**, 237-249.
- [4] 翟卫欣, 程承旗. 一种空间权重矩阵的优化方法 [J]. 武汉大学学报 (信息科学版), 2017, 42: 731-736.
- [5] 任英华, 游万海. 一种新的空间权重矩阵选择方法 [J]. 统计研究, 2012, 6: 99-105.
- [6] Mawarni M, Machdi I. Dynamic nearest neighbours for generating spatial weight matrix[C]. International Conference on Advanced Computer Science and Information Systems(ICACSIS). IEEE, 2016.
- [7] Anselin, L. *Spatial econometrics: methods and models* [M]. Berlin: Springer, 1988.
- [8] Cressie, N. *Statistics for spatial data* [M]. New York: Wiley, 1993.
- [9] Kelejian H H, Prucha I R. On the asymptotic distribution of the Moran I test statistic with applications[J].*Journal of Econometrics*, 2001, 104(2): 219-257.
- [10] Kelejian, H. H., and Prucha, L. R. A generalized moments estimator for the autoregressive parameter in a spatial model [J]. *International Economic Review*, 1999, **40(2)**, 509-533.
- [11] Lee, L. F. Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models [J]. *Econometrica*, 2004, **72(6)**, 1899-1925.
- [12] Kelejian, H. H., and Prucha, L. R. A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances [J]. *The Journal of Real Estate Finance and Economics*, 1998, **17(1)**, 99-121.
- [13] Kelejian, H. H., Prucha, I. R., and Yuzefovich, Y. Instrumental variable estimation of a spatial autoregressive model with autoregressive disturbances: large and small sample results [J]. *Spatial and Spatiotemporal Econometrics*, 2004, **18**, 163-198.

- [14] Kelejian, H. H., and Prucha, I. R. HAC estimation in a spatial framework [J]. *Journal of Econometrics*, 2006, **140(1)**, 131-154.
- [15] Arraiz, I., Drukker, D. M., Kelejian, H. H., and Prucha, I. R. A spatial cliff-ord- type model with heteroskedastic innovations: small and large sample results [J]. *Journal of Regional Science*, 2010, **50(2)**, 592-614.
- [16] Kapoor M, Kelejian H H, Prucha I R. Panel data models with spatially correlated error components[J]. *Journal of Econometrics*, 2007, 140(1): 97-130.
- [17] Lee L F, J Yu. Estimation of spatial autoregressive panel data models with fixed effects[J]. *Journal of Econometrics*. 2010, 154(2):165-85.
- [18] Lee L F, J Yu. Estimation of spatial autoregressive panel data models with fixed effects[J]. *Journal of Econometrics*, 2010, 154(2):165-185.
- [19] Fingleton B, Le Gallo J. Estimating spatial models with endogenous variables, a spatial lag and spatially dependent disturbances: finite sample properties[J]. *Papers in Regional Science*, 2008, 87(3): 319-339.
- [20] Estimation of spatial autoregressive panel data models with fixed effects[J]. *Journal of Econometrics*, 2010, 154(2): 165-185.
- [21] 文丽霞. 空间面板数据模型的研究及其应用 [D]. 广西: 广西大学, 2014.
- [22] 邓明. 时变系数空间自回归面板数据模型的极大似然估计 [J]. *统计研究*, 2016(9):96-103.
- [23] Anselin, L., 2001. Spatial econometrics. In: Baltagi, B. H. (Ed.), *A Companion to Theoretical Econometrics*. Blackwell Publishers Ltd., Massachusetts, 310-330.
- [24] Qu X, Lee L F, Yu J. QML estimation of spatial dynamic panel data models with endogenous-time varying spatial weights matrices[J]. *Journal of Econometrics*, 2017, 197(2):173-201.
- [25] Yu J, Lee L F. Estimation of Unit Root Spatial Dynamic Panel Data Models[J]. *Econometric-theory*, 2010, 26(5):1332-1362.
- [26] Baltagi B H, Fingleton B, Pirotte A. A Time-Space Dynamic Panel Data Model with Spatial Moving Average Errors[J]. *IZA Discussion Papers*, 2019,76:13-31.

- [27] Su, L., Yang, Z., 2015. QML estimation of dynamic panel data models with spatial errors. *Econometrics*, 185, 230-258.
- [28] Owen, A. B. Empirical likelihood ratio confidence regions [J]. *Ann. Statist.*, 1990, **18(1)**, 90-120.
- [29] Qin, J., and Lawless, J. Empirical likelihood and general estimating equations [J]. *Ann. Statist.*, 1994, **22(1)**, 300-325.
- [30] Owen, A. B., 1991. Empirical likelihood for linear models, *Ann. Statist.*, 19, 1725-1747.
- [31] Kolaczyk, E. D., 1994. Empirical likelihood for generalized linear models. *Statistica Sinica.*, 4, 199-218.
- [32] Qin, Y. S., 1999. Empirical likelihood ratio confidence regions in a partly linear model. *Chinese J. Applied Proba. Statist.*, 15: 363-369.
- [33] Chen, S. X., Keilegom, I. V., 2009. A review on empirical likelihood methods for regression. *Test.*, 18, 415-447.
- [34] Chen, J., Sitter, R. R., and Wu, C. Using empirical likelihood methods to obtain range restricted weights in regression estimators for surveys [J]. *Biometrika*, 2002, **89(1)**, 230-237.
- [35] Chen, J., Qin, J., 1993. Empirical likelihood estimation for finite populations and the effective usage of auxiliary information. *Biometrika.*, 80, 107-116.
- [36] Zhong, B., Rao, J. N. K., 2000. Empirical likelihood inference under stratified random sampling using auxiliary population information. *Biometrika.*, 87, 929-938.
- [37] Wu, C.B., 2004. Weighted empirical likelihood inference. *Statist. Probab. Lett.*, 66, 67-79.
- [38] Jin, F., and Lee, L. F. GEL estimation and tests of spatial autoregressive models [J]. *Journal of Econometrics*, 2019, **208(2)**, 585-612.
- [39] Qin, Y. S. Empirical likelihood for spatial autoregressive models with spatial autoregressive disturbances [J]. *Sankhyā A: The Indian Journal of Statistics*, 2021, **83(1)**, 1-25.
- [40] Qin, Y. S., Lei, Q. Z., 2021. Empirical likelihood for mixed regressive, spatial autoregressive model based on GMM. *Sankhyā A: Indian J. Statist.*, 83: 353-378.

- [41] QIN Y S. Empirical likelihood and GMM for spatial models [J]. *Communications in Statistics-Theory and Methods*, 2021, 50: 4367-4385.
- [42] LI Y H, QIN Y S, LI Y. Empirical likelihood for nonparametric regression models with spatial autoregressive errors [J]. *Journal of the Korean Statistical Society*, 2021, 50: 447-478.
- [43] RONG J R, LIU Y, QIN Y S. Empirical likelihood for spatial dynamic panel data models with spatial lags and spatial errors[J]. *Communications in Statistics-Theory and Methods*, Accepted for publication, 2021, <https://doi.org/10.1080/03610926.2022.2032172>.
- [44] 曾庆樊, 秦永松, 黎玉芳, 时变系数空间面板数据模型的经验似然推断, 广西师范大学学报 (自然科学版), 2022, 40: 30-42.
- [45] Tsao, M. Bounds on coverage probabilities of the empirical likelihood ratio confidence regions [J]. *Ann. Statist.*, 2004, **32(3)**, 1215-1221.
- [46] Owen, A. B. *Empirical Likelihood* [M]. New York: Chapman and Hall, 2001.
- [47] Bartolucci, F. A penalized version of the empirical likelihood ration for the population mean [J]. *Statist.Probab.Lett*, 2007, **77(1)**, 104-110.
- [48] Chen, J., Variyath, A. M., and Abraham, B. Adjusted empirical likelihood and its properties [J]. *Journal of Computational and Craphical Statistics*, 2008, **17(2)**, 426-443.
- [49] Emerson, S. C., and Owen A. B. Calibration of the empirical likelihood method for a vector mean [J]. *Electronic Journal of Statistics*, 2009, **3**, 1161-1192.
- [50] Shi, J., 2007. Empirical likelihood for higher dimensional linear models. *J.Sys.Sci. & Math.Scis.*, 27, 124-133.
- [51] Hjort, N. L., Eague, I. W., Eilegom, I. V., 2009. Extending the scope of empirical likelihood. *The Annals of Statistics.*, 37, 1079-1111.
- [52] Chen, S. X., Peng, L., Qin, Y. L., 2009. Effects of data dimension on empirical likelihood. *Biometrika*, 96, 711-722.
- [53] Li, G. R., Lin, L., Zhu, L. X., 2012. Empirical likelihood for a varying coefficient partially linear model with diverging number of parameters. *Journal of Multivariate Analysis*, 105, 85-111.

- [54] Tang, X. Y., Li, J. B., Lian, H., 2013. Empirical likelihood for partially linear proportional hazards models with growing dimensions . *Journal of Multivariate Analysis*, 121, 22-32.
- [55] Liu, Y., Zou, C., Wang, Z., 2013. Calibration of the empirical likelihood for high-dimensional data. *Annals of the Institute of Statistical Mathematics*, 69, 529-550.
- [56] Fang, J. L., Liu, W. R., Lu, X. W., 2017. Penalized empirical likelihood for semiparametric models with a diverging number of parameters. *Journal of Statistical Planning and Inference*, 186, 42-57.
- [57] Frank, I. E., Friedman, J. H., 1993. A statistical view of some chemometrics regression tools (with discussion). *Technometrics*, 35, 109-148.
- [58] Tibshirani, R. J., 1996, Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B*, 58, 267-288.
- [59] Antoniadis, A., 1997. Wavelets in statistics: A review. *Journal Italian Statistics Association*, 6, 97-144.
- [60] Antoniadis, A., Fan, J., 2001. Regularization of wavelets approximations. *Journal of the American Statistical Association*, 96, 939-967.
- [61] Fan, J. Q., Li, R. Z., 2001. Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American Statistical Association*, 96: 1348-1360.
- [62] Tang, X. Y., Leng, C. L., 2010. Penalized high-dimensional empirical likelihood. *Biometrika*, 97, 905-920.
- [63] Leng, C. L., Tang, X. Y., 2012. Penalized empirical likelihood and growing dimensional general estimating equations. *Biometrika*, 99, 706-716.
- [64] Lahiri, S. N. and Mukhopadhyay, S., 2012. A penalized empirical likelihood method in high dimensions. *The Annals of Statistics*, 40, 2511-2540.
- [65] Yan, L., Chen, X., 2018. Penalized quasi-Likelihood SCAD estimator in high-dimensional generalized linear models. *J. Wuhan Univ.*, 64, 533-539.
- [66] Peng, L., Qi, Y. C., Wang, R. D., 2014. Empirical likelihood test for high dimensional linear models. *Statistics and Probability Letters*, 86, 85-90.

致 谢

这是一封穿越信，证明自己存在过。

日后再回首，我一定无比感谢在师大的时光。这段熠熠生辉的日子里，我遇到许多闪闪发光的人儿，他们给我以鼓励给我以启迪给我以帮助。

首先，对永远铭记在心的老师，秦永松教授，道声感谢。秦老师在我心中是勾了金边的画，这幅画里颜色不多，黑白灰是主基调，局部有深深浅浅的亮色点缀，远看有山川有沟壑有炊烟，近看有小桥有流水有人家。老师为人谦和、低调、幽默，做事认真、细致、坚持。老师会让拘谨的你放松下来，跟你聊聊他的小时候、他的访学经历、他的小孩子，一种听大人讲故事的即视感，他也会仔细聆听你学术上的问题、听你生活上的疑惑。老师平常都把话放在肚子里，他希望问问学生论文进展到哪了可是又怕给学生压力，但是你找他，他总在。谢谢老师在学术上的指点，在问题（学习或生活）上的解答，在经历（现实或精神）上的分享。

其次，对我本科老师们，周勇老师、周志东老师、周松林老师，道声感谢。上大学的时候总是聆听你们的教诲，你们总是耐心地给学生讲述学习的重要性，把学生的事情当做自己的事情。我记得我曾一脸茫然地问周勇老师，本科结束后继续学习哪个专业比较好，老师就认真给了我建议，统计学。当时的我可真的不知道还有统计学这个专业呢！但是，老师给我指了一条路，我就朝着这个方向出发了，如今快毕业了，真的太感谢老师的建议了。谢谢老师们的指点迷津，帮助了一个又一个像我这样乡里来的孩子。过去是通往现在和未来的路，你们是我人生路上的贵人，感恩。

然后，对我研究生阶段的老师们，张军舰教授，张映辉教授，杨善朝教授，邓国和教授，熊文俊老师，唐胜达老师，晏振老师，张正家老师，黄恒振老师，…，道声感谢。各位老师的课都很丰富精彩，让坐在底下听课的我们受益匪浅。除了专业的学识水平，各位老师身上还有很多金子般闪亮的优点值得我们学习，比如刻苦、乐观、真诚等等。如同读过的书，走过的路，这些遇到过的老师都会教会我一点什么，让我收获一点什么。谢谢我的老师，你们辛苦了。

还有，感谢师兄容健荣，师姐刘雁，同门邹云龙，周雨晴，陈静，室友王斯琪，莫光妮，姚菲，朋友颜学铃，梁丽丹，许学艳，龙佳鹏，彭蒸香，等等，要不是你们的鼓励与帮助，我不会这么从容地去解决问题，要不是你们的分享与分担，我不会这么思恋研究生生活，真好有你们的相伴。

最后，感谢我的家人。逆水行舟用力撑，一篙松劲退千寻，要不是你们用力撑，我早就退千寻了。以前你们努力为我，以后我努力为你们。

论文独创性声明

本人郑重声明: 所提交的学位论文是本人在导师的指导下进行的研究工作及取得的成果. 除文中已经注明引用的内容外, 本论文不含任何个人或机构已经发表或撰写过的研究成果. 对本文的研究作出重要贡献的个人和集体, 均已在文中以明确方式标明. 本人承担本声明的法律责任.

研究生签名: _____ 日期: _____

论文使用授权声明

本人完全了解广西师范大学有关保留, 使用学位论文的规定. 广西师范大学, 中国科学技术信息研究所, 清华大学论文合作部, 有权保留本人所送交学位论文的复印件和电子文档, 可以采用影印, 缩印或其他复制手段保存论文. 本人电子文档的内容和纸质论文的内容相一致. 除在保密期间内的保密论文外, 允许论文被查阅, 可以公布 (包括刊登) 论文的全部或部分内容. 论文的公布 (包括刊登) 授权广西师范大学学位办办理.

研究生签名: _____ 日期: _____

导 师签名: _____ 日期: _____