

分类号: O212.7

单位代码: 10363

密 级: 公开

学 号: 2211020120

题 目 响应变量缺失下高维的半参数模型的
统计推断

题 目 Statistical Inference for High-dimensional
Semiparametric Models with Missing
Response Variables

学生姓名: 王 龙

指导教师: 何帮强 副教授

专 业: 数 学

研究方向: 半参数模型

论文答辩日期: 2024 年 5 月 29 日

安徽工程大学学位论文原创性声明

本人郑重声明：我恪守学术道德，崇尚严谨学风。所呈交的学位论文，是本人在导师的指导下，独立进行研究工作所取得的成果。除文中已明确注明和引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的作品及成果的内容。论文为本人亲自撰写，我对所写的内容负责，并完全意识到本声明的法律后果由本人承担。

学位论文作者签名： 子龙

日期： 2024 年 5 月 31 日

安徽工程大学学位论文版权使用授权书

学位论文作者完全了解学校有关保留、使用学位论文的规定，同意学校保留并向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅或借阅。本人授权安徽工程大学可以将本学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存和汇编本学位论文。

保密 ☐，在 _____ 年解密后适用本版权书。

本学位论文属于

不保密 ☒。

学位论文作者签名：王友

日期：2024 年 5 月 31 日

指导教师签名：何帮强

日期：2024 年 5 月 31 日

响应变量缺失下高维的半参数模型的统计推断

摘 要

半参数模型融合了参数模型中常用的方法和非参数模型中光滑的特性，并且克服了非参数模型中所谓的“维数灾祸”问题。此外，对半参数模型中参数部分附加线性约束条件可以有效减少估计偏差。当协变量的维数不断增加时，将经验似然的方法与惩罚函数相结合，可以有效进行高维协变量下的参数估计和变量选择。在科学研究和实际应用中，由于测量设备、环境条件、操作者技术水平等各种因素的影响使得数据不能精确观测，往往含有测量误差。同时部分实验结果可能由于各种原因导致记录不完整，存在数据缺失的现象。因此，本文考虑响应变量缺失下且协变量含有测量误差的半参数模型，分别研究了在附加约束条件和高维协变量下的统计推断问题。

本文首先介绍了响应变量缺失下高维的半参数模型的研究背景与意义以及半参数模型的国内外研究现状，并简要介绍了论文所涉及到的核估计、局部多项式估计和经验似然估计方法。

然后考虑缺失数据下附加线性约束条件的部分线性测量误差模型。研究当响应变量随机缺失，参数部分与非参数部分的协变量均存在测量误差时，利用局部纠偏的Profile最小二乘估计和拉格朗日乘数法在完整观测数据下与基于借补技术的方法下分别构造参数分量的约束估计，并通过使用反卷积方法在适当条件下构造了非参数分量的

估计。**Monte Carlo**模拟研究在有限样本下基于借补理论构建的约束估计要优于完整观测数据下的约束估计。同时结合香港日均因呼吸系统疾病住院人数与污染物和其他环境因素的数据进行分析,结果较好地解释了数值模拟中的结论。

最后研究在响应变量缺失下高维半参数变系数测量误差模型的参数估计与变量选择问题。首先基于逆概率加权方法分别构造了纠偏的参数部分和非参数部分估计,在适当的条件下证明纠偏的非参数估计具有渐近正态性。然后构造了参数部分经验对数似然函数,同时建议惩罚经验似然(PEL)进行变量选择,在适当的条件下证明了所提出的惩罚经验似然估计具有**Oracle**特征且在零假设下服从渐近卡方分布。**Monte Carlo**模拟研究表明建议的估计在有限样本表现较好,并给出了一个实例分析。

关键词: 半参数模型, 随机缺失, 约束估计, 逆概率加权, 惩罚经验似然

Statistical Inference for High-dimensional Semiparametric Models with Missing Response Variables

ABSTRACT

The semiparametric models combine the common methods of parametric models with the smooth properties of nonparametric models and overcome the so-called "dimensionality disaster" problem of nonparametric models. In addition, adding linear constraints to the parameter part of the semiparametric models can effectively reduce the estimation bias. When the dimension of covariates increases, the empirical likelihood method combined with the penalty function can effectively estimate parameters and select variables under high-dimensional covariates. In scientific research and practical application, due to the influence of various factors such as measuring equipment, environmental conditions, and operator's technical level, the data can not be accurately observed, and often contain measurement errors. Meanwhile, some experimental results may be incomplete due to various reasons, and there is a phenomenon of missing data. Therefore, in this thesis, we consider the semiparametric

models with missing response variables and covariates containing measurement errors, and study the statistical inference problem under additional constraints and high-dimensional covariates respectively. Thesis first introduces the research background and significance of high-dimensional semiparametric models with missing response variables and the research status of semiparametric models at home and abroad, and briefly introduces kernel estimation, local polynomial estimation, and empirical likelihood estimation methods involved in this thesis.

Then a partial linear measurement error model with additional linear constraints on missing data is considered. In this thesis, when the response variables are missing randomly and there are measurement errors in both parametric and nonparametric covariates, the locally corrected Profile least squares estimation and Lagrange multiplier method are used to construct the restricted estimation of the parameter components under the complete observation data and the method based on the imputation technique, respectively. The nonparametric component estimates are constructed under appropriate conditions by using the deconvolution method. The restricted estimation based on the theory of imputation in Monte Carlo simulation is better than the restricted estimation based on the complete observation data. Meanwhile, the data of daily hospitalizations due to respiratory diseases in Hong Kong, pollutants and other environmental factors are analyzed, and the results explain the conclusions of numerical

simulation well.

Finally, the problem of parameter estimation and variable selection in the measurement error model with a high-dimensional semiparametric variable coefficient is studied. Firstly, based on the inverse probability weighting method, the parametric and nonparametric estimates of bias correction are constructed respectively, and the asymptotic normality of the nonparametric estimates of bias correction is proved under appropriate conditions. Then the parametric partial empirical log-likelihood function is constructed, and the penalized empirical likelihood (PEL) is suggested for variable selection. Under appropriate conditions, it is proved that the proposed penalized empirical likelihood estimation has Oracle characteristics and follows an asymptotic chi-square distribution under the null hypothesis. A Monte Carlo simulation study shows that the proposed estimates perform well in a limited sample, and an example analysis is provided.

Wang Long (Mathematics)

Supervised by He Bangqiang

KEY WORDS: semiparametric models, random missing, restricted estimation, inverse probability weighting, penalized empirical likelihood

目 录

第 1 章 绪论.....	1
1.1 研究背景与研究意义.....	1
1.2 国内外研究现状.....	2
1.3 主要研究内容.....	4
第 2 章 估计方法.....	6
2.1 核估计.....	6
2.2 局部多项式估计.....	7
2.3 经验似然.....	8
第 3 章 响应变量缺失下部分线性测量误差模型的约束估计.....	10
3.1 模型框架.....	10
3.2 完整观测数据下的约束估计.....	11
3.3 基于借补技术的估计方法.....	14
3.4 数值模拟.....	16
3.5 实际例子.....	18
3.6 主要结果的证明.....	19
3.7 本章小结.....	28
第 4 章 响应变量缺失下高维半参数变系数测量误差模型的惩罚经验似然.....	30
4.1 模型框架.....	30
4.2 估计方法与主要结论.....	31
4.2.1 修正的经验似然.....	31
4.2.2 惩罚经验似然.....	34
4.2.3 算法.....	37
4.3 数值模拟.....	37
4.4 实际例子.....	40
4.5 主要结果的证明.....	41
4.6 本章小结.....	54

第 5 章 总结与展望.....	56
5.1 总结.....	56
5.2 展望.....	57
参考文献.....	58
硕士期间发表的论文.....	64
致 谢.....	65

第 1 章 绪论

1.1 研究背景与研究意义

在农业、地质、生物医药及经济管理等领域，人们往往根据现有数据对未来发展进行预测，因此回归模型成为研究经济社会发展的重要工具。回归模型由早期的参数模型发展为非参数模型，后又逐渐深化和扩大为半参数模型。参数模型通常依据经验和历史资料对变量间的关系进行假设，但是在实际中，往往会存在模型的设定误差，当假设不成立时，统计推断结果可能会很差甚至产生错误的结论。非参数模型假定响应变量和协变量之间的关系未知，通过数据本身对整个回归模型进行估计，与参数模型相比更为贴合实际。当协变量的维数越来越高时，非参数模型的估计精度会大幅下降，即产生“维数灾祸”的问题。半参数模型的提出既解决了数据降维的问题又保留了非参数模型光滑的优点，较单独的参数或非参数模型有更强的适应性和解释力。

在医学和流行病学等应用领域，数据缺失处处存在。比如在医学实验中有些个体因为严重的副作用而中途退出实验导致数据缺失，而过多的缺失数据将造成统计推断的结果出现偏差。所以在缺失数据下的统计推断具有一定的实际意义。根据缺失机制的不同可将缺失数据分为完全随机缺失、随机缺失和非随机缺失。完全随机缺失指数据集中某一变量的缺失、非缺失分组和与之相关的其他变量之间不存在任何的系统差异。随机缺失指一个变量的缺失不是由其本身所决定而是由数据集中与该变量相关的其他变量所决定。不符合上述两种缺失方式的称为非随机缺失。由于在实际应用中很少有数据符合完全随机缺失的情况并且非随机缺失的数据处理比较繁杂，因此数据随机缺失的情况受到了广泛的关注。本文主要考虑响应变量随机缺失下半参数模型的统计推断问题。

在统计推断中除依据样本信息外还可以根据样本试验之前获得的先验信息即经验和历史资料来提升统计推断结果的准确性。例如在部分线性模型中，给定线性形式的先验信息即对参数部分附加线性约束条件得到参数的约束估计量将更为有效。同时在实际研究中，往往由于各种原因使得数据不能精确观测，而是

含有测量误差。所以在附加线性约束条件的基础上进一步结合缺失数据与测量误差数据的研究更契合实际应用的需求。

随着社会科学技术的进步，金融经济、信息技术以及生物医药等领域时时刻刻都在产生大量的数据，并且数据间所蕴含的信息也会越来越多，相应的复杂程度也在不断加深。例如生物学中的基因组数据，其维数随着样本容量的增加而增加。这虽然能更加完整地将某一现象的信息表现出来，但是在数据分析上会存在重大的挑战与困难。而通过将经验似然方法与惩罚函数结合起来可以有效地解决高维协变量下的参数估计与变量选择问题。在此基础上结合缺失数据与测量误差数据可较好地拓宽半参数模型的应用范围。

1.2 国内外研究现状

Engle等^[1]在研究电力需求与气候之间的关系时首次提出了部分线性模型，既避免了线性模型的局限性也解决了非参数模型中可能产生的“维数灾祸”问题，具有广泛的应用研究价值。

$$Y = X^T \beta + g(U) + \varepsilon, \quad (1.1)$$

其中 Y 为响应变量， X ， U 为协变量， β 为 p 维参数向量， $g(\cdot)$ 为未知函数， ε 为随机误差且满足 $E(\varepsilon | X, U) = 0$ 。Speckman^[2]在该模型基础上提出了Profile最小二乘估计。Liang等^[3]和Liang^[4]分别考虑部分线性模型中参数部分与非参数部分的协变量存在测量误差的估计问题。

半参数变系数模型作为进一步的推广具有如下的形式

$$Y = X^T \beta + Z^T g(U) + \varepsilon, \quad (1.2)$$

其中 Y 为响应变量， X ， Z ， U 为协变量， $g(\cdot) = (g_1(\cdot), \dots, g_q(\cdot))^T$ 为 q 维未知函数向量， β 为 p 维参数向量， ε 为随机误差且满足 $E(\varepsilon | X, Z, U) = 0$ 。Fan和Huang^[5]在半参数变系数模型中提出了参数分量的Profile最小二乘估计，并研究了所提出的Profile似然比检验量在零假设下服从渐近卡方分布。You和Chen^[6]在半参数变系数模型中通过衰减校正，提出了参数分量的修正Profile最小二乘估计和非参数分量的局部多项式估计。魏传华和吴喜之^[7]与Wei^[8]分别在半参数变系数模型与

含测量误差的半参数变系数模型中研究附加约束条件的参数估计问题。Yang和Park^[9]考虑不同的非参数系数有不同的光滑变量的半参数变系数模型,将光滑反拟合方法与剖面技术相结合,得到参数部分的有效估计。Aman等^[10]提出了一种半参数部分线性变系数模态回归,以模态的残差和作为损失函数,构造了一个关于系数函数假设的拟合优度检验统计量。

经验似然由Thomas和Grunkemeier^[11]首次提出,随后Owen^[12]将该方法应用在总体均值的置信域构造中。经验似然是一种非参数统计方法,在构造置信域时,不需要构造枢轴统计量并且置信域的形状完全取决于数据自身,因此该方法受到了广泛关注。Shi和Lau^[13]与Li和Xue^[14]将经验似然方法分别应用在部分线性模型与含测量误差的部分线性模型中,并且构造了参数的经验似然置信域。Fan等^[15]将经验似然方法应用于异方差部分线性测量误差模型,在误差方差已知和未知的两种情况下,构造了参数部分的两种不同的经验对数似然比,并证明了其服从渐近卡方分布。You和Zhou^[16]研究了半参数变系数模型的经验似然。Fan等^[17]研究了部分时变系数含测量误差模型的经验似然推断,提出了修正的经验对数似然比函数。Yang等^[18]提出了纵向数据下变系数模型的经验似然。Fan等^[19]在非参数部分含有测量误差的半参数变系数模型下,构建了非参数部分的经验似然比统计量。Sun和Liu^[20]针对纵向数据下的半参数变系数模型,提出了一种基于指数平方损失函数和杠杆权重的稳健经验似然方法。

在统计分析中,数据缺失的现象非常普遍,而数据缺失过多将造成统计结果出现偏差。Wang和Sun^[21]利用借补、半参数回归替代和逆概率加权方法分别估计缺失数据下部分线性模型的参数和非参数。Xiao和Li^[22]借助借补技术给出在响应变量缺失下的半参数变系数模型的纠偏剖面最小二乘估计。范国良等^[23]主要研究了响应变量随机缺失情况下变系数部分非线性模型中参数和非参数的经验似然问题。Yan等^[24]基于经验似然方法研究了具有随机缺失响应的半参数变系数分位数模型的置信区间的构造。Zou等^[25]应用剖面最小二乘法和逆概率加权方法来定义协变量随机缺失时半参数变系数模型中误差方差的估计并同时构造了误差方差的Jackknife经验似然统计量。

随着科学技术的发展,对高维半参数模型的统计研究也是目前的热点之一。Zou和Zhang^[26]提出了一种结合二次正则化和自适应加权套索收缩优点的自适应

弹性网方法解决参数维数发散时线性回归模型的变量选择与估计问题。Zhu和Zhu^[27]在具有高维且高度相关的预测因子的半参数模型中提出了分布加权最小二乘估计。Wu和Li^[28]研究了当预测因子的数量随样本量的增大而趋于无穷大时一类充分降维估计量的渐近性质。Wu等^[29]针对删失数据下的线性回归模型，研究了在预测变量极大时的变量选择问题。Chen和Mao^[30]与He和Lv^[31]分别利用惩罚经验似然的方法研究了部分线性测量误差模型与鞅差误差序列下的部分线性测量误差模型的变量选择与参数估计问题。Wang和Liang^[32]主要针对超高维下带有缺失观测的半参数变系数模型，基于变系数函数的B样条方法，研究了仅使用非零系数的协变量得到的Oracle估计量的相合性。

结合上述研究成果可知半参数模型是目前的研究热点之一。将缺失数据，测量误差数据与约束条件结合起来考虑半参数模型的统计推断问题比较契合目前的实际应用。同时在高维半参数模型下运用惩罚经验似然方法解决参数估计问题受到了广泛关注，进一步考虑响应变量缺失与协变量含有测量误差的情形可拓宽半参数模型的应用范围。

1.3 主要研究内容

本文分别研究了附加约束条件的部分线性模型和高维半参数变系数模型在协变量含有测量误差且响应变量随机缺失下的统计推断问题。

具体内容如下：

1、考虑部分线性模型中参数与非参数部分的协变量均含有测量误差且响应变量随机缺失下的约束估计问题。首先通过反卷积的方法给出非参数部分在完整观测到数据下的估计量。同时对参数部分附加线性约束条件，利用局部纠偏的Profile最小二乘法得到参数部分的纠偏约束估计量。根据借补理论对缺失值进行借补，仍通过上述方法得到非参数部分的估计量与参数部分的借补纠偏约束估计量。在适当的条件下，研究了参数部分估计量的渐近性质与非参数部分估计量的一致性。根据模型与假设条件，对参数部分的纠偏约束估计量与借补的纠偏约束估计量进行数值模拟，检验估计方法的有效性。最后通过实际数据验证所研究模型的适应性。

2、研究了高维半参数变系数模型在响应变量随机缺失与协变量含有测量误

差下的参数估计与变量选择问题。首先通过参数方法选取概率函数，运用逆概率加权并对测量误差进行纠偏得到非参数函数的估计。在此基础上构造了参数部分的辅助随机向量并结合适当的条件，研究了辅助随机向量的渐近性质。然后依据经验似然的统计方法，定义了参数部分的经验对数似然比函数。当参数部分的维数发散时，结合惩罚函数得到了参数部分的惩罚检验似然函数并进行变量选择。在适当的条件下，证明了惩罚经验似然估计具有Oracle特征。此外，通过数值模拟可知建议的惩罚经验似然估计在有限样本下表现较好，并通过实际数据应用来验证所提出方法的有效性。

第 2 章 估计方法

本章将简要介绍论文运用到的主要研究方法，包括核估计、局部多项式估计和经验似然方法。

2.1 核估计

Parzen^[33]提出了核估计方法这一非参数方法用于函数估计，其思想如下：估计函数 $f(x)$ 在 x 点的取值，与更靠近 x 的样本比远离 x 的样本作用更大，那么选取 x 附近的样本并通过加权平均来对函数 $f(x)$ 进行估计，而权重为某一函数的核，即所谓的核估计方法。如下给出函数 $f(x)$ 的估计：

$$\hat{f}_K(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{X_i - x}{h}\right), \quad (2.1)$$

其中 $K(\cdot)$ 为核密度函数， h 为设定的带宽，常见的核函数如下：

均匀核：

$$K(t) = \frac{1}{2} I(|t| \leq 1),$$

Epanechnikov 核：

$$K(t) = \frac{3}{4} (1 - t^2) I(|t| \leq 1),$$

四次核：

$$K(t) = \frac{15}{16} (1 - t^2)^2 I(|t| \leq 1),$$

高斯核：

$$K(t) = \frac{1}{\sqrt{2\pi}} \exp(-t^2/2).$$

常见的核估计为 Nadaraya^[34]和 Waston^[35]分别提出的非参数函数 $f(x)$ 核估计：

$$\hat{f}_{NW}(x) = \frac{\sum_{i=1}^n K_h(X_i - x)Y_i}{\sum_{i=1}^n K_h(X_i - x)}. \quad (2.2)$$

则称 $\hat{f}_{NW}(x)$ 为 $f(x)$ 的 Nadaraya-Waston 估计, 简称为 N-W 估计。

在适当的正则性条件下, N-W 估计是相合的, 即当 $h \rightarrow 0$ 且 $nh \rightarrow \infty$ 时,

$$\hat{f}_{NW}(x) \xrightarrow{p} f(x),$$

然而非参数估计是有偏的, 其中偏差可以通过减小带宽而减小, 但是 $\hat{f}_{NW}(x)$ 的方差将增大。于是, 通常使用交叉验证的方法选择最优带宽来平衡核估计的偏差与方差。

2.2 局部多项式估计

核估计方法存在着一定的局限性, 如只是进行局部加权平均, 会产生较大的偏差, 并且偏差受协变量的密度函数影响。其次存在边界效应, 即在边界处收敛于实际函数的速度相较于在内点处的收敛速度要慢。而局部多项式估计方法相较于 N-W 估计, 减少了偏差, 并且不存在边界效应, 因此在统计研究中广泛应用。Cleveland^[36]对局部多项式估计方法作了详细的阐述, 下面简单介绍该方法。

以一元回归函数 $m(x)$ 为例, 假设其在 x_0 的邻域内存在连续的 p 阶导数, 那么由泰勒展式可得,

$$m(x) \approx \sum_{j=0}^p \frac{m^{(j)}(x_0)}{j!} (x - x_0)^j \equiv \sum_{j=0}^p \beta_j (x - x_0)^j, \quad (2.3)$$

其中 x_0 为 x 邻域内的点。因此, 求 $m(x)$ 的估计量可以归结为二次极小化问题

$$\min_{\beta \in R^{p+1}} \sum_{i=1}^n \left\{ Y_i - \sum_{j=0}^p \beta_j (X_i - x)^j \right\}^2 K_h(X_i - x), \quad (2.4)$$

其中 $\beta = (\beta_0, \dots, \beta_p)^T$ 为系数向量, $K_h(\cdot) = h^{-1}K(\cdot/h)$, $K(\cdot)$ 是核函数, h 是带宽。

局部多项式估计在边界点处表现很好, 同时在估计回归函数时, 也给出了回归函数的各阶导函数, 该部分的相关内容可见 Fan 和 Gijbels^[37]和 Ruppert^[38]。

2.3 经验似然

经验似然作为一种非参数统计方法，因其在构造置信域时的域不变性，不需要估计渐近方差等优良性质，受到了广泛的关注。下面将简要介绍经验似然方法的统计思想。

设 $X_1, \dots, X_n$ 为来源于总体 $X \in R^p$ ($p \geq 1$) 的彼此独立的随机样本，同时具有相同的分布 F_0 ，样本的经验分布函数有如下形式

$$F_n(A) = \frac{1}{n} \sum_{i=1}^n I(X_i \in A), \quad (2.5)$$

其中 $I(X_i \in A)$ 表示任一事件 $A \subseteq R^p$ 的示性函数，则定义 $F(x)$ 的非参数似然函数如下

$$L(F) = \prod_{i=1}^n F(X_i) = \prod_{i=1}^n p_i, \quad (2.6)$$

其中 $p_i = F(X_i) = P\{X = X_i\}$, $i = 1, \dots, n$ 是 $F(X_i)$ 在 X_i 处的概率。Owen^[39]证明了：若存在 $F \neq F_n$ ，则有 $L(F) < L(F_n)$ 。该结论表明 $X_1, \dots, X_n$ 的经验分布函数 $F_n(x) = n^{-1} \sum_{i=1}^n I(X_i \leq x)$ 可使式(2.6)达到极大，即 $F_n(x)$ 为 $F(x)$ 的非参数极大似然估计。那么类似可定义非参数似然比如下

$$\mathbb{R}(F) = \{L(F) / L(F_n)\}.$$

通过简单计算可得

$$\mathbb{R}(F) = \prod_{i=1}^n np_i.$$

假设 $T(\cdot)$ 为某一函数且感兴趣的中心点为 $T(F_0)$ 。 $T(F_0)$ 的非参数极大似然估计为 $T(F_n)$ 。对任意正常数 $c < 1$ ，Owen^[12]在适当的条件下证明了集合 $\{T(F) | \mathbb{R}(F) \geq c\}$ 可被用作 $T(F_0)$ 的置信域，此外通常限制 $F \ll F_n$ ，可使所得的域具有良好的性质。同时在适当条件下证明了 $-2\log(\mathbb{R}(F))$ 依分布收敛到自由度为

p 的 χ_p^2 , 其中 p 代表估计参数的个数。该渐近性质也被称为非参数版本的 Wilks 定理。

下面通过类似于参数似然的思想来说明经验似然。假设某个函数 $T(\cdot)$, 且对参数 $b = T(F)$ 感兴趣, 其中 F 为某分布族 $\mathcal{F}$ 中的元素, Owen^[12]定义如下函数

$$\mathcal{R}^*(b) = \sup\{\mathbb{R}(F) \mid T(F) = b, F \in \mathcal{F}\}. \quad (2.7)$$

对于某个临界值 r_0 , 当 $\mathcal{R}^*(b_0) < r_0$ 时, 经验似然检验拒绝 $\mathcal{R}^*(b_0) < r_0$ 。 b 的经验似然置信域为 $\{b \mid \mathcal{R}^*(b) \geq r_0\}$ 。那么可证明 $\mathcal{R}^*(b)$ 具有非参数版的 Wilks 定理。其中通过 χ^2 分布的分位点可得出临界值 r_0 。

假设 X 的均值为 $\mu_0 = (\mu_{10}, \dots, \mu_{p0}) \in R^p$, 且其方差 V_0 满秩, 考虑均值 μ_0 经验似然置信域构造问题, Owen^[12]定义经验似然比函数如下

$$\mathcal{R}^*(\mu) = \max \left\{ \prod_{i=1}^n np_i \mid p_i \geq 0, \sum_{i=1}^n p_i = 1, \sum_{i=1}^n p_i X_i = \mu \right\}. \quad (2.8)$$

通过约束条件 $\sum_{i=1}^n p_i X_i = \mu$ 将参数 μ 引入到经验似然比中, 得到关于 μ 的非参数对数似然比函数并且在该约束条件下达到极大。

Owen^[40]在适当的条件下证明了 $-2\log(\mathcal{R}^*(\mu)) \xrightarrow{d} \chi_p^2$, $\xrightarrow{d}$ 表示依分布收敛。那么, 对于任意给定的 $0 < \alpha < 1$, 参数 μ 的经验似然置信域可构造如下

$$I(\mu) = \{\mu \mid -2\log \mathcal{R}^*(\mu) \leq \chi_p^2(1-\alpha)\}.$$

其中 $\chi_p^2(1-\alpha)$ 满足 $P\{\chi_p^2 \leq \chi_p^2(1-\alpha)\} = 1-\alpha$, 称为自由度为 p 的 χ^2 分布的 $1-\alpha$ 分位点。

第 3 章 响应变量缺失下部分线性测量误差模型的约束估计

本章主要研究部分线性模型中的参数部分与非参数部分均含有测量误差且响应变量缺失的约束统计。通过反卷积方法分别得到完整观测数据与基于借补理论的非参数部分的估计，并结合局部纠偏的Profile最小二乘估计和拉格朗日乘数法分别得到参数部分的约束估计与纠偏约束估计，在适当的条件下证明了其渐近性质。最后通过数值模拟与实际例子分析所提方法的有效性。

3.1 模型框架

考虑如下的含有测量误差的部分线性模型：

$$\begin{cases} Y = X^\tau \beta + g(T) + \varepsilon, \\ U = X + v, \\ Z = T + \eta, \end{cases} \quad (3.1)$$

其中 τ 代表向量或矩阵的转置， Y 为响应变量， $X \in R^p$ 和 $T \in [0,1]$ 为协变量， $g(\cdot)$ 是未知的光滑函数。 $\beta = (\beta_1, \dots, \beta_p)^\tau$ 是未知参数变量。在协变量给定时，误差项 ε 与 (X, T) 独立且 $E(\varepsilon | X, T) = 0$ ， $E(\varepsilon^2 | X, T) = \sigma^2$ 。本章假定模型(3.1)中 X 和 T 是存在测量误差，即往往只能观测到 X 和 T 的替代变量 U 和 Z 。测量误差 v ， η 和 (X, T, ε) 独立且彼此之间也相互独立，其均值为 0，协方差分别为 Σ_v ， Σ_η 。Ma 等^[41]研究了异方差的部分线性模型，根据修正的加权估计方程，构造了参数部分和非参数部分的相合估计，并证明了参数部分的估计具有渐近正态性。

在实际分析中，数据缺失的存在会造成最终的统计结果出现偏差。Sun 等^[42]基于来自归因和逆概率加权方法的两个完整数据集，提出了两个基于经验过程的统计量来检验部分线性模型在有变量缺失时是否适合用来拟合数据。Wang^[43]考虑协变量随机缺失下的部分线性模型，通过提出模型校准方法与加权方法，定义了模型中参数部分和非参数部分的估计量，并证明了估计量的大样本性质。Xue

和 Xue^[44]在响应变量缺失下的部分线性模型中构建了参数部分的经验似然比统计量和非参数部分的经验似然置信带，并得到该参数估计量的渐近分布。在本章中，引入示性变量 δ ，来描述响应变量缺失的情况，当 $\delta=1$ 时，表示 Y 可被观测到；当 $\delta=0$ 时，表示 Y 存在缺失。本章主要考虑响应变量随机缺失的情况，假设 Y 遵循如下的缺失机制：

$$P(\delta=1|Y, X, T) = P(\delta=1|X, T) = \pi(X, T). \quad (3.2)$$

在统计应用中，模型的参数分量若满足线性约束时，可有效提升模型的参数估计效果。本章假定参数分量满足如下条件：

$$A\beta = b, \quad (3.3)$$

其中， A 为 $k \times p$ 的已知矩阵， b 是 $k \times 1$ 的已知向量，且 $\text{rank}(A) = k \leq p$ 。

Feng和Xue^[45]考虑了部分线性变系数模型在参数部分存在线性约束条件的情况下且非参数部分有测量误差的统计推断问题，给出了参数分量局部纠偏Profile最小二乘估计，并研究了其在一定正则性条件下的渐近性质。Xia等^[46]进一步研究了参数部分与非参数部分均含有测量误差的部分线性变系数模型的约束估计问题。在此基础上，本章主要考虑响应变量随机缺失且参数部分与非参数部分协变均含有测量误差的部分线性模型的偏差校正与约束估计问题。利用局部纠偏的Profile最小二乘方法和拉格朗日乘数法在完整观测数据与借补缺失值下分别构造参数分量的约束估计，并在适当条件下证明了其渐近正态性。最后通过数值模拟在缺失率，样本容量以及测量误差不同的情况下比较分析了两类参数分量的标准差与均方误差。

3.2 完整观测数据下的约束估计

假设 $\{Y_i, \delta_i, U_i, Z_i\}_{i=1}^n$ 是来自以下部分线性测量误差模型的随机样本

$$\begin{cases} Y_i = X_i^T \beta + g(T_i) + \varepsilon_i, \\ U_i = X_i + v_i, \\ Z_i = T_i + \eta_i, \end{cases} \quad (3.4)$$

其中 β 是 p 维未知参数向量，误差项 ε_i 与 (X_i, T_i) 独立且 $E(\varepsilon_i | X_i, T_i) = 0$ ，

$E(\varepsilon_i^2 | X_i, T_i) = \sigma^2$ 。测量误差 v_i ， η_i 和 $(X_i, T_i, \varepsilon_i)$ 独立且彼此之间也相互独立，其均值为 0，协方差分别为 Σ_v ， Σ_η 。有以下的等式：

$$\delta_i Y_i = \delta_i X_i^\tau \beta + \delta_i g(T_i) + \delta_i \varepsilon_i, \quad i = 1, 2, \dots, n. \quad (3.5)$$

对于非参数部分有测量误差的部分线性模型，Fan 和 Truong^[47]假定 T 具有未知密度函数 $f(t)$ ， η 具有已知的特征函数分布 $\varphi_\eta(t)$ 。通过使用反卷积的方法，对于给定的 β ，可以得到如下的 $g(\cdot)$

$$g_n(t) = \sum_{i=1}^n W_{ni}^\delta(t) (Y_i - U_i^\tau \beta), \quad (3.6)$$

其中

$$\hat{f}_n(t) = \frac{1}{nh} \sum_{j=1}^n K_n \left(\frac{t - Z_j}{h} \right), \quad (3.7)$$

$$K_n(Z) = \frac{1}{2\pi} \int_{\mathbb{R}^1} \exp(-itz) \frac{\varphi_K(t)}{\varphi_\eta(t/h)} dt, \quad (3.8)$$

$$W_{ni}^\delta(\cdot) = K_n \left(\frac{\cdot - Z_i}{h} \right) \delta_i / \sum_{j=1}^n K_n \left(\frac{\cdot - Z_j}{h} \right) \delta_j, \quad (3.9)$$

$\delta g(t) = E[\delta(Y - X^\tau \beta) | T = t] = E[\delta(Y - U^\tau \beta) | T = t]$ 且 $\varphi_K(t)$ 是核函数 $K(\cdot)$ 的 Fourier 变换。由于 X 存在测量误差，则进行局部纠偏，可以得到 β 的纠偏 Profile 最小二乘估计值如下

$$\hat{\beta}_c = \arg \min_{\beta \in \mathbb{R}^p} \left[\sum_{i=1}^n \delta_i (\tilde{Y}_i - \tilde{U}_i^\tau \beta)^2 - \sum_{i=1}^n \delta_i \beta^\tau \Sigma_v \beta \right] = \left[\sum_{i=1}^n \delta_i (\tilde{U}_i \tilde{U}_i^\tau - \Sigma_v) \right]^{-1} \sum_{i=1}^n \delta_i \tilde{U}_i \tilde{Y}_i, \quad (3.10)$$

其中 $\tilde{Y}_i = Y_i - \sum_{j=1}^n W_{nj}^\delta(Z_i) Y_j$ ， $\tilde{U}_i = U_i - \sum_{j=1}^n W_{nj}^\delta(Z_i) U_j$ ，因此非参数分量 $g(t)$ 的估计值为

$$\hat{g}_n(t) = \sum_{i=1}^n W_{ni}^\delta(t) (Y_i - U_i^\tau \hat{\beta}_c). \quad (3.11)$$

可以定义模型误差方差的估计如下

$$\hat{\sigma}_n^2 = n^{-1} \sum_{i=1}^n \delta_i [Y_i - U_i^\tau \hat{\beta}_c - \hat{g}_n(Z_i)]^2 - n^{-1} \sum_{i=1}^n \delta_i \hat{\beta}_c^\tau \Sigma_v \hat{\beta}_c. \quad (3.12)$$

若要较好地估计纠偏的Profile最小二乘估计值 $\hat{\beta}_c$ 的一致性与渐近正态性, 应考虑约束条件, 因为约束条件的有效性能降低估计产生的偏差。本节将构造一个约束估计量, 它不仅是一致的, 而且满足线性约束(3.3)。为了应用拉格朗日乘数法, 定义如下与约束条件(3.3)相对应的拉格朗日函数

$$F(\beta, \lambda) = (\tilde{Y} - \tilde{U}\beta)^\tau \Delta (\tilde{Y} - \tilde{U}\beta) - k_n \beta^\tau \Sigma_v \beta + 2\lambda^\tau (A\beta - b), \quad (3.13)$$

其中 $\Delta = \text{diag}(\delta_1, \dots, \delta_n)$, $k_n = \sum_{i=1}^n \delta_i$, λ 是 $k \times 1$ 的拉格朗日乘子向量, 将 $F(\beta, \lambda)$

分别对 β 和 λ 求导, 可得

$$\begin{cases} \frac{\partial F(\beta, \lambda)}{\partial \beta} = -2\tilde{U}^\tau \Delta \tilde{Y} - 2k_n \Sigma_v \beta + 2\tilde{U}^\tau \Delta \tilde{U} \beta + 2A^\tau \lambda = 0, \\ \frac{\partial F(\beta, \lambda)}{\partial \lambda} = 2(A\beta - b) = 0. \end{cases} \quad (3.14)$$

问题(3.14)的解如下

$$\begin{cases} \hat{\beta}_r = \hat{\beta}_c - (\tilde{U}^\tau \Delta \tilde{U} - k_n \Sigma_v)^{-1} A^\tau [A(\tilde{U}^\tau \Delta \tilde{U} - k_n \Sigma_v)^{-1} A^\tau]^{-1} (A\hat{\beta}_c - b), \\ \hat{\lambda}_1 = [A(\tilde{U}^\tau \Delta \tilde{U} - k_n \Sigma_v)^{-1} A^\tau]^{-1} (A\hat{\beta}_c - b). \end{cases} \quad (3.15)$$

$g(t)$ 的约束估计可被定义如下

$$\hat{g}_{nr}(t) = \sum_{i=1}^n W_{ni}^\delta(t) (Y_i - U_i^\tau \hat{\beta}_r), \quad (3.16)$$

σ^2 的约束估计为

$$\hat{\sigma}_{nr}^2 = n^{-1} \sum_{i=1}^n \delta_i [Y_i - U_i^\tau \hat{\beta}_r - \hat{g}_{nr}(Z_i)]^2 - n^{-1} \sum_{i=1}^n \delta_i \hat{\beta}_r^\tau \Sigma_v \hat{\beta}_r. \quad (3.17)$$

下面的定理将给出 $\hat{\beta}_r$ 的渐近正态性与 $\hat{g}_{nr}(Z_i)$ 的一致性。记 $\zeta = X - E(X|T)$, I_p

为 $p \times p$ 的单位矩阵。

定理 3.1 (i) 假设条件(A1)-(A10)成立, 则当 $n \rightarrow \infty$ 时有

$$\sqrt{n}(\hat{\beta}_r - \beta) \xrightarrow{d} N(0, (I_p - D_1) \Sigma_1^{-1} \Sigma_2 \Sigma_1^{-1} (I_p - D_1)^\tau),$$

其中

$$\Sigma_1 = E(\delta \zeta \zeta^\tau), \Sigma_2 = E(\delta \zeta \zeta^\tau) + E(\delta w w^\tau \varepsilon \varepsilon^\tau) + E(\delta \zeta v^\tau \beta)^{\otimes 2}, D_1 = \Sigma_1^{-1} A^\tau (A \Sigma_1^{-1} A^\tau)^{-1} A.$$

(ii) 假设条件(A1)-(A10)成立, 则当 $n \rightarrow \infty$ 时有

$$\sup_{1 \leq i \leq n} |\hat{g}_{nr}(Z_i) - g(Z_i)| = o_p(1), \quad \hat{\sigma}_{nr}^2 = \delta \sigma^2 + o_p(1).$$

3.3 基于借补技术的估计方法

上一节是在完整观测数据下定义的估计值 $\hat{\beta}_c$ ，但是当 Y_i 缺失时，该过程可能会降低 β 估计的效率，这是由于没有充分利用样本信息而导致的。而处理缺失数据时，借补技术很普遍，已应用于各种半参数模型。该方法的主要思想是首先为每个缺失的数据计算一个合理的值，然后进行统计推断。

定义如下

$$H_i = \delta_i Y_i + (1 - \delta_i)[U_i^\tau \beta + g(Z_i)], \quad (3.18)$$

当 $\delta_i = 0$ 时，有 $H_i = U_i^\tau \beta + g(Z_i)$ 且 $E(H | X, T) = X^\tau \beta + g(T)$ ，则

$$H_i = X_i^\tau \beta + g(T_i) + e_i, \quad (3.19)$$

其中，误差项 e_i 与 (X_i, T_i, v_i, η_i) 相互独立且 $E(e_i | X_i, T_i) = 0$ ，但是 X_i 和 T_i 都存在测量误差，所以可以通过如下等式估计 H_i

$$\hat{H}_i = \delta_i Y_i + (1 - \delta_i)[U_i^\tau \hat{\beta}_c + \hat{g}_n(Z_i)]. \quad (3.20)$$

定义

$$m_1(t) = E(X | T = t) = E(U | T = t), \quad m_2(t) = E(Y | T = t) = E(H | T = t),$$

$$m_x(t) = E(\delta X | T = t) / E(\delta | T = t), \quad m_y(t) = E(\delta Y | T = t) / E(\delta | T = t).$$

$m_1(t)$ 与 $m_2(t)$ 的核估计定义如下

$$\hat{m}_1(t) = \sum_{i=1}^n W_{ni}(t) U_i, \quad \hat{m}_2(t) = \sum_{i=1}^n W_{ni}(t) \hat{H}_i.$$

引入如下记号

$$W_{ni}(\cdot) = K_n \left(\frac{\cdot - Z_i}{h} \right) \bigg/ \sum_{j=1}^n K_n \left(\frac{\cdot - Z_j}{h} \right),$$

$$\bar{U}_i = U_i - \hat{m}_1(Z_i), \quad \bar{H}_i = \hat{H}_i - \hat{m}_2(Z_i), \quad \hat{H}_i = (\hat{H}_1, \dots, \hat{H}_n)^\tau.$$

根据 Yang 等^[48]中式(2.5)知参数 β 的借补纠偏估计值 $\hat{\beta}_I$ 可通过极小化下式得到

$$\hat{\beta}_I = \arg \min_{\beta \in R^p} \left[\sum_{i=1}^n (\bar{H}_i - \bar{U}_i^T \beta)^2 - n\beta^T \Sigma_v \beta + \sum_{i=1}^n (1 - \delta_i) \hat{\beta}_c^T \Sigma_v \beta \right], \quad (3.21)$$

式(3.21)中的第二项用于校正由测量误差引起的平方损失函数中的偏差, 式(3.21)

中的第三项是在 $\hat{H}_i$ 包含 U_i 的情况下校正的偏差。通过式(3.21)可得

$$\begin{aligned} \hat{\beta}_I = & \left[\sum_{i=1}^n \{U_i - \hat{m}_1(Z_i)\} \{U_i - \hat{m}_1(Z_i)\}^T - n\Sigma_v \right]^{-1} \\ & \cdot \sum_{i=1}^n [\{U_i - \hat{m}_1(Z_i)\} \{\hat{H}_i - \hat{m}_2(Z_i)\} - (1 - \delta_i) \Sigma_v \hat{\beta}_c / 2]. \end{aligned} \quad (3.22)$$

根据式(3.19)对关于 T 取期望可知

$$g(t) = m_2(t) - m_1^T(t)\beta, \quad (3.23)$$

且非参数分量 $g(t)$ 的估计定义如下

$$\bar{g}_n(t) = \sum_{i=1}^n W_{ni}(t) (\hat{H}_i - U_i^T \hat{\beta}_I). \quad (3.24)$$

应用拉格朗日乘数法, 定义如下与约束条件(3.3)相对应的下列拉格朗日函数

$$G(\beta, \lambda) = (\bar{H} - \bar{U}\beta)^T (\bar{H} - \bar{U}\beta) - n\beta^T \Sigma_v \beta + (n - k_n) \hat{\beta}_c^T \Sigma_v \beta + 2\lambda^T (A\beta - b). \quad (3.25)$$

将 $G(\beta, \lambda)$ 分别对 β 和 λ 求导

$$\begin{cases} \frac{\partial G(\beta, \lambda)}{\partial \beta} = -2\bar{U}^T \bar{H} - 2n\Sigma_v \beta + 2\bar{U}^T \bar{U} \beta + (n - k_n) \Sigma_v \hat{\beta}_c + 2A^T \lambda = 0, \\ \frac{\partial G(\beta, \lambda)}{\partial \lambda} = 2(A\beta - b) = 0. \end{cases} \quad (3.26)$$

问题(3.26)的解为

$$\begin{cases} \hat{\beta}_I^R = \hat{\beta}_I - (\bar{U}^T \bar{U} - n\Sigma_v)^{-1} A^T [A(\bar{U}^T \bar{U} - n\Sigma_v)^{-1} A^T]^{-1} (A\hat{\beta}_I - b), \\ \hat{\lambda}_2 = [A(\bar{U}^T \bar{U} - n\Sigma_v)^{-1} A^T]^{-1} (A\hat{\beta}_I - b). \end{cases} \quad (3.27)$$

下面的定理将给出 $\hat{\beta}_I$, $\hat{\beta}_I^R$ 的渐近正态性和非参数部分 $\bar{g}_n(t)$ 的一致性。记

$$\bar{X} = X - m_x(t), \Sigma_3 = E[\pi(X, T) \bar{X} \bar{X}^T], \Sigma_4 = E[\zeta \zeta^T], \Sigma_5 = E[(1 - \pi(X, T)) \zeta \bar{X}^T],$$

$$D_2 = \Sigma_4^{-1} A^T (A \Sigma_4^{-1} A^T)^{-1} A, \Omega^* = E[\delta((\varepsilon - v^T \beta) \bar{X})^{\otimes 2}] + E(\delta v v^T \varepsilon^2),$$

$$\Omega_1 = \Omega^* + E[\delta((v v^T - (1 + \delta) \Sigma_v / 2) \beta)^{\otimes 2}],$$

$$\Omega_2 = \Sigma_5 \Sigma_3^{-1} \{\Omega^* + E[\delta((v v^\tau - \Sigma_v) \beta)^{\otimes 2}]\} \Sigma_3^{-1} \Sigma_5,$$

$$\Omega_3 = 2 \Sigma_5 \Sigma_3^{-1} \Omega^* + \Sigma_5 \Sigma_3^{-1} E[(1 + \delta)(\Sigma_v \beta)^{\otimes 2}].$$

定理 3.2 (i) 假设条件(A1)-(A10)成立, 则当 $n \rightarrow \infty$ 时有

$$\sqrt{n}(\hat{\beta}_I - \beta) \xrightarrow{d} N(0, \Sigma^*),$$

其中 $\Sigma^* = \Sigma_4^{-1} \Gamma \Sigma_4^{-1}$, $\Gamma = \Omega_1 + \Omega_2 + \Omega_3$ 。

(ii) 假设条件(A1)-(A10)成立, 且当 $h_1 = h_2 = O_p(n^{-1/3})$, $n \rightarrow \infty$ 时有

$$\sup_{1 \leq i \leq n} |\bar{g}_n(Z_i) - g(Z_i)| = o_p(1).$$

定理 3.3 假设条件(A1)-(A10)成立, 则当 $n \rightarrow \infty$ 时有

$$\sqrt{n}(\hat{\beta}_I^R - \beta) \xrightarrow{d} N(0, (I_p - D_2) \Sigma_4^{-1} \Gamma \Sigma_4^{-1} (I_p - D_2)^\tau).$$

3.4 数值模拟

本节重点是利用数值模拟来评估所提方法在有限样本下的性能, 模拟中的数据由如下模型产生

$$\begin{cases} Y_i = X_i^\tau \beta + g(T_i) + \varepsilon_i, \\ U_i = X_i + v_i, \\ Z_i = T_i + \eta_i, \end{cases}$$

其中 $\beta = (1, 2)^\tau$, 非参数函数 $g(T_i) = 2 + \cos(2\pi T_i)$, 协变量 $Z = (1/n, 2/n, \dots, 1)^\tau$,

$X_i \sim N(\mu, \Sigma_x)$, $\mu = (1, 1)^\tau$, $\Sigma_x = I_2$, 随机误差 $\varepsilon_i \sim N(0, 1)$, 测量误差 $v_i \sim N(0, \Sigma_v)$,

$\eta_i \sim N(0, \Sigma_\eta)$, 这里取两种测量误差, 即 $\Sigma_v = 0.1 I_2$, $\Sigma_\eta = 0.1$ 和 $\Sigma_v = 0.2 I_2$, $\Sigma_\eta = 0.2$ 。

其中 I_2 是 2×2 的单位矩阵。假定参数 β 满足式(3.3), 其中 $A = (1, 1)$, $b = 3$ 。考虑如下两种缺失机制:

(i) $\pi_1(x, t) = P(\delta = 1 | X = x, T = t) = 0.6$,

(ii) $\pi_2(x, t) = P(\delta = 1 | X = x, T = t) = 0.9$ 。

可知以上两种不同缺失函数的平均的相应变量的完整率为 $E[\pi_1(x, t)] = 0.6$ 和

$E[\pi_2(x,t)]=0.9$ ，即数据缺失的概率分别为0.4,0.1。核函数根据式(3.7)和式(3.8)可选取为 $k(t)=\left(\sqrt{2\pi}\right)^{-1}\exp\{(-t^2/2)[1-(t^2-1)/9h^2]\}$ ，窗宽可结合条件(A4)选取为 $h=sd(T)\cdot n^{-1/9}$ 。下面针对模型中的参数部分进行模拟研究，在不同的缺失概率下比较仅用完全观测到的数据得到的约束纠偏估计量 $\hat{\beta}_r$ 的标准差SD1和均方误差MSE1与基于借补的约束纠偏估计量 $\hat{\beta}_I^R$ 的标准差SD2和均方误差MSE2。在模拟的过程中，样本容量 n 分别取100，300，500，在每个样本容量下，重复模拟1000次，结果如表3-1所示。

表 3-1 $\hat{\beta}_r$ 和 $\hat{\beta}_I^R$ 的标准差和均方误差

缺失率	样本容量	测量误差	SD1	MSE1	SD2	MSE2
0.4	100	$0.1I_2$	0.3635	0.1321	0.2075	0.0732
	100	$0.2I_2$	0.3829	0.1466	0.2971	0.0883
	300	$0.1I_2$	0.1930	0.0373	0.1399	0.0196
	300	$0.2I_2$	0.2083	0.0434	0.1554	0.0242
	500	$0.1I_2$	0.1504	0.0226	0.1094	0.0120
	500	$0.2I_2$	0.1619	0.0262	0.1188	0.0141
0.1	100	$0.1I_2$	0.1663	0.0276	0.1370	0.0188
	100	$0.2I_2$	0.1966	0.0386	0.1559	0.0243
	300	$0.1I_2$	0.0947	0.0090	0.0752	0.0057
	300	$0.2I_2$	0.1053	0.0111	0.0853	0.0073
	500	$0.1I_2$	0.0735	0.0054	0.0553	0.0031
	500	$0.2I_2$	0.0838	0.0070	0.0676	0.0046

从表 3-1 可得

- (1) $\hat{\beta}_I^R$ 比 $\hat{\beta}_r$ 有相对更小的标准差与均方误差。
- (2) 在缺失率，样本容量相同的情况下，随着测量误差的减少，参数估计量的标准差与均方误差也在减少。
- (3) 在缺失率，测量误差相同的情况下，随着样本容量的增加，参数估计量的标准差与均方误差在减少。

(4) 在样本容量，测量误差相同的情况下，随着缺失率的增加，参数估计量的标准差与均方误差在增加。

3.5 实际例子

本节将利用 Profile 最小二乘方法，分析 1994 年 1 月至 1997 年 12 月期间香港呼吸系统疾病住院人数与大气污染物以及温度、湿度等环境因素的关系。所使用的数据集由 $n=1451$ 个观测值组成，主要包括呼吸系统疾病住院人数(RESPI)以及二氧化氮浓度(NO_2)、二氧化硫浓度(SO_2)、可吸入悬浮颗粒物浓度(RSP)、臭氧浓度(O_3)、温度(TEMP)和湿度(HUM)的每日平均测量值。本章将响应变量 RESPI 用 Y 表示，协变量 NO_2 , SO_2 , RSP, O_3 , TEMP, HUM 分别用 $X_1, X_2, X_3, X_4, X_5, X_6$ 表示。同时令协变量 $Z_1 = (1/n, 2/n, \dots, 1)^T$ ，测量误差 $\eta_1 \sim N(0, 0.1^2)$ ，并假设 X_1 带有测量误差，即将部分线性测量误差模型构造如下：

$$\begin{cases} Y = \sum_{i=1}^6 X_i \beta_i + g(T_1) + \varepsilon_1, \\ U_1 = X_1 + v_1, \\ Z_1 = T_1 + \eta_1, \end{cases}$$

其中 $\varepsilon_1 \sim N(0, 0.1^2)$ ， $v_1 \sim N(0, 0.1^2)$ 。为了更好地说明数据缺失的情况，这里考虑数值模拟中缺失率为 0.4 和 0.1 的缺失机制，核函数与窗宽仍然按照数值模拟部分中的情况选取。参考Feng和Xue^[45]实际例子分析中对式(3.3)的设置，即令 $\beta_2 = 0$ ， $\beta_4 = 0$ ， $A = [0, 1, 0, 1, 0, 0]$ 和 $b = 0$ 。下面针对模型中的参数部分进行分析，在不同的缺失概率下比较仅用完全观测到的数据得到的约束纠偏估计量 $\hat{\beta}_{ri}$ ， $i=1, \dots, 6$ ，的估计系数和均方误差MSE3 与基于借补的约束纠偏估计量 $\hat{\beta}_{ri}^R$ ， $i=1, \dots, 6$ ，的估计系数和均方误差MSE4，结果如表 3-2 所示。

表 3-2 $\hat{\beta}_{ri}$ 和 $\hat{\beta}_{ri}^R$ 的估计系数和均方误差

缺失率	$\hat{\beta}_{ri}$		$\hat{\beta}_{ri}^R$		MSE3	MSE4
0.4	$\hat{\beta}_{r1}$	0.2655	$\hat{\beta}_{r1}^R$	0.2614	0.1472	0.1409
	$\hat{\beta}_{r2}$	0	$\hat{\beta}_{r2}^R$	0	0	0

续表 3-2 $\hat{\beta}_{ri}$ 和 $\hat{\beta}_{li}^R$ 的估计系数和均方误差

缺失率	$\hat{\beta}_{ri}$		$\hat{\beta}_{li}^R$		MSE3	MSE4
0.4	$\hat{\beta}_{r3}$	-0.1373	$\hat{\beta}_{I3}^R$	-0.1325	0.1203	0.0952
	$\hat{\beta}_{r4}$	0	$\hat{\beta}_{I4}^R$	0	0	0
	$\hat{\beta}_{r5}$	-0.7286	$\hat{\beta}_{I5}^R$	-0.8175	0.1909	0.1823
	$\hat{\beta}_{r6}$	0.4178	$\hat{\beta}_{I6}^R$	0.4372	0.0995	0.0701
0.1	$\hat{\beta}_{r1}$	0.2428	$\hat{\beta}_{I1}^R$	0.2535	0.0397	0.0269
	$\hat{\beta}_{r2}$	0	$\hat{\beta}_{I2}^R$	0	0	0
	$\hat{\beta}_{r3}$	-0.1284	$\hat{\beta}_{I3}^R$	-0.1275	0.0154	0.0125
	$\hat{\beta}_{r4}$	0	$\hat{\beta}_{I4}^R$	0	0	0
	$\hat{\beta}_{r5}$	-0.8240	$\hat{\beta}_{I5}^R$	-0.8008	0.1798	0.1571
	$\hat{\beta}_{r6}$	0.4193	$\hat{\beta}_{I6}^R$	0.4249	0.0421	0.0392

从表 3-2 可得

(1) 二氧化氮浓度, 湿度会对因呼吸系统疾病住院的人数产生正向的影响, 温度对因呼吸系统疾病住院的人数有较大的负向影响。

(2) 随着缺失率的下降, 参数估计量的均方误差是在减少的, 并且 $\hat{\beta}_{li}^R$ 比 $\hat{\beta}_{ri}$, $i=1,3,5,6$ 具有相对更小的均方误差。

3.6 主要结果的证明

在对主要结论进行证明之前, 本章首先给出了结论成立的定义及正则性条件。记 η 的分布被称为 γ 阶平滑的, 如果它的特征函数 $\varphi_n(\cdot)$ 满足 $t \rightarrow \infty$ 时,

$$d_0 |t|^{-\gamma} \leq |\varphi_n(t)| \leq d_1 |t|^{-\gamma},$$

其中 d_0, d_1, γ 均为正数。

令 X_{ij} 为 X_i 的第 j 个分量, $h_j(t) = E(X_{ij} | T_i = t)$, $\zeta_{ij} = X_{ij} - h_j(T_i)$, $\zeta_i = (\zeta_{i1}, \dots, \zeta_{ip})^T$, $1 \leq i \leq n$, $1 \leq j \leq p$ 。一些假设条件如下:

(A1) 函数 $g(\cdot)$ 和 $h_j(\cdot)$ ($1 \leq i \leq p$) 是一阶 Lipschitz 连续的。

(A2) $0 < \inf_{t \in [0,1]} f(t) \leq \sup_{t \in [0,1]} f(t) < \infty$, $f(t)$ 是有界的 m 阶导数, m 是正整

数。

(A3) 核函数 $K(\cdot)$ 满足 m 阶对称, 即满足 $K(x) = K(-x)$, $\int_{-\infty}^{\infty} K(x)dx = 1$, $\int_{-\infty}^{\infty} x^m K(x)dx \neq 0$, $\int_{-\infty}^{\infty} x^j K(x)dx = 0$, 其中 $1 \leq j \leq m-1$ 。

(A4) 误差分布是 γ 阶平滑的, 取窗宽 $h = dn^{-1/(2m+2\gamma+1)}$, 其中 $d > 0$, $2m > 2\gamma+1$, 并假定 $t \rightarrow \infty$, 常数 $c \neq 0$ 时, $t^\gamma \varphi_\eta(t) \rightarrow c$, $t^{\gamma+1} \varphi'_\eta(t) = O(1)$, 且满足

$$\int_{-\infty}^{\infty} |t|^{\gamma+1} \{\varphi_K(t) + \varphi'_K(t)\} dt < \infty, \int_{-\infty}^{\infty} |t|^{\gamma+1} \varphi_K(t)^2 dt < \infty.$$

(A5) $E(\zeta \zeta^\tau)$, $E[\pi(X, T) \zeta \zeta^\tau]$ 是正定矩阵并且 Σ_1 和 Σ_v 的特征值是有界的。

(A6) $\inf_{x,t} \pi(X, T) > 0$, $\pi(X, T)$ 的偏导数直到 2 阶均是有界的。

(A7) $m_1(\cdot)$, $m_2(\cdot)$, $m_x(\cdot)$, $m_y(\cdot)$ 的导数直到 2 阶均是有界的。

(A8) $nh_1 h_2 \rightarrow \infty$, $nh_1^4 \rightarrow 0$, $nh_2^4 \rightarrow 0$, $(h_2^2 / h_1) \rightarrow 0$ 。

(A9) $\sup_{x,t} E[Y^2 | X = x, T = t] < \infty$, $\sup_t E[\|X\|^2 | T = t] < \infty$ 。

(A10) $\sup_t E(\varepsilon^2 | T = t) < \infty$, $\sup_t E(e^2 | T = t) < \infty$, $E(\|v^2\|) < \infty$, $E(\|\eta^2\|) < \infty$ 。

引理 3.1 若条件(A1)-(A10)成立, 则

$$\max_{1 \leq i \leq n} |g(T_i) - \sum_{j=1}^n W_{nj}^\delta(Z_i) g(T_j)| = o_p(n^{-1/2}).$$

证明 参考 He 和 Lv^[31] 中引理 A.1 的证明可得到本引理。

引理 3.2 若条件(A1)-(A10)成立, 则当 $n \rightarrow \infty$ 时, 有

$$n^{-1} \sum_{i=1}^n \delta_i(\tilde{U}_i \tilde{U}_i^\tau - \Sigma_v) \xrightarrow{p} \Sigma_1.$$

证明 记 $\tilde{U}_i = \sum_{j=1}^n W_{nj}^\delta(Z_i) U_j$, 根据 $m_1(Z_i)$ 的定义, 则

$$\begin{aligned} n^{-1} \sum_{i=1}^n \delta_i(\tilde{U}_i \tilde{U}_i^\tau - \Sigma_v) &= n^{-1} \sum_{i=1}^n \delta_i\{(U_i - \tilde{U}_i)(U_i - \tilde{U}_i)^\tau - \Sigma_v\} \\ &= n^{-1} \sum_{i=1}^n \delta_i\{[U_i - m_1(Z_i)][U_i - m_1(Z_i)]^\tau - \Sigma_v\} \\ &\quad + 2n^{-1} \sum_{i=1}^n \delta_i[U_i - m_1(Z_i)][m_1(Z_i) - \tilde{U}_i]^\tau \end{aligned}$$

$$+n^{-1}\sum_{i=1}^n\delta_i[m_1(Z_i)-\check{U}_i][m_1(Z_i)-\check{U}_i]^r$$

$$=Y_1+Y_2+Y_3.$$

由大数定律可知 $Y_1 \xrightarrow{p} \Sigma_1$, 根据 Wang 和 Sun^[21]定理 2.1 的证明过程可知 Y_2 和 Y_3 依概率收敛到 0, 证毕。

引理 3.3 若条件(A1)-(A10)成立, 则当 $n \rightarrow \infty$ 时, 有

$$\sqrt{n}(\hat{\beta}_c - \beta) \xrightarrow{d} N(0, \Sigma_1^{-1} \Sigma_2 \Sigma_1^{-1}).$$

证明 定义如下 $\nabla = \sum_{i=1}^n \delta_i (\tilde{U}_i \tilde{U}_i^\tau - \Sigma_v)$, $\tilde{g}(Z_i) = g(Z_i) - \sum_{j=1}^n W_{nj}^\delta(Z_i) g(Z_j)$, $\tilde{\varepsilon}_i = \varepsilon_i - \sum_{j=1}^n W_{nj}^\delta(Z_i) \varepsilon_j$, $\tilde{v}_i = v_i - \sum_{j=1}^n W_{nj}^\delta(Z_i) v_j$, 则

$$\begin{aligned} \hat{\beta}_c - \beta &= \nabla^{-1} \sum_{i=1}^n \delta_i \tilde{U}_i \tilde{Y}_i - \beta \\ &= \nabla^{-1} \sum_{i=1}^n \delta_i \tilde{U}_i^\tau \tilde{Y}_i - \nabla^{-1} \sum_{i=1}^n \delta_i (\tilde{U}_i \tilde{U}_i^\tau - \Sigma_v) \beta \\ &= \nabla^{-1} \sum_{i=1}^n \delta_i \Sigma_v \beta + \nabla^{-1} \sum_{i=1}^n \delta_i \tilde{U}_i (\tilde{Y}_i - \tilde{U}_i^\tau \beta) \\ &= \nabla^{-1} \sum_{i=1}^n \delta_i \Sigma_v \beta + \nabla^{-1} \sum_{i=1}^n \delta_i \tilde{U}_i \tilde{g}(T_i) + \nabla^{-1} \sum_{i=1}^n \delta_i \tilde{U}_i \tilde{\varepsilon}_i - \nabla^{-1} \sum_{i=1}^n \delta_i \tilde{U}_i \tilde{v}_i \beta \\ &= C_1 + C_2 + C_3 - C_4. \end{aligned}$$

通过简单计算可知,

$$C_2 = o_p(n^{-1/2}), C_3 = \left(\frac{\nabla}{n}\right)^{-1} \left[n^{-1} \sum_{i=1}^n \delta_i \tilde{X}_i \varepsilon_i + n^{-1} \sum_{i=1}^n \delta_i v_i \varepsilon_i + o_p(n^{-1/2}) \right],$$

$$C_4 = \left(\frac{\nabla}{n}\right)^{-1} \left[n^{-1} \sum_{i=1}^n \delta_i \zeta_i v_i^\tau \beta + n^{-1} \sum_{i=1}^n \delta_i \Sigma_v \beta + o_p(n^{-1/2}) \right].$$

$$\text{则 } \hat{\beta}_c - \beta = \left(\frac{\nabla}{n}\right)^{-1} \left[n^{-1} \sum_{i=1}^n \delta_i \tilde{X}_i \varepsilon_i + n^{-1} \sum_{i=1}^n \delta_i v_i \varepsilon_i - n^{-1} \sum_{i=1}^n \delta_i \zeta_i v_i^\tau \beta + o_p(n^{-1/2}) \right].$$

应用中心极限定理, 即可知引理 3.3 成立。

引理 3.4 若条件(A1)-(A10)成立, 则当 $n \rightarrow \infty$ 时, 有

$$\sup_{1 \leq i \leq n} |\hat{g}_n(Z_i) - g(Z_i)| = o_p(1).$$

$$\begin{aligned} \text{证明} \quad & \sup_{1 \leq i \leq n} |\hat{g}_n(Z_i) - g(Z_i)| \\ & \leq \sup_{1 \leq i \leq n} |\hat{g}_n(Z_i) - g_n(Z_i)| + \sup_{1 \leq i \leq n} |g_n(Z_i) - g(Z_i)| \\ & = \sup_{1 \leq i \leq n} \left| \sum_{j=1}^n W_{nj}^\delta(Z_i)(Y_j - U_j^\tau \hat{\beta}_c) - \sum_{j=1}^n W_{nj}^\delta(Z_i)(Y_j - U_j^\tau \hat{\beta}) \right| \\ & \quad + \sup_{1 \leq i \leq n} \left| \sum_{j=1}^n W_{nj}^\delta(Z_i)[v_j + g(T_j) + g(Z_i) - g(Z_i) + \varepsilon_j] - g(Z_i) \right| \\ & \leq \sup_{1 \leq i \leq n} \left| g(Z_i) - \sum_{j=1}^n W_{nj}^\delta(Z_i)g(Z_j) \right| + \sup_{1 \leq i \leq n} \left| \sum_{j=1}^n W_{nj}^\delta(Z_i)[g(T_j) - g(Z_j)] \right| \\ & \quad + \sup_{1 \leq i \leq n} \left| \sum_{j=1}^n W_{nj}^\delta(Z_i)U_j^\tau(\hat{\beta}_c - \beta) \right| + \sup_{1 \leq i \leq n} \left| \sum_{j=1}^n W_{nj}^\delta(Z_i)v_j \right| + \sup_{1 \leq i \leq n} \left| \sum_{j=1}^n W_{nj}^\delta(Z_i)\varepsilon_j \right| \\ & = F_1 + F_2 + F_3 + F_4 + F_5. \end{aligned}$$

由引理 3.1 知 $F_1 = o_p(n^{-1/2})$, 根据 $\sum_{j=1}^n W_{nj}^\delta(\cdot)$ 的定义, 可知 $\max_{1 \leq i, j \leq n} |W_{nj}^\delta(Z_i)| = o(1)$, 则 $F_2 = F_4 = o_p(1)$ 。由引理 3.3 可知 $\hat{\beta}_c - \beta = O_p(n^{-1/2})$, 通过条件(A3), (A5)可得 $\max_{1 \leq i, j \leq n} |\tilde{U}_i| = o_p(n^{1/2})$, 即 $F_3 = o_p(1)$ 。由 Liang^[4]中的引理 A.2 的证明过程可知 $F_5 = o_p(1)$, 则可证明引理 3.4 成立。

引理 3.5 若条件(A1)-(A10)成立, 则当 $n \rightarrow \infty$ 时, 有

$$\hat{\sigma}_n^2 = \delta \sigma^2 + o_p(1).$$

证明 记 $\tilde{\varepsilon}_i = \sum_{j=1}^n W_{ni}^\delta(Z_i)\varepsilon_j$, $\tilde{v}_i = \sum_{j=1}^n W_{ni}^\delta(Z_i)v_j$, 根据式(3.11)和式(3.12)可将 $\hat{\sigma}_n^2$ 作如下分解

$$\begin{aligned} \hat{\sigma}_n^2 &= n^{-1} \sum_{i=1}^n \delta_i [Y_i - U_i^\tau \hat{\beta}_c - \hat{g}_n(Z_i)]^2 - n^{-1} \sum_{i=1}^n \delta_i \hat{\beta}_c^\tau \Sigma_v \hat{\beta}_c \\ &= n^{-1} \sum_{i=1}^n \delta_i (\tilde{Y}_i - \tilde{U}_i^\tau \hat{\beta}_c)^2 - n^{-1} \sum_{i=1}^n \delta_i \hat{\beta}_c^\tau \Sigma_v \hat{\beta}_c \\ &= n^{-1} \sum_{i=1}^n \delta_i \varepsilon_i^2 + n^{-1} \sum_{i=1}^n \delta_i (\hat{\beta}_c - \beta)^\tau \tilde{U}_i \tilde{U}_i^\tau (\hat{\beta}_c - \beta) + n^{-1} \sum_{i=1}^n \delta_i \tilde{g}^2(T_i) \end{aligned}$$

$$\begin{aligned}
 & +n^{-1} \sum_{i=1}^n \delta_i \tilde{\varepsilon}_i^2 - 2n^{-1} \sum_{i=1}^n \delta_i (\hat{\beta}_c - \beta)^\tau \tilde{U}_i \varepsilon_i + 2n^{-1} \sum_{i=1}^n \delta_i (\hat{\beta}_c - \beta)^\tau \tilde{U}_i \tilde{\varepsilon}_i \\
 & - 2n^{-1} \sum_{i=1}^n \delta_i \varepsilon_i \tilde{\varepsilon}_i + 2n^{-1} \sum_{i=1}^n \delta_i \tilde{g}(T_i) \varepsilon_i + 2n^{-1} \sum_{i=1}^n \delta_i \beta^\tau \tilde{v}_i \varepsilon_i - 2n^{-1} \sum_{i=1}^n \delta_i \beta^\tau v_i \varepsilon_i \\
 & - 2n^{-1} \sum_{i=1}^n \delta_i \beta^\tau \tilde{v}_i \tilde{\varepsilon}_i + 2n^{-1} \sum_{i=1}^n \delta_i \beta^\tau v_i \tilde{\varepsilon}_i - 2n^{-1} \sum_{i=1}^n \delta_i \tilde{g}(T_i) \tilde{U}_i^\tau (\hat{\beta}_c - \beta) \\
 & + 2n^{-1} \sum_{i=1}^n \delta_i \tilde{g}(T_i) \beta^\tau \tilde{v}_i - 2n^{-1} \sum_{i=1}^n \delta_i \tilde{g}(T_i) \beta^\tau v_i + 2n^{-1} \sum_{i=1}^n \delta_i \beta^\tau v_i \tilde{U}_i^\tau (\hat{\beta}_c - \beta) \\
 & - 2n^{-1} \sum_{i=1}^n \delta_i \beta^\tau \tilde{v}_i \tilde{U}_i^\tau (\hat{\beta}_c - \beta) - 2n^{-1} \sum_{i=1}^n \delta_i \beta^\tau v_i \tilde{v}_i \beta + n^{-1} \sum_{i=1}^n \delta_i \beta^\tau \tilde{v}_i \tilde{v}_i^\tau \beta \\
 & - n^{-1} \sum_{i=1}^n \delta_i (\hat{\beta}_c - \beta)^\tau \Sigma_v (\hat{\beta}_c - \beta) - 2n^{-1} \sum_{i=1}^n \delta_i \tilde{g}(T_i) \bar{\varepsilon}_i - 2n^{-1} \sum_{i=1}^n \delta_i (\hat{\beta}_c - \beta)^\tau \Sigma_v \beta \\
 & + n^{-1} \sum_{i=1}^n \delta_i \beta^\tau (v_i v_i^\tau - \Sigma_v) \beta \\
 & =: \sum_{i=1}^{23} M_{ni}.
 \end{aligned}$$

对于 M_{n1} 来说, 应用大数定律可得 M_{n1} 依概率收敛到 $\delta\sigma^2$ 。根据引理 3.2 与引理 3.3 可知 $M_{n2} = o_p(1)$ 。根据引理 3.4 可知 $M_{n3} = M_{n4} = o_p(1)$ 。由 Cauchy-Schwarz 不等式有,

$$|M_{n5}| \leq O_p(n^{-1/2}) n^{-1} \sum_{i=1}^n \|\tilde{U}_i \varepsilon_i\| \leq O_p(n^{-1/2}) \left(n^{-1} \sum_{i=1}^n \|\tilde{U}_i\|^2 \right)^{-1/2} \left(n^{-1} \sum_{i=1}^n |\varepsilon_i|^2 \right)^{1/2} = o_p(1),$$

再由 $\max_{1 \leq i \leq n} |\tilde{\varepsilon}_i| = o(1)$ 有

$$|M_{n6}| \leq O_p(n^{-1/2}) \max_{1 \leq i \leq n} \|\tilde{U}_i\| \max_{1 \leq i \leq n} |\tilde{\varepsilon}_i| = o_p(1).$$

类似可以得到 $M_{ni} = o_p(1), i = 7, \dots, 23$ 。证毕。

定理 3.1 的证明

首先证明(i), 根据式(3.15)中 $\hat{\beta}_r$ 的定义与式(3.3)可知

$$\hat{\beta}_r - \beta = \hat{\beta}_c - \beta - (\tilde{U}^\tau \Delta \tilde{U} - k_n \Sigma_v)^{-1} A^\tau [A(\tilde{U}^\tau \Delta \tilde{U} - k_n \Sigma_v)^{-1} A^\tau]^{-1} (A\hat{\beta}_c - b)$$

$$\begin{aligned}
 &= \hat{\beta}_c - \beta - (\tilde{U}^\tau \Delta \tilde{U} - k_n \Sigma_v)^{-1} A^\tau [A(\tilde{U}^\tau \Delta \tilde{U} - k_n \Sigma_v)^{-1} A^\tau]^{-1} A(\hat{\beta}_c - \beta) \\
 &= \{I_p - (\tilde{U}^\tau \Delta \tilde{U} - k_n \Sigma_v)^{-1} A^\tau [A(\tilde{U}^\tau \Delta \tilde{U} - k_n \Sigma_v)^{-1} A^\tau]^{-1} A\}(\hat{\beta}_c - \beta),
 \end{aligned}$$

结合引理 3.2 与引理 3.3, 依据中心极限定理即证。

下证定理 3.1(ii)

$$\sup_{1 \leq i \leq n} |\hat{g}_{nr}(Z_i) - g(Z_i)| \leq \sup_{1 \leq i \leq n} |\hat{g}_n(Z_i) - g(Z_i)| + \sup_{1 \leq i \leq n} \left| \sum_{j=1}^n W_{nj}^\delta(Z_i) U_j^\tau (\hat{\beta}_r - \hat{\beta}_c) \right|,$$

根据 Cauchy-Schwarz 不等式可得

$$\begin{aligned}
 \sup_{1 \leq i \leq n} \left| \sum_{j=1}^n W_{nj}^\delta(Z_i) U_j^\tau (\hat{\beta}_r - \hat{\beta}_c) \right| &\leq \sup_{1 \leq i \leq n} \left| \sum_{j=1}^n [W_{nj}^\delta(Z_i)]^2 \right|^{1/2} \cdot \left| (\hat{\beta}_r - \hat{\beta}_c)^\tau \sum_{i=1}^n \tilde{U}_i \tilde{U}_i^\tau (\hat{\beta}_r - \hat{\beta}_c) \right|^{1/2} \\
 &\leq \sup_{1 \leq i \leq n} |W_{nj}^\delta(Z_i)|^{1/2} \cdot \sup_{1 \leq i \leq n} \left| \sum_{j=1}^n W_{nj}^\delta(Z_i) \right|^{1/2} \cdot O_p(1) = o_p(1).
 \end{aligned}$$

再结合引理 3.4 即可得到定理 3.1(ii)的第一个结论。

根据式(3.16)和式(3.17)可对 $\hat{\sigma}_{nr}^2$ 作如下分解

$$\begin{aligned}
 \hat{\sigma}_{nr}^2 &= n^{-1} \sum_{i=1}^n \delta_i [Y_i - U_i^\tau \hat{\beta}_r - \hat{g}_{nr}(Z_i)]^2 - n^{-1} \sum_{i=1}^n \delta_i \hat{\beta}_r^\tau \Sigma_v \hat{\beta}_r \\
 &= n^{-1} \sum_{i=1}^n \delta_i (\tilde{Y}_i - \tilde{U}_i^\tau \hat{\beta}_r)^2 - n^{-1} \sum_{i=1}^n \delta_i \hat{\beta}_r^\tau \Sigma_v \hat{\beta}_r \\
 &= n^{-1} \sum_{i=1}^n \delta_i [(\tilde{Y}_i - \tilde{U}_i^\tau \hat{\beta}_c) + \tilde{U}_i^\tau (\hat{\beta}_c - \hat{\beta}_r)]^2 - n^{-1} \sum_{i=1}^n \delta_i (\hat{\beta}_r - \hat{\beta}_c + \hat{\beta}_c)^\tau \Sigma_v (\hat{\beta}_r - \hat{\beta}_c + \hat{\beta}_c) \\
 &= \hat{\sigma}_n^2 + 2n^{-1} \sum_{i=1}^n \delta_i (\tilde{Y}_i - \tilde{U}_i^\tau \hat{\beta}_c) \tilde{U}_i^\tau (\hat{\beta}_c - \hat{\beta}_r) + n^{-1} \sum_{i=1}^n \delta_i (\hat{\beta}_r - \hat{\beta}_c)^\tau \tilde{U}_i \tilde{U}_i^\tau (\hat{\beta}_r - \hat{\beta}_c) \\
 &\quad - n^{-1} \sum_{i=1}^n \delta_i (\hat{\beta}_r - \hat{\beta}_c)^\tau \Sigma_v (\hat{\beta}_r - \hat{\beta}_c) - 2n^{-1} \sum_{i=1}^n \delta_i (\hat{\beta}_r - \hat{\beta}_c)^\tau \Sigma_v \hat{\beta}_c \\
 &= \hat{\sigma}_n^2 + Q_1 + Q_2 - Q_3 - Q_4.
 \end{aligned}$$

根据 Cauchy-Schwarz 不等式可得

$$|Q_1| \leq 2 \left| n^{-1} \sum_{i=1}^n [\delta_i (\tilde{Y}_i - \tilde{U}_i^\tau \hat{\beta}_c)]^2 \right|^{1/2} \cdot \left| n^{-1} (\hat{\beta}_r - \hat{\beta}_c)^\tau \sum_{i=1}^n \tilde{U}_i \tilde{U}_i^\tau (\hat{\beta}_r - \hat{\beta}_c) \right|^{1/2}$$

$$= 2 \left| \hat{\sigma}_n^2 + n^{-1} \sum_{i=1}^n \delta_i \hat{\beta}_c^T \Sigma_v \hat{\beta}_c \right|^{1/2} \cdot o_p(1) = o_p(1).$$

由引理 3.3 和定理 3.1(i)可知 $Q_2 = Q_3 = Q_4 = o_p(1)$ ，证毕。

定理 3.2 的证明

定义如下

$$\Phi = n^{-1} \sum_{i=1}^n (\bar{U}_i \bar{U}_i^T - \Sigma_v),$$

$$\Lambda_n = n^{-1/2} \sum_{i=1}^n \{ [U_i - \hat{m}_1(Z_i)] \{ \hat{H}_i - \hat{m}_2(Z_i) - [U_i - \hat{m}_1(Z_i)]^T \beta \} + \Sigma_v \beta - (1 - \delta_i) \Sigma_v \hat{\beta}_c / 2 \}.$$

则 $\sqrt{n}(\hat{\beta}_I - \beta) = \Phi^{-1} \Lambda_n$ ，根据 $m_1(Z_i)$ 与 $\hat{m}_1(Z_i)$ 的定义，计算可得

$$\begin{aligned} n^{-1} \sum_{i=1}^n (\bar{U}_i \bar{U}_i^T - \Sigma_v) &= n^{-1} \sum_{i=1}^n \{ [U_i - \hat{m}_1(Z_i)] [U_i - \hat{m}_1(Z_i)]^T - \Sigma_v \} \\ &= n^{-1} \sum_{i=1}^n \{ [U_i - m_1(Z_i)] [U_i - m_1(Z_i)]^T - \Sigma_v \} \\ &\quad + 2n^{-1} \sum_{i=1}^n [U_i - m_1(Z_i)] [m_1(Z_i) - \hat{m}_1(Z_i)]^T \\ &\quad + n^{-1} \sum_{i=1}^n [m_1(Z_i) - \hat{m}_1(Z_i)] [m_1(Z_i) - \hat{m}_1(Z_i)]^T \\ &= \Theta_1 + \Theta_2 + \Theta_3. \end{aligned}$$

由大数定律可知 $\Theta_1 \xrightarrow{p} \Sigma_4$ ，根据 Wang 和 Sun^[21]定理 2.1 的证明过程可知 Θ_2 和

Θ_3 依概率收敛到 0。对于 Λ_n 可做如下分解

$$\Lambda_n = n^{-1/2} \sum_{i=1}^n \Psi_{1i} + n^{-1/2} \sum_{i=1}^n \Psi_{2i} = \Lambda_{n1} + \Lambda_{n2},$$

其中

$$\begin{aligned} \Psi_{1i} &= [U_i - m_1(Z_i)] [\delta_i Y_i + (1 - \delta_i)(U_i^T \hat{\beta}_c + \hat{g}_n(Z_i)) - \hat{m}_2(Z_i) - (U_i - \hat{m}_1(Z_i))^T \beta] \\ &\quad + \Sigma_v \beta - (1 - \delta_i) \Sigma_v \hat{\beta}_c / 2, \end{aligned}$$

$$\Psi_{2i} = [m_1(Z_i) - \hat{m}_1(Z_i)] [\delta_i Y_i + (1 - \delta_i)(U_i^T \hat{\beta}_c + \hat{g}_n(Z_i)) - \hat{m}_2(Z_i) - (U_i - \hat{m}_1(Z_i))^T \beta].$$

可得

$$\begin{aligned}
 \Lambda_{n1} &= n^{-1/2} \sum_{i=1}^n \{ [U_i - m_1(Z_i)] [\delta_i Y_i + (1 - \delta_i)(U_i^\tau \hat{\beta}_c + g(Z_i)) - m_2(Z_i) \\
 &\quad - (U_i - m_1(Z_i))^\tau \beta] + \delta_i \Sigma_v \beta \} \\
 &\quad + n^{-1/2} \sum_{i=1}^n (1 - \delta_i) \left\{ [U_i - m_1(Z_i)] \left[U_i - \sum_{j=1}^n W_{nj}^\delta(Z_i) U_j \right]^\tau - \Sigma_v \right\} (\hat{\beta}_c - \beta) \\
 &\quad + n^{-1/2} \sum_{i=1}^n (1 - \delta_i) \left\{ [U_i - m_1(Z_i)] \left[\sum_{j=1}^n W_{nj}^\delta(Z_i) U_j \right]^\tau \right\} (\hat{\beta}_c - \beta) \\
 &\quad + n^{-1/2} \sum_{i=1}^n \{ [U_i - m_1(Z_i)] (1 - \delta_i) [\hat{g}_n(Z_i) - g(Z_i)] \} \\
 &\quad + n^{-1/2} \sum_{i=1}^n \{ [U_i - m_1(Z_i)] [m_2(Z_i) - \hat{m}_2(Z_i)] \} \\
 &\quad - n^{-1/2} \sum_{i=1}^n \{ [U_i - m_1(Z_i)] [m_1(Z_i) - \hat{m}_1(Z_i)]^\tau \beta \} \\
 &\quad + n^{-1/2} \sum_{i=1}^n (1 - \delta_i) \Sigma_v (\hat{\beta}_c - \beta) / 2 \\
 &\quad + n^{-1/2} \sum_{i=1}^n (1 - \delta_i) \Sigma_v \beta / 2 \\
 &= \Lambda_{n11} + \Lambda_{n12} + \Lambda_{n13} + \Lambda_{n14} + \Lambda_{n15} - \Lambda_{n16} + \Lambda_{n17} + \Lambda_{n18}.
 \end{aligned}$$

由 $\sum_{j=1}^n W_{nj}(\cdot)$ 的定义和 $\hat{\beta}_c - \beta = O_p(n^{-1/2})$ 并且类似 Wang 和 Sun^[21] 中的定理 2.1 的证明与大数定律可得

$$\begin{aligned}
 \Lambda_{n11} &= n^{-1/2} \sum_{i=1}^n \delta_i \{ [X_i + v_i - m_1(Z_i)] (\varepsilon_i - v_i^\tau \beta) + \Sigma_v \beta \}, \\
 \Lambda_{n12} &= E \{ [1 - \pi(X, T)] [X - m_1(Z)] [X - m_x(Z)]^\tau \} \\
 &\quad \cdot n^{-1/2} \Sigma_3^{-1} \sum_{i=1}^n \delta_i \{ [X_i - m_x(Z_i) + v_i] (\varepsilon_i - v_i^\tau \beta) + \Sigma_v \beta \} + o_p(1), \\
 \Lambda_{n14} &= -n^{-1/2} \sum_{i=1}^n \delta_i (\varepsilon_i - v_i^\tau \beta) [m_x(Z_i) - m_1(Z_i)] + o_p(1),
 \end{aligned}$$

$$\Lambda_{n13} = o_p(1), \Lambda_{n15} = o_p(1), \Lambda_{n16} = o_p(1), \Lambda_{n17} = o_p(1),$$

$$\Lambda_{n18} = n^{-1/2} \sum_{i=1}^n (1 - \delta_i) \Sigma_v \beta / 2.$$

结合 Λ_{n11} 至 Λ_{n18} 可得

$$\begin{aligned} \Lambda_{n1} &= n^{-1/2} \sum_{i=1}^n \delta_i [(\bar{X}_i + v_i)(\varepsilon_i - v_i^\tau \beta) + (1 + \delta_i) \Sigma_v \beta / 2] \\ &\quad + E\{[1 - \pi(X, T)] \zeta \bar{X}^\tau\} n^{-1/2} \Sigma_3^{-1} \sum_{i=1}^n \delta_i [(\bar{X}_i + v_i)(\varepsilon_i - v_i^\tau \beta) + \Sigma_v \beta] + o_p(1). \end{aligned}$$

通过类似的证明可得 $\Lambda_{n2} = o_p(1)$ 。应用中心极限定理，即可知定理 3.2 的第一个结论成立。

下证定理 3.2(ii)

根据式(3.23) $g(t)$ 与式(3.24) $\bar{g}_n(t)$ 的定义可知

$$\begin{aligned} \bar{g}_n(Z_i) - g(Z_i) &= \hat{m}_2(Z_i) - m_2(Z_i) - [\hat{m}_1(Z_i) - m_1(Z_i)]^\tau \beta \\ &\quad - [\hat{m}_1(Z_i) - m_1(Z_i)]^\tau (\hat{\beta}_1 - \beta) - m_1^\tau(Z_i) (\hat{\beta}_1 - \beta). \end{aligned}$$

根据 $H_i = \delta_i Y_i + (1 - \delta_i)[U_i^\tau \beta + g(Z_i)]$, $\hat{m}_2(Z_i)$ 的定义和式(3.20) $\hat{H}_i$ 的定义可知

$$\begin{aligned} \hat{m}_2(Z_i) - m_2(Z_i) &= \sum_{i=1}^n W_{ni}(Z_i) \hat{H}_i - m_2(Z_i) \\ &= \sum_{i=1}^n W_{ni}(Z_i) \{\delta_i Y_i + (1 - \delta_i)[U_i^\tau \hat{\beta}_c + \hat{g}_n(Z_i)]\} - m_2(Z_i) \\ &= \sum_{i=1}^n W_{ni}(Z_i) H_i - m_2(Z_i) + \sum_{i=1}^n W_{ni}(Z_i) (1 - \delta_i) [U_i^\tau (\hat{\beta}_c - \beta)] \\ &\quad + \sum_{i=1}^n W_{ni}(Z_i) (1 - \delta_i) [\hat{g}_n(Z_i) - g(Z_i)]. \end{aligned}$$

类似 Wang 和 Sun^[21]中的定理 2.2 的证明可知

$$\begin{aligned} \sup_{1 \leq i \leq n} \left| \sum_{i=1}^n W_{ni}(Z_i) H_i - m_2(Z_i) \right| &= O_p(nh_1)^{-1/2} + O_p(h_1), \\ \sup_{1 \leq i \leq n} |\hat{m}_1(Z_i) - m_1(Z_i)| &= O_p(nh_2)^{-1/2} + O_p(h_2). \end{aligned}$$

由定理 3.2(i)可知 $\hat{\beta}_I - \beta = O_p(n^{-1/2})$, 根据 $\sum_{j=1}^n W_{nj}(\cdot)$ 的定义和 $\hat{\beta}_c - \beta = O_p(n^{-1/2})$ 两个条件与引理 3.4, 在 $h_1 = h_2 = O_p(n^{-1/3})$ 时可知

$$\begin{aligned} \sup_{1 \leq i \leq n} |\bar{g}_n(Z_i) - g(Z_i)| &= O_p(nh_1)^{-1/2} + O_p(h_1) + O_p(n^{-1/2}) + o_p(1) + O_p(nh_2)^{-1/2} \\ &\quad + O_p(h_2) + [O_p(nh_2)^{-1/2} + O_p(h_2)]O_p(n^{-1/2}) + O_p(n^{-1/2}) \\ &= o_p(1). \end{aligned}$$

证毕。

定理 3.3 的证明

根据式(3.26)中 $\hat{\beta}_I^R$ 的定义与式(3.3)可知

$$\begin{aligned} \hat{\beta}_I^R - \beta &= \hat{\beta}_I - \beta - (\bar{U}^T \bar{U} - n\Sigma_v)^{-1} A^T [A(\bar{U}^T \bar{U} - n\Sigma_v)^{-1} A^T]^{-1} (A\hat{\beta}_I - b) \\ &= \hat{\beta}_I - \beta - (\bar{U}^T \bar{U} - n\Sigma_v)^{-1} A^T [A(\bar{U}^T \bar{U} - n\Sigma_v)^{-1} A^T]^{-1} A(\hat{\beta}_I - \beta) \\ &= \{I_p - (\bar{U}^T \bar{U} - n\Sigma_v)^{-1} A^T [A(\bar{U}^T \bar{U} - n\Sigma_v)^{-1} A^T]^{-1} A\}(\hat{\beta}_I - \beta), \end{aligned}$$

记 $D_{21} = (\bar{U}^T \bar{U} - n\Sigma_v)^{-1} A^T [A(\bar{U}^T \bar{U} - n\Sigma_v)^{-1} A^T]^{-1} A$, 则

$$\sqrt{n}(\hat{\beta}_I^R - \beta) = (I_p - D_{21})\sqrt{n}(\hat{\beta}_I - \beta),$$

则根据大数定律与引理 3.2 可知,

$$(I_p - D_{21}) \xrightarrow{p} I_p - \Sigma_4^{-1} A^T (A \Sigma_4^{-1} A^T)^{-1} A.$$

记 $D_2 = \Sigma_4^{-1} A^T (A \Sigma_4^{-1} A^T)^{-1} A$, 根据定理 3.2 的结论 $\sqrt{n}(\hat{\beta}_I - \beta) \xrightarrow{d} N(0, \Sigma_4^{-1} \Gamma \Sigma_4^{-1})$,

则 $\sqrt{n}(\hat{\beta}_I^R - \beta) \xrightarrow{d} N(0, (I_p - D_2) \Sigma_4^{-1} \Gamma \Sigma_4^{-1} (I_p - D_2)^T)$, 证毕。

3.7 本章小结

本章主要研究了部分线性模型在响应变量缺失, 协变量存在测量误差且参数分量满足线性约束条件下的统计推断, 分别利用完整观测数据和借补技术得到了参数 β 的约束纠偏估计量 $\hat{\beta}_r$ 和 $\hat{\beta}_I^R$, 并证明了其渐近正态性质。由于非参数函数部分存在测量误差, 通过使用反卷积的方法分别在完整观测数据和借补技术下得到了非参数分量的估计, 并在适当的条件下证明了其一致性。然后进行数值模拟,

考虑在不同的缺失概率, 测量误差, 样本容量下比较 $\hat{\beta}_r$ 和 $\hat{\beta}_I^R$ 的标准差与均方误差, 可知借补约束纠偏的估计量 $\hat{\beta}_I^R$ 要优于完整观测数据下的约束纠偏估计量 $\hat{\beta}_r$ 。最后给出实际数据分析, 比较好地解释了数值模拟中的相关结论。本章的研究进一步发展了复杂数据下部分线性模型在参数分量满足线性约束条件下的统计推断理论, 为复杂数据的统计建模提供了一种高效可行的方法。

第4章 响应变量缺失下高维半参数变系数测量误差模型的 惩罚经验似然

本章主要研究含有测量误差的高维半参数变系数模型的参数估计问题。首先基于逆概率加权方法给出了非参数部分的局部纠偏估计,然后将经验似然方法与惩罚函数相结合,提出参数部分的惩罚经验似然估计,并在适当的条件下证明其具有Oracle特征。最后通过数值模拟与实际例子验证所提方法的有效性。

4.1 模型框架

考虑参数部分和非参数部分均含有测量误差的半参数变系数模型:

$$\begin{cases} Y = X^T \beta + Z^T g(U) + \varepsilon, \\ W = X + v, \\ \xi = Z + \eta, \end{cases} \quad (4.1)$$

其中 Y 是响应变量, $X \in R^p$, $Z \in R^q$, U 是协变量, $\beta = (\beta_1, \beta_2, \dots, \beta_p)^T$ 是 p 维的未知参数向量, $g(\cdot) = (g_1(\cdot), g_2(\cdot), \dots, g_q(\cdot))^T$ 是 q 维的未知函数向量, ε 是均值为0的随机误差项。其中测量误差 v , η 均值为0, 且假设与 X , Z , U 独立。Cui^[49]构建了部分线性测量误差模型中参数估计并研究了其渐近性质。Hu等^[50]和Li等^[51]分别研究参数含有测量误差和参数的维数发散时的半参数变系数模型的经验似然。Sun和Sun^[52]和陈盼盼等^[53]利用逆概率加权方法分别研究在协变量缺失下变系数模型和半参数变系数模型的参数估计。Xv等^[54]考虑了协变量随机缺失的半参数变系数测量误差模型的估计, 研究了未知参数的经验对数似然比统计量的渐近分布。

由于数据采集和存储技术的发展, 出现大量且复杂的数据, 例如多媒体数据。其维数会随着样本容量的增加而增加, 因此对高维半参数模型的研究受到了广泛关注。Fan和Li^[55]提出了一种通过非凹惩罚最小二乘法进行变量选择。Tang和Leng^[56]首次提出惩罚经验似然方法(PEL)。PEL方法保留经验似然的优点, 如仅

使用数据确定置信区域的形状和方向并且不需要估计变量的方差。Leng和Tang^[57]将PEL方法应用于一般方程的参数估计和变量选择。Fan等^[58]研究了非参数部分协变量含有测量误差的高维半参数变系数模型的参数估计问题。本章在此基础上考虑响应变量随机缺失的情况，基于逆概率加权方法纠偏构造参数部分和非参数部分的估计，然后根据参数部分纠偏的随机辅助向量构造经验对数似然比函数。同时利用PEL进行变量选择，在适当的条件下证明所提出的PEL估计具有Oracle特征且在零假设下服从渐近卡方分布。

4.2 估计方法与主要结论

4.2.1 修正的经验似然

假设 $\{Y_i, W_i, \xi_i, U_i\}_{i=1}^n$ 是来自以下半参数变系数模型的随机样本

$$\begin{cases} Y_i = X_i^T \beta + Z_i^T g(U_i) + \varepsilon_i, \\ W_i = X_i + v_i, \\ \xi_i = Z_i + \eta_i, \end{cases} \quad (4.2)$$

其中 β 为 p 维未知参数， ε_i 为均值为 0 的随机误差。由于 X_i ， Z_i 具有测量误差，只能观测到其替代变量 W_i ， ξ_i ，测量误差 v_i ， η_i 与 $(X_i, Z_i, U_i, \varepsilon_i)$ 独立并且彼此之间也相互独立，其均值均为 0，协方差矩阵分别为 Σ_v ， Σ_η 。考虑 Yi 等^[59]提出的有测量误差情形下 Y_i 的随机缺失机制：

$$P(\delta_i = 1 | Y_i, U_i, W_i, \xi_i) = P(\delta_i = 1 | U_i, W_i, \xi_i) = \pi(U_i, W_i, \xi_i).$$

上式表明给定 U_i, W_i, ξ_i ，则 δ_i 和 Y_i 条件独立。其中称 $\pi(U_i, W_i, \xi_i)$ 为概率函数。概率函数在实际情况下一般是未知的，采用参数方法进行估计。即选取概率函数：

$$\pi(U_i, W_i, \xi_i, \theta) = \frac{\exp(\theta_0 + \theta_1 U_i + \theta_2^T W_i + \theta_3^T \xi_i)}{1 + \exp(\theta_0 + \theta_1 U_i + \theta_2^T W_i + \theta_3^T \xi_i)},$$

其中 $\theta = (\theta_0, \theta_1, \theta_2^T, \theta_3^T)^T$ 是未知参数向量，可以通过极大似然估计方法获得 θ 的极大似然估计量 $\hat{\theta}$ 。令 $V_i = (U_i, W_i^T, \xi_i^T)^T$ ，则选取概率函数的 $\pi(V_i, \theta)$ 的估计为 $\hat{\pi}_i = \pi(V_i, \hat{\theta})$ 。

如果 β 是已知的, 那么模型(4.2)的第一个方程可以写成如下形式:

$$Y_i - \sum_{j=1}^p X_{ij} \beta_j = \sum_{j=1}^q Z_{ij} g_j(U_i) + \varepsilon_i, \quad i=1, 2, \dots, n. \quad (4.3)$$

式(4.3)视为变系数模型, 采用局部线性估计方法对变系数函数 $g_j(\cdot), j=1, 2, \dots, q$ 进行估计。

若变系数函数 $g_j(\cdot)$ 具有连续的二阶导数, 任意给定 U 邻域内一点 u , 有

$$g_j(U_i) \approx g_j(u) + g'_j(u)(U_i - u) \equiv a_j + b_j(U_i - u), \quad j=1, 2, \dots, q,$$

其中 $g'_j(u) = \partial g_j(u) / \partial u$ 。由于响应变量 Y_i 是缺失的, 利用逆概率加权的方法, 系数函数可以通过极小化下式估计

$$\sum_{i=1}^n \frac{\delta_i}{\pi(V_i, \hat{\theta})} \left[Y_i - \sum_{j=1}^p X_{ij} \beta_j - \sum_{j=1}^q \{a_j + b_j(U_i - u)\} Z_{ij} \right]^2 K_h(U_i - u), \quad (4.4)$$

其中 $K_h(\cdot) = K(\cdot/h)/h$ 是核函数, h 称为窗宽。

为了书写简便, 引入以下记号

$$D_u^Z = \begin{pmatrix} Z_1^T & \frac{U_1 - u}{h} Z_1^T \\ \vdots & \vdots \\ Z_n^T & \frac{U_n - u}{h} Z_n^T \end{pmatrix},$$

$$Y = (Y_1, Y_2, \dots, Y_n)^T, X = (X_1, X_2, \dots, X_n)^T, \hat{\Delta} = \text{diag}(\delta_1 / \hat{\pi}_1, \delta_2 / \hat{\pi}_2, \dots, \delta_n / \hat{\pi}_n),$$

$$M = (Z_1^T g(U_1), Z_2^T g(U_2), \dots, Z_n^T g(U_n))^T, \omega_u = \text{diag}(K_h(U_1 - u), \dots, K_h(U_n - u)).$$

式(4.4)的解为

$$\{\hat{a}_1(u), \dots, \hat{a}_q(u), h\hat{b}_1(u), \dots, h\hat{b}_q(u)\}^T = ((D_u^Z)^T \omega_u \hat{\Delta} D_u^Z)^{-1} (D_u^Z)^T \omega_u \hat{\Delta} (Y - X \beta).$$

因此, 得到变系数函数的估计

$$g^*(u, \beta) = (I_q \ 0_q) ((D_u^Z)^T \omega_u \hat{\Delta} D_u^Z)^{-1} (D_u^Z)^T \omega_u \hat{\Delta} (Y - X \beta), \quad (4.5)$$

其中 I_q 和 0_q 分别为 $q \times q$ 的单位矩阵和零矩阵。

然而, X_i, Z_i 是不能直接观测到的, 观测到的是其替代变量 W_i 和 ξ_i 。为了消

除测量误差造成的影响，对式(4.5)进行局部纠偏，可得

$$\hat{g}(u, \beta) = (I_q \ 0_q)((D_u^\xi)^T \omega_u \hat{\Delta} D_u^\xi - \Omega)^{-1} (D_u^\xi)^T \omega_u \hat{\Delta} (Y - W\beta), \quad (4.6)$$

其中 $W = (W_1, W_2, \dots, W_n)^T$ ，用 ξ_i 代替 Z_i 得到与 D_u^Z 具有相同形式的 D_u^ξ ，

$$\Omega = \sum_{i=1}^n \frac{\delta_i}{\pi(V_i, \hat{\theta})} \Sigma_\eta \otimes \begin{pmatrix} 1 & \frac{U_i - u}{h} \\ \frac{U_i - u}{h} & \left(\frac{U_i - u}{h}\right)^2 \end{pmatrix} K_h(U_i - u),$$

$\otimes$ 表示 Kronecker 乘积。

同时引入如下记号

$$Q = (Q_1^T, Q_2^T, \dots, Q_n^T)^T, Q_i = (I_q \ 0_q)((D_{U_i}^\xi)^T \omega_{U_i} \hat{\Delta} D_{U_i}^\xi - \Omega)^{-1} (D_{U_i}^\xi)^T \omega_{U_i} \hat{\Delta},$$

$$S = (Q_1^T \xi_1, Q_2^T \xi_2, \dots, Q_n^T \xi_n), \tilde{Y} = (I - S)Y, \tilde{X} = (I - S)X,$$

$$\tilde{W} = (I - S)W, \tilde{\xi} = (I - S)\xi.$$

因此，可以通过最小化下式得到 β 的 Profile 最小二乘估计

$$\hat{\beta}_p = \arg \min_{\beta} \sum_{i=1}^n \frac{\delta_i}{\pi(V_i, \hat{\theta})} [\{Y_i - W_i^T \beta - \xi^T \hat{g}(U_i)\}^2 - \hat{g}^T(U_i) \Sigma_\eta \hat{g}(U_i) - \beta^T \Sigma_v \beta]. \quad (4.7)$$

由于 $\hat{\beta}_p$ 是下式的解

$$\sum_{i=1}^n \frac{\delta_i}{\pi(V_i, \hat{\theta})} \{\tilde{W}_i(\tilde{Y}_i - \tilde{W}_i^T \beta) - W^T Q_i^T \Sigma_\eta Q_i (Y - W\beta) + \Sigma_v \beta\} = 0, \quad (4.8)$$

因此，引入如下的关于 β 的辅助向量

$$\varphi_i(\beta) = \frac{\delta_i}{\pi(V_i, \hat{\theta})} \{\tilde{W}_i(\tilde{Y}_i - \tilde{W}_i^T \beta) - W^T Q_i^T \Sigma_\eta Q_i (Y - W\beta) + \Sigma_v \beta\}, \quad (4.9)$$

定义 β 的经验对数似然比函数如下

$$L(\beta) = -\max \left\{ \sum_{i=1}^n \log(np_i) \mid p_i \geq 0, \sum_{i=1}^n p_i = 1, \sum_{i=1}^n p_i \varphi_i(\beta) = 0 \right\}.$$

由 Lagrange 乘数法， $L(\beta)$ 可以表示为

$$L(\beta) = \sum_{i=1}^n \log\{1 + \lambda^T \varphi_i(\beta)\}, \quad (4.10)$$

其中 λ 是下面方程的解

$$\frac{1}{n} \sum_{i=1}^n \frac{\varphi_i(\beta)}{1 + \lambda^T \varphi_i(\beta)} = 0. \quad (4.11)$$

4.2.2 惩罚经验似然

为了达到变量选择的目的，在经验对数似然比函数的基础上添加惩罚函数，得到

$$L_p(\beta) = \sum_{i=1}^n \log\{1 + \lambda^T \varphi_i(\beta)\} + n \sum_{j=1}^p p_\tau(|\beta_j|), \quad (4.12)$$

其中 $p_\tau(\cdot)$ 是惩罚函数，惩罚函数的一阶导数满足

$$p'_\tau(t) = \tau \left\{ I(t \leq \tau) + \frac{(a\tau - t)_+}{(a-1)\tau} I(t > \tau) \right\}, \quad (4.13)$$

其中 $t > 0$, $a = 3.7$ 且 $\tau > 0$ 是调整参数。

假设 $\mathcal{H} = \{j: \beta_{0j} \neq 0\}$ 代表由 β_0 的非零元素组成的集合且 $d = |\mathcal{H}|$ 代表 $\mathcal{H}$ 的维度。不失一般性，记 $\beta = (\beta_1^T, \beta_2^T)^T$ ，其中 $\beta_1 \in R^d$, $\beta_2 \in R^{p-d}$ 代表非零元素与零元素，即 $\beta_0 = (\beta_{10}^T, \beta_{20}^T) = (\beta_{10}^T, 0^T)^T$ 。类似地，记 $\hat{\beta} = (\hat{\beta}_1^T, \hat{\beta}_2^T)^T$ ，其中 $\hat{\beta}_1$, $\hat{\beta}_2$ 分别是 β_1 , β_2 的惩罚经验似然估计值。

进一步假设如下

$$\begin{aligned} \Sigma_1 &= E(XX^T) - E(\Phi^T(U)\Gamma^{-1}(U)\Phi(U)), \\ \Sigma_2 &= E\{\pi^{-1}(V)[\varepsilon - \eta^T g(U) - v^T \beta]^2\} \Sigma_1 \\ &\quad + E\{\pi^{-1}(V)(\varepsilon - v^T \beta)^2 E[\Phi^T(U)\Gamma^{-1}(U)\Sigma_\eta \Gamma^{-1}(U)\Phi(U)]\} \\ &\quad + E\{\pi^{-1}(V)[\Phi^T(U)\Gamma^{-1}(U)(\eta\eta^T - \Sigma_\eta)g(U)]^{\otimes 2}\} \\ &\quad + E\{\pi^{-1}(V)(\varepsilon - \eta^T g(U))^2\} \Sigma_v \\ &\quad + E\{\pi^{-1}(V)(v v^T - \Sigma_v)\beta\beta^T(v v^T - \Sigma_v)\}, \end{aligned}$$

其中 $A^{\otimes 2} = AA^T$, $\Gamma(U) = E(ZZ^T | U)$, $\Phi(U) = E(ZX^T | U)$, $\Sigma = \Sigma_1^{-1}\Sigma_2\Sigma_1^{-1}$ 。对应地，

将 Σ 划分为矩阵块 $\Sigma_{ij} (i=1, 2; j=1, 2)$ 。

在给出定理之前，罗列了一些正则条件。记 $a_n = (p/n)^{1/2}$ ， $c_n = (\log n / (nh))^{1/2} + h^2$ ， $D_n = \{\beta : \|\beta - \beta_0\| \leq Ca_n\}$ ， C 表示正常数。

(C1) 随机变量 U 具有有界支撑 Q ，它的密度函数 $f(u)$ 是 Lipschitz 连续，并且在其支撑集 Q 上有下界且远离 0。

(C2) 对于任何 $u \in Q$ ，矩阵 $\Gamma(u) = E(ZZ^T | U = u)$ 和 $E(XX^T | U = u)$ 是非奇异的， $\Gamma(u)$ ， $\Gamma^{-1}(u)$ ， $E(XX^T | U = u)$ 和 $\Phi(u) = E(ZX^T | U = u)$ 均是 Lipschitz 连续的。

(C3) 存在正数 $r \geq 4$ ，使得 $E\|X\|^r < \infty$ ， $E\|Z\|^r < \infty$ ， $E\|\varepsilon\|^r < \infty$ ， $E\|v\|^r < \infty$ 和 $E\|\eta\|^r < \infty$ 。

(C4) 对于任何 $u \in Q$ ， $\{g_j(u), j=1, 2, \dots, q\}$ 具有连续的二阶导数。

(C5) 核函数 $K(\cdot)$ 是具有紧支撑的对称密度函数，Lipschitz 连续，且存在 $\mathbb{D} < 2 - r^{-1}$ ，且窗宽 h 满足： $nh^6 \rightarrow 0$ ， $nh^3 / (\log n)^3 \rightarrow \infty$ 和 $n^{2\mathbb{D}-1}h \rightarrow \infty$ 。

(C6) Σ_1 和 Σ_2 为正定矩阵，且其特征值均有界。

(C7) 选取概率 $\pi(V, \theta)$ 关于 θ 二阶可导且存在常数 C_0 ，使 $\inf\{\pi(V, \theta)\} \geq C_0 > 0$ 。

(C8) $\{\varepsilon_i, i=1, 2, \dots, n\}$ 为 α -混合序列，存在正数 $\gamma > 2$ ，有混合系数 $\alpha(n) = O(n^{-\gamma})$ 。

(C9) $\phi = [X - \Phi^T(U)\Gamma^{-1}(U)Z][\varepsilon - \eta^T g(U) - v^T \beta]$ 且 ϕ_{nj} 是 ϕ_n 的第 j 个变量， $j=1, 2, \dots, p$ 。对于条件 C3 中的 r ，存在正常数 c ，当 $n \rightarrow \infty$ ，有

$$\max_{1 \leq \mathbb{L}_1, \mathbb{L}_2, \mathbb{L}_3 \leq p} E(\phi_{n\mathbb{L}_1} \phi_{n\mathbb{L}_2} \phi_{n\mathbb{L}_3})^2 < \infty, E(\|\phi_n(\beta) / \sqrt{p}\|^r) < c, E(\|XZ^T / \sqrt{p}\|^r) < c,$$

$$E(\|\Phi^T(U)\Gamma^{-1}(U)ZZ^T / \sqrt{p}\|^r) < c, E(\|\Phi^T(U)\Gamma^{-1}(U)(\eta\eta^T - \Sigma_\eta) / \sqrt{p}\|^r) < c,$$

$$\frac{1}{p} \sum_{\mathbb{L}=1}^p E\left\{\|\phi_{n\mathbb{L}}\| \left(\|XZ^T / \sqrt{p}\|^4 + \|\Phi^T(U)\Gamma^{-1}(U)ZZ^T / \sqrt{p}\|^4\right)\right\} < c.$$

(C10) 当 $n \rightarrow \infty$, $\tau(n/p)^{1/2} \rightarrow \infty$ 和 $\min_{j \in \mathcal{H}} \beta_{0j} / \tau \rightarrow \infty$ 。

(C11) 惩罚函数 $p_\tau(\cdot)$ 满足如下

$$\max_{j \in \mathcal{H}} p'_\tau(|\beta_{0j}|) = o((np)^{-1/2}), \max_{j \in \mathcal{H}} p''_\tau(|\beta_{0j}|) = o(p^{-1/2}).$$

定理 4.1 假设条件(C1)-(C7)成立, 则

$$\sqrt{nh} \left\{ \hat{g}(u) - g(u) - \frac{1}{2} h^2 \frac{\mu_2^2 - \mu_1 \mu_3}{\mu_2 - \mu_1^2} g''(u) \right\} \xrightarrow{d} N(0, \nabla),$$

其中 $\xrightarrow{d}$ 表示依分布收敛, 且

$$\nabla = \frac{\mu_2^2 e_0 - 2\mu_1 \mu_2 e_1 + \mu_1^2 e_2}{(\mu_2 - \mu_1^2)^2} f^{-1}(u) \Sigma^*,$$

$$\Sigma^* = \Gamma^{-1}(u) [E\{(\varepsilon - v^T \beta)^2\} \Gamma(u) + E\{(\varepsilon - v^T \beta)^2\} \Sigma_\eta + E\{\zeta g(u) g^T(u) \zeta^T\} \Gamma^{-1}(u)],$$

$$\zeta = \Sigma_\eta - \eta \eta^T - Z \eta^T, \mu_k = \int_{-\infty}^{+\infty} u^k K(u) du, e_k = \int_{-\infty}^{+\infty} u^k K^2(u) du, k = 0, 1, 2, 3, 4.$$

由定理 4.1, 可以得到非参数部分的渐近偏差和渐近方差, 获得非参数部分估计的渐近正态性。

定理 4.2 假设条件(C1)-(C11)成立, 且 $p^5/n \rightarrow 0$, 当 $n \rightarrow \infty$ 时, 有

(i) 稀疏性: $P(\hat{\beta}_2 = 0) \rightarrow 1$ 。

(ii) 渐近正态性: $\sqrt{n} \mathcal{A}_n \Sigma_d^{-1/2} (\hat{\beta}_1 - \beta_{10}) \xrightarrow{d} N(0, \mathcal{B})$, 其中 $\mathcal{A}_n$ 是 $s \times d$ 矩阵且

$$\mathcal{A}_n \mathcal{A}_n^T \rightarrow \mathcal{B}, \Sigma_d = \Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21}.$$

定理 4.2 说明所提估计量具有 Oracle 特征, 检验假设:

$$H_0: M_n \beta_{10} = 0, H_1: M_n \beta_{10} \neq 0,$$

其中 M_n 是 $s \times d$ 矩阵且对于固定的 s 有 $M_n M_n^T = I_s$ 。构造惩罚经验似然比统计量如下

$$\mathcal{L}_n = -2 \left\{ L_p(\hat{\beta}) - \min_{\beta, M_n \beta_1 = 0} L_p(\beta) \right\}.$$

定理 4.3 假设条件(C1)-(C11)和零假设 H_0 成立, 且 $p^5/n \rightarrow 0$, 则当 $n \rightarrow \infty$ 时, 有

$$\mathcal{L}_n \xrightarrow{d} \chi_s^2,$$

其中 χ_s^2 表示自由度为 s 的卡方分布。根据定理 4.3，对任意给定的 $0 < \alpha_1 < 1$ ，令 $\chi_s^2(1-\alpha_1)$ 为 χ_s^2 的 $(1-\alpha_1)$ 的分位数。 β 的惩罚经验似然置信域构造如下

$$I_{\alpha_1}(\beta) = \{\beta \mid \mathcal{L}_n \leq \chi_s^2(1-\alpha_1)\}.$$

4.2.3 算法

考虑到在经验损失上，目标函数具有非二次性质，因此通过采用迭代算法使目标函数最小化。即在求解 λ 和 β 之间进行迭代。结合 Fan 和 Li^[55] 提出的局部二次逼近方法，对于每一个 $j=1, 2, \dots, p$ ，在给定非零 β_{j_0} 的邻域中， $p_\tau(\beta_j)$ 可近似为

$$p_\tau(\beta_j) \approx p_\tau(\beta_{j_0}^{(k)}) + \frac{1}{2} \{p'_\tau(\beta_{j_0}^{(k)}) / \beta_{j_0}^{(k)}\} \{\beta_j^2 - (\beta_{j_0}^{(k)})^2\},$$

其中 $\beta_{j_0}^{(k)}$ 是 β_{j_0} 的第 k 步估计。然后应用 Owen^[39] 中讨论的算法，通过非线性优化得到(4.12)的最小值，重复这个过程直到收敛。根据 Wang 等^[60] 使用 BIC 类型标准来选择调整参数 τ ，该参数定义为

$$BIC(\tau) = -2L_p(\beta_\tau) + \log(n)df_\tau,$$

其中 β_τ 是调整参数为 τ 时的 β 的估计值， df_τ 是 β_τ 中非零系数的个数。

4.3 数值模拟

本节主要通过数值模拟来研究所提出的方法在有限样本下的表现，模拟过程中的数据由如下模型产生：

$$\begin{cases} Y_i = X_i^T \beta + Z_i^T g(U_i) + \varepsilon_i, \\ W_i = X_i + v_i, \\ \xi_i = Z_i + \eta_i, \end{cases}$$

其中非参数函数 $g(U) = (g_1(U), g_2(U))^T$ ， $U = (1/n, 2/n, \dots, 1)^T$ ， $g_1(U) = \sin(\pi U)$ ， $g_2(U) = 2 + \cos(2\pi U)$ 。参数分量 $\beta = (\beta_1, \beta_2, \beta_3, \beta_4, \dots, \beta_p)^T = (1, 0, 2, 0, \dots, 0)^T$ 。协变量 $X_i \sim N(\mu_x, \Sigma_x)$ ， $\mu_x = (2, 2, \dots, 2)^T$ ， $\Sigma_x = (\rho_{ij})_{p \times p}$ ，其中 $\rho_{ij} = \rho^{|i-j|}$ 且 $\rho = 0.5$ 。

$Z_i \sim N(\mu_z, \Sigma_z)$, $\mu_z = (2, 2)^T$, $\Sigma_z = ((1, 0.5)^T, (1, 0.5)^T)^T$, 测量误差 $v \sim N(0, \Sigma_v)$, $\eta \sim N(0, \Sigma_\eta)$ 。这里取两种测量误差, 即 $\Sigma_v = 0.1I_p$, $\Sigma_\eta = 0.1I_2$, 和 $\Sigma_v = 0.2I_p$, $\Sigma_\eta = 0.2I_2$ 。其中 I_p 是 $p \times p$ 的单位矩阵。 ε_i 由如下的 AR(1) 模型生成, $\varepsilon_i = 0.5\varepsilon_{i-1} + \kappa_i$, 其中 $\kappa_i \sim N(0, 1)$, $\varepsilon_0 \sim N(0, 1)$ 。在模拟中核函数选用高斯核函数: $K(t) = (1/\sqrt{2\pi})\exp(-t^2/2)$, 窗宽的选择采用最小化交叉验证的方法, 即最小化 $CV(h) = n^{-1} \sum_{i=1}^n (Y_i - W_i^T \hat{\beta}_i - \xi_i^T \hat{g}(U_i))^2$, 其中 $\hat{\beta}_i, \hat{g}(U_i)$ 分别是 $\beta_i, g(U_i)$ 的估计值。

为了比较不同缺失率对模型的影响, 利用 Logistic 回归模型选取缺失机制。即 $\pi(U_i, W_i, \xi_i, \theta) = 1 / \{1 + \exp[-(\theta_0 + \theta_1 U_i + \theta_2^T W_i + \theta_3^T \xi_i)]\}$, $\theta = (\theta_0, \theta_1, \theta_2^T, \theta_3^T)^T$, 其中选取的 $\theta_0 = 0$, $\theta_1 = 0.5$, $\theta_2 = (\theta_{21}, \theta_{22}, \dots, \theta_{2p})^T = (0.3, 0.3, 0, 0, \dots, 0)^T$, $\theta_3 = (0.3, 0.2)^T$ 和 $\theta_0 = 0$, $\theta_1 = 0.1$, $\theta_2 = (\theta_{21}, \theta_{22}, \dots, \theta_{2p})^T = (0.15, 0.2, 0, 0, \dots, 0)^T$, $\theta_3 = (0.05, 0.02)^T$, 对应的平均缺失率大约为 0.1 和 0.3。

下面针对模型中的参数部分进行模拟研究。在模拟的过程中, 维数 p 取 $10(3n)^{1/5.1} - 20$ 的整数部分, 样本容量分别取 300 和 500, 然后维数 p 补充取 30。考虑在不同的缺失概率, 测量误差下, 利用惩罚经验似然方法给出估计的零系数平均值与估计值 $\hat{\beta}_1$ 的均方误差, 在每个样本容量下, 重复模拟 500 次, 模拟结果如表 4-1 所示:

表 4-1 不同情况下估计的零系数平均值与估计值 $\hat{\beta}_1$ 的均方误差

缺失率	n	p	测量误差	零系数平均值	均方误差
0.3	300	18	$0.1I_p$	16.32	0.4105
	300	18	$0.2I_p$	16.51	0.4545
	300	30	$0.1I_p$	28.35	0.5153
	300	30	$0.2I_p$	28.57	0.5232
	500	22	$0.1I_p$	20.17	0.2571
	500	22	$0.2I_p$	19.73	0.2707
	500	30	$0.1I_p$	27.75	0.3173
	500	30	$0.2I_p$	27.76	0.3326

续表 4-1 不同情况下估计的零系数平均值与估计值 $\hat{\beta}_1$ 的均方误差

缺失率	n	p	测量误差	零系数平均值	均方误差
0.1	300	18	$0.1I_p$	15.77	0.1685
	300	18	$0.2I_p$	15.79	0.1983
	300	30	$0.1I_p$	27.76	0.2018
	300	30	$0.2I_p$	27.86	0.2337
	500	22	$0.1I_p$	20.10	0.0461
	500	22	$0.2I_p$	19.85	0.0733
	500	30	$0.1I_p$	28.05	0.0981
	500	30	$0.2I_p$	28.14	0.1106

从表 4-1 中可以看出

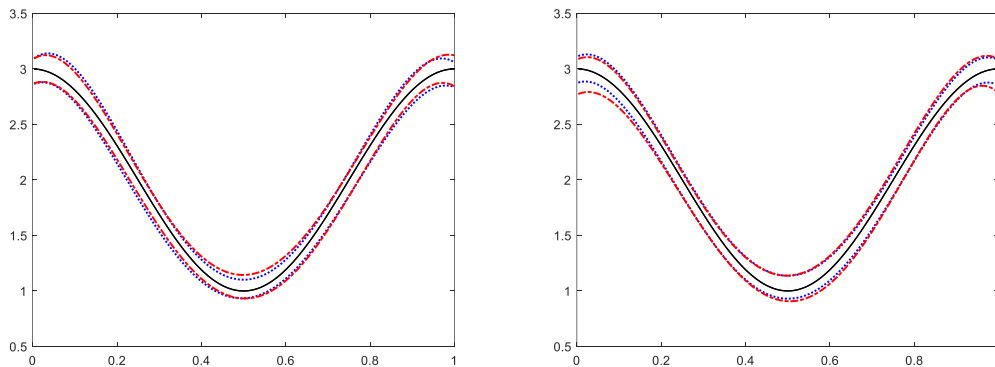
(1) 在缺失率、样本容量、维数 p 相同的情况下，测量误差越小，估计的零系数平均值更接近真实的 $p-2$ 值，均方误差越小。

(2) 在样本容量、维数 p 、测量误差相同的情况下，缺失率的增加会使估计的零系数平均值更偏离真实的 $p-2$ 值，均方误差越大。

(3) 在缺失率、样本容量、测量误差相同的情况下，随着维数 p 的减少，惩罚经验似然估计值的均方误差在降低。

这说明惩罚经验似然方法可以选择几乎所有的两个真实预测因子，估计的零系数的平均值接近 $p-2$ ，所选模非常接近真实模型。

下面给出非参数部分 $g_1(U)$ 和 $g_2(U)$ 在置信水平为 95%，样本容量 $n=300$ ， $p=30$ ，缺失率分别为 0.1 和 0.3，测量误差分别为 $0.1I_p$ 和 $0.2I_p$ 下的置信带。



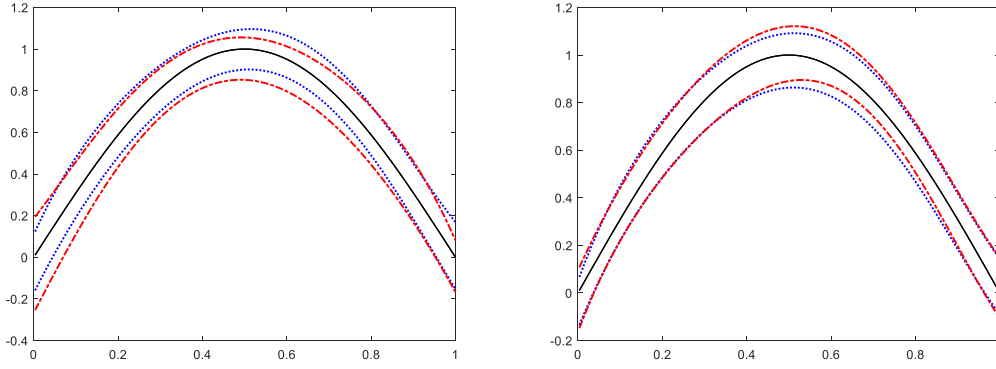


图 4-1 置信水平为 95% 非参数部分 $g_1(U)$ 和 $g_2(U)$ 的置信带，其中样本容量为 $n = 300$ ，缺失率为 0.1(左)及 0.3(右)，测量误差 $0.1I_p$ (点线) 以及 $0.2I_p$ (点划线)

从图 4-1 可以看出，在样本容量和缺失率相同的情况下，测量误差越小，置信带越窄，测量误差越大，置信带越宽；在样本容量和测量误差相同的情况下，缺失率越小，置信带越窄，缺失率越大，置信带越宽。

4.4 实际例子

在本节中将惩罚经验似然方法与 Profile 最小二乘方法对 1994 年 1 月至 1997 年 12 月期间香港因呼吸系统疾病住院人数与污染物和一些环境因素的每日平均测量值的比较分析。这个数据集由 $n = 1451$ 个观测值组成，研究响应变量呼吸系统疾病住院人数 (RESP) 和协变量二氧化氮浓度 (NO_2) (克/立方米)，二氧化硫浓度 (SO_2) (克/立方米)，可吸入悬浮颗粒物浓度 (RSP) (克/立方米)，臭氧浓度 (O_3) (克/立方米)，温度 (TEMP) 以及湿度 (HUM) 之间的关系。为了更清晰地表示响应变量和各个协变量之间的关系，本节将响应变量 RESP 用 Y 表示，协变量 NO_2 , SO_2 , RSP, O_3 , TEMP, HUM 分别用 X_1, X_2, X_3, X_4, Z_1 和 Z_2 表示。令 $U = (1/n, 2/n, \dots, 1)^T$ ，并假设 X_1, Z_1 带有测量误差，即半参数变系数模型如下：

$$\begin{cases} Y = X_1\beta_1 + X_2\beta_2 + X_3\beta_3 + X_4\beta_4 + Z_1g_1(U) + Z_2g_2(U) + \varepsilon, \\ W_1 = X_1 + v_1, \\ \xi_1 = Z_1 + \eta_1, \end{cases}$$

其中 $v_1 \sim N(0, 0.1^2)$, $\eta_1 \sim N(0, 0.1^2)$, ε 由如下的 AR(1) 模型生成, $\varepsilon_i = 0.5\varepsilon_{i-1} + \kappa_i$,

$\kappa_i \sim N(0, 0.1^2)$, $\varepsilon_0 \sim N(0, 1)$ 。选取概率函数如下：

$$\pi(W, X, \xi, Z) = 1 / \{1 + \exp[-(0.007W_1 + 0.03X_2 + 0.01X_3 + 0.01X_4 + 0.01\xi_1 + 0.004Z_2)]\},$$

其对应的平均缺失率大约为 0.1。在实际研究中，核函数与窗宽仍然采用数值模拟部分的高斯核函数与交叉验证得到的窗宽。

采用 Profile 最小二乘方法，得到的参数估计值为 $\hat{\beta}_{o1} = 0.7412$, $\hat{\beta}_{o2} = -0.2068$, $\hat{\beta}_{o3} = 0.3505$, $\hat{\beta}_{o4} = 0.1345$ 。采用惩罚经验似然方法，得到的参数估计值为 $\hat{\beta}_1 = 0.9169$, $\hat{\beta}_2 = 0$, $\hat{\beta}_3 = 0.1688$, $\hat{\beta}_4 = 0.1044$ 。从两种方法的结果均可以看到 NO_2 , RSP , O_3 和 RESP 的相关系数均为正，且 NO_2 在因呼吸系统疾病住院的人数中有着非常显著影响。在 Profile 最小二乘方法得到的结果中 NO_2 与 RESP 的相关系数为负，即在适当的条件下，随着空气中二氧化硫浓度的增加会导致因呼吸系统疾病住院的人数减少。这一结论在实际生活中并不合理，通过惩罚经验似然的方法，进行变量选择得到参数估计值 $\hat{\beta}_2 = 0$ ，并考虑在零假设 $H_0: \beta_2 = 0$ 下，检验 NO_2 是否对 RESP 有显著影响。可得相应的 P 值为 0.3197，因此 NO_2 对 RESP 无显著影响。则根据筛选得到的最终结果可认为在适当的条件下，空气污染物因素之一的二氧化氮浓度，可吸入悬浮颗粒物浓度，臭氧浓度均会对因呼吸系统疾病住院的人数产生正向的影响。

4.5 主要结果的证明

引理 4.1 若条件(C1)-(C6)成立，则当 $n \rightarrow \infty$ 时，

$$\begin{aligned} \sup_{u \in Q} \left| \frac{1}{nh} \sum_{i=1}^n K\left(\frac{U_i - u}{h}\right) K\left(\frac{U_i - u}{h}\right)^k Z_{ij_1} Z_{ij_2} - f(u) \Gamma_{j_1, j_2}(u) \mu_k \right| &= O(c_n) \quad a.s., \\ \sup_{u \in Q} \left| \frac{1}{nh} \sum_{i=1}^n K\left(\frac{U_i - u}{h}\right) K\left(\frac{U_i - u}{h}\right)^k Z_{ij} \varepsilon_i \right| &= O\{(\log n / (nh))^{1/2}\} \quad a.s., \\ \sup_{u \in Q} \left| \frac{1}{nh} \sum_{i=1}^n K\left(\frac{U_i - u}{h}\right) K\left(\frac{U_i - u}{h}\right)^k Z_{ij} \eta_{ij} \right| &= O\{(\log n / (nh))^{1/2}\} \quad a.s., \end{aligned}$$

其中 $j, j_1, j_2 = 1, 2, \dots, q$, $k = 0, 1, \dots, 4$, $\Gamma_{j_1, j_2}(u)$, 是矩阵 $\Gamma(u)$ 的第 (j_1, j_2) 元素。

证明 类似 Xia 和 Li^[61]引理 A2 可以得到本引理的证明。

引理 4.2 若条件(C1)-(C7)成立, 则当 $n \rightarrow \infty$ 时, 有

$$\begin{aligned}(D_u^\varepsilon)^T \omega_u \hat{\Delta}(D_u^\varepsilon) - \Omega &= \sum_{i=1}^n \frac{\delta_i}{\pi(V_i, \hat{\theta})} f(u) \Gamma(u) \otimes \begin{pmatrix} 1 & \mu_1 \\ \mu_1 & \mu_2 \end{pmatrix} \{1 + O_p(c_n)\}, \\ (D_u^\varepsilon)^T \omega_u \hat{\Delta} W &= \sum_{i=1}^n \frac{\delta_i}{\pi(V_i, \hat{\theta})} f(u) \Phi(u) \otimes (1, \mu_1) \{1 + O_p(c_n)\}, \\ (D_u^\varepsilon)^T \omega_u \hat{\Delta} \xi &= \sum_{i=1}^n \frac{\delta_i}{\pi(V_i, \hat{\theta})} f(u) \Gamma(u) \otimes (1, \mu_1) \{1 + O_p(c_n)\}.\end{aligned}$$

证明 类似 Fen 和 Xue^[45]的引理 2 可以得到本引理的证明。

引理 4.3 若条件(C1)-(C7)成立, 当 $n \rightarrow \infty$ 时, 则

$$\frac{1}{n} \sum_{i=1}^n \frac{\delta_i}{\pi(V_i, \hat{\theta})} \{\tilde{W}_i \tilde{W}_i^T - W^T Q_i^T \Sigma_\eta Q_i W - \Sigma_v\} \xrightarrow{P} E[XX^T] - E[\Phi^T(U) \Gamma^{-1}(U) \Phi(U)].$$

证明 通过简单计算可得

$$\begin{aligned}& \frac{1}{n} \sum_{i=1}^n \frac{\delta_i}{\pi(V_i, \hat{\theta})} \{\tilde{W}_i \tilde{W}_i^T - W^T Q_i^T \Sigma_\eta Q_i W - \Sigma_v\} \\&= \frac{1}{n} \sum_{i=1}^n \frac{\delta_i}{\pi(V_i, \hat{\theta})} \{(X_i X_i^T - W^T Q_i^T Z_i X_i^T - X_i Z_i^T Q_i W + W^T Q_i^T Z_i Z_i^T Q_i W) \\& \quad - W^T Q_i^T \eta_i (X_i^T - Z_i^T) Q_i W - (X_i - W^T Q_i^T Z_i) \eta_i^T Q_i W - v_i Z_i^T Q_i W - v_i \eta_i^T Q_i W \\& \quad - W^T Q_i^T Z_i v_i^T - W^T Q_i^T \eta_i v_i^T + W^T Q_i^T (\eta_i \eta_i^T - \Sigma_\eta) Q_i W + (v_i v_i^T - \Sigma_v) + X_i v_i^T + v_i X_i^T\} \\&:= \sum_{i=1}^{14} \mathcal{F}_i.\end{aligned}\tag{4.14}$$

根据 Q_i 的定义, 引理 4.2 以及大数定律, 可得

$$\begin{aligned}\mathcal{F}_1 &= E(XX^T), \quad \mathcal{F}_2 = \mathcal{F}_3 = -E[\Phi^T(U) \Gamma^{-1}(U) \Phi(U)], \\ \mathcal{F}_4 &= E[\Phi^T(U) \Gamma^{-1}(U) \Phi(U)], \quad \mathcal{F}_i = o_p(1), \quad i = 5, 6, \dots, 14.\end{aligned}$$

即证。

引理 4.4 若条件(C1)-(C8)成立, 当 $n \rightarrow \infty$ 时, 则

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n \varphi_i(\beta) \xrightarrow{d} N(0, \Sigma_2).$$

$$\begin{aligned}
 \text{证明} \quad \frac{1}{\sqrt{n}} \sum_{i=1}^n \varphi_i(\beta) &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{\delta_i}{\pi(V_i, \hat{\theta})} \{ \tilde{W}_i (\tilde{Y}_i - \tilde{W}_i^T \beta) + \Sigma_v \beta \} \\
 &\quad - \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{\delta_i}{\pi(V_i, \hat{\theta})} W^T Q_i^T \Sigma_\eta Q_i (Y - W \beta) \} \\
 &:= \mathcal{G}_1 - \mathcal{G}_2.
 \end{aligned} \tag{4.15}$$

记 $\tilde{W}_i = W_i - \hat{W}_i$, $\tilde{v}_i = v_i - \hat{v}_i$, $\tilde{\varepsilon}_i = \varepsilon_i - \hat{\varepsilon}_i$, $\tilde{\xi}_i = \xi_i - \hat{\xi}_i$, $\tilde{\eta}_i = \eta_i - \hat{\eta}_i$, 通过对 $\mathcal{G}_1$ 简单计算, 可得

$$\begin{aligned}
 \mathcal{G}_1 &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{\delta_i}{\pi(V_i, \hat{\theta})} \{ (W_i - \hat{W}_i) [\varepsilon_i - \hat{\varepsilon}_i - (v_i - \hat{v}_i)^T \beta \\
 &\quad + (\xi_i - \hat{\xi}_i)^T g(U_i) - (\eta_i - \hat{\eta}_i)^T g(U_i)] + \Sigma_v \beta \} \\
 &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{\delta_i}{\pi(V_i, \hat{\theta})} \{ [W_i - \Phi^T(U_i) \Gamma^{-1}(U_i) \xi_i] [\varepsilon_i - v_i^T \beta - \eta_i^T g(U_i)] + \Sigma_v \beta \} \\
 &\quad + \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{\delta_i}{\pi(V_i, \hat{\theta})} \{ [\Phi^T(U_i) \Gamma^{-1}(U_i) \xi_i - [((D_{U_i}^\xi)^T \omega_{U_i} \hat{\Delta} D_{U_i}^\xi - \Omega)^{-1} (D_{U_i}^\xi)^T \omega_{U_i} \hat{\Delta} W_i]^T \xi_i] \\
 &\quad \cdot [\varepsilon_i - v_i^T \beta - \eta_i^T g(U_i)] \} \\
 &\quad + \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{\delta_i}{\pi(V_i, \hat{\theta})} \{ (W_i - \hat{W}_i) (\hat{\varepsilon}_i - \hat{v}_i^T \beta) \} \\
 &\quad + \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{\delta_i}{\pi(V_i, \hat{\theta})} \{ (W_i - \hat{W}_i) [\xi_i^T g(U_i) - \hat{\xi}_i^T g(U_i) + \hat{\eta}_i^T g(U_i)] \} \\
 &:= \mathbb{D}_1 + \mathbb{D}_2 + \mathbb{D}_3 + \mathbb{D}_4,
 \end{aligned} \tag{4.16}$$

类似 You 和 Chen^[6]引理 A.4 的证明, 可得 $\mathbb{D}_i = o_p(1)$, $i = 2, 3, 4$, 则

$$\begin{aligned}
 \mathbb{D}_1 &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{\delta_i}{\pi(V_i, \hat{\theta})} \{ [(X_i + v_i) - \Phi^T(U_i) \Gamma^{-1}(U_i) (Z_i + \eta_i)] [\varepsilon_i - v_i^T \beta - \eta_i^T g(U_i)] + \Sigma_v \beta \} \\
 &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{\delta_i}{\pi(V_i, \hat{\theta})} \{ [X_i - \Phi^T(U_i) \Gamma^{-1}(U_i) Z_i] [\varepsilon_i - v_i^T \beta - \eta_i^T g(U_i)] \} \\
 &\quad - \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{\delta_i}{\pi(V_i, \hat{\theta})} \{ \Phi^T(U_i) \Gamma^{-1}(U_i) \eta_i (\varepsilon_i - v_i^T \beta) \}
 \end{aligned}$$

$$\begin{aligned}
 & + \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{\delta_i}{\pi(V_i, \hat{\theta})} \{\Phi^T(U_i) \Gamma^{-1}(U_i) \eta_i \eta_i^T g(U_i)\} \\
 & + \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{\delta_i}{\pi(V_i, \hat{\theta})} \{v_i [\varepsilon_i - \eta_i^T g(U_i)]\} \\
 & - \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{\delta_i}{\pi(V_i, \hat{\theta})} \{(v_i v_i^T - \Sigma_v) \beta\}.
 \end{aligned} \tag{4.17}$$

类似 Feng 和 Xue^[45]引理 4 的证明可得

$$\begin{aligned}
 \mathcal{G}_2 &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{\delta_i}{\pi(V_i, \hat{\theta})} \Phi^T(U_i) \Gamma^{-1}(U_i) \Sigma_\eta g(U_i) \{1 + O_p(c_n)\} \\
 &+ \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{\delta_i}{\pi(V_i, \hat{\theta})} \Phi^T(U_i) \Gamma^{-1}(U_i) \Sigma_\eta \mathbf{1}_q O_p(c_n).
 \end{aligned} \tag{4.18}$$

结合 $\mathcal{G}_1$, $\mathcal{G}_2$ 可得

$$\begin{aligned}
 \frac{1}{\sqrt{n}} \sum_{i=1}^n \varphi_i(\beta) &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{\delta_i}{\pi(V_i, \hat{\theta})} \{[X_i - \Phi^T(U_i) \Gamma^{-1}(U_i) Z_i][\varepsilon_i - \eta_i^T g(U_i) - v_i^T \beta] \\
 &\quad - \Phi^T(U_i) \Gamma^{-1}(U_i) \eta_i (\varepsilon_i - v_i^T \beta) + \Phi^T(U_i) \Gamma^{-1}(U_i) (\eta_i \eta_i^T - \Sigma_\eta) g(U_i) \\
 &\quad + v_i [\varepsilon_i - \eta_i^T g(U_i)] - (v_i v_i^T - \Sigma_v) \beta\} + o_p(1) \\
 &:= \frac{1}{\sqrt{n}} \sum_{i=1}^n J_i.
 \end{aligned} \tag{4.19}$$

因此, 通过证明 $\frac{1}{\sqrt{n}} \sum_{i=1}^n J_i \xrightarrow{d} N(0, \Sigma_2)$ 即可证明出引理 4.4, 鉴于 Cramér-Wold 方

法, 则只需验证 $p \times 1$ 不为 0 的向量 $\mathbb{A}$, 令 $B_i = \mathbb{A}^T J_i$, 有

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n B_i \xrightarrow{P} N(0, \mathbb{A}^T \Sigma_2 \mathbb{A}). \tag{4.20}$$

记 $\mathbb{F} := \mathbb{F}_n, \mathbb{G} := \mathbb{G}_n$ 则 $\mathbb{F}/n \rightarrow 0, \mathbb{G}/\mathbb{F} \rightarrow 0$ 和 $n\alpha(\mathbb{G})/\mathbb{F} \rightarrow 0$ (如 $\mathbb{F} = n^{1/2}, \mathbb{G} = n^{1/4}$).

令 $\mathbb{H} := \mathbb{H}_n = \lfloor n/(\mathbb{F} + \mathbb{G}) \rfloor$. 采用 Bernstein 大块小块方法证明(4.20)。用长度为 $\mathbb{F}$ 的

大块和长度为 $\mathbb{G}$ 的小块将集合 $\{1, 2, \dots, n\}$ 划分为 $2\mathbb{H}+1$ 个子集。定义 G_{mn} , G'_{mn} ,

G''_{mn}

$$G_{mn} = \sum_{i=k_m}^{k_m+\mathbb{F}-1} B_i, G'_{mn} = \sum_{j=l_m}^{l_m+\mathbb{G}-1} B_j, G''_{\mathbb{H}n} = \sum_{k=\mathbb{H}(\mathbb{F}+\mathbb{G})+1}^n B_k,$$

其中 $k_m = (m-1)(\mathbb{F} + \mathbb{G}) + 1$, $l_m = (m-1)(\mathbb{F} + \mathbb{G}) + \mathbb{F} + 1$, $m = 1, 2, \dots, \mathbb{H}$ 。则

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n B_i = \frac{1}{\sqrt{n}} \left\{ \sum_{m=1}^{\mathbb{H}} G_{mn} + \sum_{m=1}^{\mathbb{H}} G'_{mn} + G''_{\mathbb{H}n} \right\} := \frac{1}{\sqrt{n}} \{s'_n + s''_n + s'''_n\}. \quad (4.21)$$

所以要证明(4.20), 即应证明如下

$$n^{-1} E(s''_n)^2 \rightarrow 0, \quad n^{-1} E(s'''_n)^2 \rightarrow 0, \quad (4.22)$$

$$\text{Var}(n^{-1/2} s'_n) \rightarrow \mathbb{A}^T \Sigma_2 \mathbb{A}, \quad (4.23)$$

$$\left| E \exp \left(it \sum_{m=1}^{\mathbb{H}} n^{-1/2} G_{mn} \right) - \prod_{m=1}^{\mathbb{H}} E \exp(itn^{-1/2} G_{mn}) \right| \rightarrow 0, \quad (4.24)$$

$$\frac{1}{n} \sum_{m=1}^{\mathbb{H}} E G_{mn}^2 I(|G_{mn}| > \nu \sqrt{n}) \rightarrow 0, \forall \nu > 0. \quad (4.25)$$

式(4.22)表明 s''_n 和 s'''_n 是渐近可忽略不计的。式(4.24)说明 G_{mn} 是渐近独立的, 式(4.23)和式(4.25)是证明 s'_n 渐近正态性质的 Lindeberg-Feller 条件。对于 $i \neq j$, 记

$E(B_i B_j) = 0$, $\text{Cov}(G'_{in}, G'_{jn}) = 0$ 。由于 $EB_i^2 < \infty$, 则

$$\frac{1}{n} E(s'''_n)^2 = \frac{1}{n} \sum_{m=1}^{\mathbb{H}} \sum_{i=l_m}^{l_m+\mathbb{F}-1} EB_i^2 = O(\mathbb{H}\mathbb{G}/n) = O(\mathbb{G}/\mathbb{F}) = o(1).$$

与之类似

$$\frac{1}{n} E(s''_n)^2 = \frac{1}{n} \sum_{i=\mathbb{H}(\mathbb{F}+\mathbb{G})+1}^n EB_i^2 \leq C(n - \mathbb{H}(\mathbb{F} + \mathbb{G}))/n \leq C\mathbb{F}/n \rightarrow 0,$$

显然

$$\begin{aligned} \frac{1}{n} \text{Var}(s'_n) &= \frac{1}{n} \sum_{m=1}^{\mathbb{H}} \text{Var}(G_{mn}) = n^{-1} \sum_{i=1}^n \text{Var} B_i \\ &\quad - n^{-1} \sum_{m=1}^{\mathbb{H}} \sum_{i=l_m}^{l_m+\mathbb{G}+1} \text{Var} B_i - n^{-1} \sum_{i=\mathbb{H}(\mathbb{F}+\mathbb{G})+1}^n \text{Var} B_i \rightarrow \mathbb{A}^T \Sigma_2 \mathbb{A}. \end{aligned}$$

对于(4.24), 结合 Volkonskii 和 Rozanov^[62]引理 1.1 可得

$$\left| E \exp \left(it \sum_{m=1}^{\mathbb{H}} n^{-1/2} G_{mn} \right) - \prod_{m=1}^{\mathbb{H}} E \exp(itn^{-1/2} G_{mn}) \right| \leq 16\mathbb{H}\alpha(\mathbb{G}+1) \leq Cn\alpha(\mathbb{G})/\mathbb{F} \rightarrow 0.$$

对于(4.25), 由式(4.19)令

$$J_{A1} = \frac{\delta_i}{\pi(V_i, \hat{\theta})} \{ [X_i - \Phi^T(U_i)\Gamma^{-1}(U_i)Z_i] - \Phi^T(U_i)\Gamma^{-1}(U_i)\eta_i + v_i \} \varepsilon_i,$$

$$J_{A2} = \frac{\delta_i}{\pi(V_i, \hat{\theta})} \{ [X_i - \Phi^T(U_i)\Gamma^{-1}(U_i)Z_i] [-\eta_i^T g(U_i) - v_i^T \beta]$$

$$+ \Phi^T(U_i)\Gamma^{-1}(U_i)[\eta_i v_i^T \beta + (\eta_i \eta_i^T - \Sigma_\eta)g(U_i)] - v_i \eta_i^T g(U_i) - (v_i v_i^T - \Sigma_v)\beta \},$$

其中 J_{A1} 为 α -混合。对于任意小的 $\nu > 0$, 当 $\gamma \geq 2(2+\nu)/(2-\nu)$, 且 $0 < \rho_a \leq \nu/2$,

结合 Yang^[63]定理 2.2 可得

$$\begin{aligned} \frac{1}{n} \sum_{m=1}^{\mathbb{H}} EG_{mn}^2 I(|G_{mn}| > \nu \sqrt{n}) &\leq C n^{-1-\nu/2} \sum_{m=1}^{\mathbb{H}} E \left| \sum_{i=k_m}^{k_m+\mathbb{F}-1} B_i \right|^{2+\nu} \\ &\leq C n^{-1-\nu/2} \sum_{m=1}^{\mathbb{H}} \left\{ E \left| \sum_{i=k_m}^{k_m+\mathbb{F}-1} \mathbb{A}^T J_{A1} \right|^{2+\nu} + E \left| \sum_{i=k_m}^{k_m+\mathbb{F}-1} \mathbb{A}^T J_{A2} \right|^{2+\nu} \right\} \\ &\leq C n^{-1-\nu/2} \sum_{m=1}^{\mathbb{H}} \left\{ \mathbb{F}^{\rho_a} \sum_{i=k_m}^{k_m+\mathbb{F}-1} E \left| \mathbb{A}^T J_{A1} \right|^{2+\nu} + \left[\sum_{i=k_m}^{k_m+\mathbb{F}-1} \left\| \mathbb{A}^T J_{A1} \right\|_4^2 \right]^{(2+\nu)/2} \right. \\ &\quad \left. + \sum_{i=k_m}^{k_m+\mathbb{F}-1} E \left| \mathbb{A}^T J_{A2} \right|^{2+\nu} + \left[\sum_{i=k_m}^{k_m+\mathbb{F}-1} E \left| \mathbb{A}^T J_{A2} \right|^2 \right]^{1+\nu/2} \right\} \\ &\leq C(\mathbb{F}/n)^{\nu/2} \rightarrow 0. \end{aligned}$$

证毕。

引理 4.5 若条件(C1)-(C9)成立, 且 $p^{3+2/(r-2)}/n \rightarrow 0$, 当 $n \rightarrow \infty$ 时, 则

$$\frac{\left\{ n^{-1/2} \sum_{i=1}^n \varphi_i^T(\beta) \right\} \Sigma_2^{-1} \left\{ n^{-1/2} \sum_{i=1}^n \varphi_i(\beta) \right\} - p}{\sqrt{2p}} \xrightarrow{d} N(0,1), \quad (4.26)$$

$$\max_{1 \leq i \leq n} \|\varphi_i(\beta)\| = o_p(\sqrt{n/p}). \quad (4.27)$$

证明 利用 Hall 和 Heyde^[64]的鞅中心极限定理, 并结合引理 4.4, 式(4.26)容易得证。这里仅证明式(4.27), 通过简单计算有

$$\begin{aligned} \max_{1 \leq i \leq n} \|\varphi_i(\beta)\| &\leq \left[\max_{1 \leq i \leq n} \|X_i\| + \max_{1 \leq i \leq n} \|\Phi^T(U_i)\Gamma^{-1}(U_i)\xi_i\| + O_p(c_n) \max_{1 \leq i \leq n} \|\xi_i\| \right] \\ &\quad \cdot \left[\max_{1 \leq i \leq n} \|\varepsilon_i - v_i^T \beta\| + O_p(c_n) \max_{1 \leq i \leq n} \|\xi_i\| + O_p(c_n) \max_{1 \leq i \leq n} \|Z_i^T g(U_i)\| \right] \end{aligned}$$

$$\begin{aligned}
 & + \max_{1 \leq i \leq n} \|\eta_i^T g(U_i)\| \{1 + O_p(c_n)\} \\
 & + \max_{1 \leq i \leq n} \|\Phi^T(U_i)\Gamma^{-1}(U_i)\eta_i\| \cdot \max_{1 \leq i \leq n} \|\varepsilon_i - v_i^T \beta\| \\
 & + \max_{1 \leq i \leq n} \|\Phi^T(U_i)\Gamma^{-1}(U_i)\Sigma_\eta g(U_i)\| \{1 + O_p(c_n)\} \\
 & + O_p(c_n) \max_{1 \leq i \leq n} \|\Phi^T(U_i)\Gamma^{-1}(U_i)\Sigma_\eta \mathbf{1}_q\| \\
 & + \max_{1 \leq i \leq n} \|v_i\| \cdot \max_{1 \leq i \leq n} \|\xi_i - \eta_i^T g(U_i)\| \\
 & + C \max_{1 \leq i \leq n} \|v_i v_i^T - \Sigma_v\|.
 \end{aligned}$$

类似 Li 等^[51]中(A.6)和(A.7)的证明, 可知

$$\max_{1 \leq i \leq n} \|\varphi_i(\beta)\| = O_p\left(\sqrt{pn^{1/r}}\right) = o_p\left(\sqrt{n/p}\right).$$

引理 4.6 若定理 4.2 的条件成立, 则 $\|\lambda(\beta)\| = O_p(a_n)$ 。

证明 记 $\lambda = \lambda(\beta) = \rho\Theta$, 其中 $\rho \geq 1$, $\|\Theta\| = 1$ 。令

$$\begin{aligned}
 \bar{\varphi}(\beta) &= \frac{1}{n} \sum_{i=1}^n \varphi_i(\beta), \quad \varphi^*(\beta) = \max_{1 \leq i \leq n} \|\varphi_i(\beta)\|, \quad S(\beta) = \frac{1}{n} \sum_{i=1}^n \varphi_i(\beta) \varphi_i^T(\beta), \\
 0 &= \frac{1}{n} \sum_{i=1}^n \frac{\Theta^T \varphi_i(\beta)}{1 + \rho \Theta^T \varphi_i(\beta)} = \frac{1}{n} \sum_{i=1}^n \Theta^T \varphi_i(\beta) - \rho \frac{1}{n} \sum_{i=1}^n \frac{(\Theta^T \varphi_i(\beta))^2}{1 + \rho \Theta^T \varphi_i(\beta)} \\
 &\leq \Theta^T \bar{\varphi}(\beta) - \frac{\rho}{1 + \rho \varphi^*(\beta)} \Theta^T S(\beta) \Theta.
 \end{aligned}$$

因此, $\rho[\Theta^T S(\beta) \Theta - \varphi^*(\beta) |\Theta^T \bar{\varphi}(\beta)|] \leq |\Theta^T \bar{\varphi}(\beta)|$ 。则

$$\begin{aligned}
 \varphi^*(\beta) &\leq \varphi^*(\beta_0) + \max_{1 \leq i \leq n} \left\| (\tilde{W}_i \tilde{W}_i^T - W^T Q_i^T \Sigma_\eta Q_i W - \Sigma_v)(\beta - \beta_0) \right\| \\
 &\leq \varphi^*(\beta_0) + \max_{1 \leq i \leq n} \left\| (X_i - W^T Q_i^T Z_i)(X_i^T - Z_i^T Q_i W) - W^T Q_i^T \eta_i (X_i^T - Z_i^T) Q_i W \right. \\
 &\quad \left. - (X_i - W^T Q_i^T Z_i) \eta_i^T Q_i W - v_i Z_i^T Q_i W - v_i \eta_i^T Q_i W - W^T Q_i^T Z_i v_i^T - W^T Q_i^T \eta_i v_i^T \right. \\
 &\quad \left. + W^T Q_i^T (\eta_i \eta_i^T - \Sigma_\eta) Q_i W + (v_i v_i^T - \Sigma_v) + X_i v_i^T + v_i X_i^T \right\| (\beta - \beta_0). \quad (4.28)
 \end{aligned}$$

结合引理 4.2 可知

$$X_i - W^T Q_i^T Z_i = X_i - \Phi^T(U_i) \Gamma^{-1}(U_i) Z_i + O_p\left(\sqrt{pc_n}\right) Z_i. \quad (4.29)$$

令 $\mathcal{X}_i = X_i - \Phi^T(U_i) \Gamma^{-1}(U_i) Z_i$ 且根据 Minkowski 不等式可知,

$$\text{Var}(\|\mathcal{X}_i\|^{r/2}) \leq E(\|\mathcal{X}_i\|^r) = E\left\{\left(\sum_{j=1}^p \mathcal{X}_{ij}^2\right)^{r/2}\right\} \leq \left\{\sum_{j=1}^p \{E(\mathcal{X}_{ij}^r)\}^{2/r}\right\}^{r/2} = O(p^{r/2}),$$

则

$$\begin{aligned} \max_{1 \leq i \leq n} \|\mathcal{X}_i\| &\leq \left[\left\{ \text{Var}(\|\mathcal{X}_i\|^{r/2}) \right\}^{1/2} \max_{1 \leq i \leq n} \left\{ \frac{\|\mathcal{X}_i\|^{r/2} - E\|\mathcal{X}_i\|^{r/2}}{\{\text{Var}(\|\mathcal{X}_i\|^{r/2})\}^{1/2}} \right\} + E\|\mathcal{X}_i\|^{r/2} \right]^{2/r} \\ &= o_p(\sqrt{pn}^{1/r}) + O_p(\sqrt{p}) = o_p(\sqrt{pn}^{1/r}). \end{aligned} \quad (4.30)$$

结合式(4.29)可知

$$\max_{1 \leq i \leq n} \|X_i - W^T Q_i^T Z_i\| = o_p(\sqrt{pn}^{1/r}). \quad (4.31)$$

应用 Shi 和 Lau^[13]的引理 1 可知

$$W^T Q_i^T (\eta_i \eta_i^T - \Sigma_\eta) Q_i W = O(n^{2/r} p \ln n), \quad (X_i - W^T Q_i^T Z_i) \eta_i^T Q_i W = O(n^{2/r} p \ln n).$$

结合引理 4.5 与式(4.28)-(4.31)可知

$$\varphi^*(\beta) = O_p(\sqrt{pn}^{1/r}) + o_p(n^{-1/2+2/r} p^{3/2} \ln n). \quad (4.32)$$

根据引理 4.4 与引理 4.5 可知 $\Theta^T \bar{\varphi}(\beta) = O_p(a_n)$, 则 $\varphi^*(\beta) | \Theta^T \bar{\varphi}(\beta) | = o_p(1)$ 。

类似 Li 等^[51]引理 B.4 的证明, 可得 $\text{tr}[(S(\beta_0) - \Sigma_2)^2] = O_p(p^2(c_n^4 + 1/n))$, 即

$S(\beta_0)$ 的所有特征值以 $O_p(p(c_n^2 + 1/\sqrt{n}))$ 的速率均匀收敛于 Σ_2 的特征值。则

$$\begin{aligned} S(\beta) &= \frac{1}{n} \sum_{i=1}^n \left\{ \frac{\delta_i}{\pi(V_i, \hat{\theta})} \{ \tilde{W}_i (\tilde{Y}_i - \tilde{W}_i^T \beta_0) - W^T Q_i^T \Sigma_\eta Q_i (Y - W \beta_0) + \Sigma_v \beta_0 \} \right. \\ &\quad \left. - \frac{\delta_i}{\pi(V_i, \hat{\theta})} \{ \tilde{W}_i \tilde{W}_i^T (\beta - \beta_0) + W^T Q_i^T \Sigma_\eta Q_i W (\beta - \beta_0) + \Sigma_v (\beta - \beta_0) \} \right\}^{\otimes 2} \\ &= S(\beta_0) + \frac{1}{n} \sum_{i=1}^n \left\{ \frac{\delta_i}{\pi(V_i, \hat{\theta})} (\tilde{W}_i \tilde{W}_i^T - W^T Q_i^T \Sigma_\eta Q_i W - \Sigma_v) (\beta - \beta_0) \right\}^{\otimes 2} \\ &\quad - \frac{2}{n} \sum_{i=1}^n \left\{ \frac{\delta_i}{\pi(V_i, \hat{\theta})} \tilde{W}_i (\tilde{Y}_i - \tilde{W}_i^T \beta_0) - W^T Q_i^T \Sigma_\eta Q_i (Y - W \beta_0) + \Sigma_v \beta_0 \right\} \\ &\quad \cdot \left\{ \frac{\delta_i}{\pi(V_i, \hat{\theta})} (\tilde{W}_i \tilde{W}_i^T - W^T Q_i^T \Sigma_\eta Q_i W - \Sigma_v) (\beta - \beta_0) \right\}. \end{aligned} \quad (4.33)$$

结合条件 (C6) 可知 $\Theta^T S(\beta) \Theta = \Theta^T \Sigma_2 \Theta + o_p(1) \xrightarrow{p} C$, 因此, 通过计算可得

$\rho \leq C | \Theta^T \bar{\varphi}(\beta) | = O_p(a_n)$, 则 $\|\lambda(\beta)\| = O_p(a_n)$ 。

引理 4.7 若条件(C1)-(C11)成立, 当 $n \rightarrow \infty$ 时, 依概率趋近于 1, $L_p(\beta)$ 在 D_n 上取最小值。

证明 由引理 4.5 和引理 4.6 可知, 当 $\beta \in D_n$ 时有

$$\max_{1 \leq i \leq n} |\lambda^T(\beta) \varphi_i(\beta)| \leq \|\lambda^T(\beta)\| \max_{1 \leq i \leq n} \|\varphi_i(\beta)\| = o_p(1), \quad (4.34)$$

将式(4.11)展开可得

$$\lambda = S^{-1}(\beta) \bar{\varphi}(\beta) + o_p(a_n), \quad (4.35)$$

将式(4.10)进行泰勒展开可得

$$-\sum_{i=1}^n \log(1 + \lambda^T \varphi_i(\beta)) = -\sum_{i=1}^n \left[\lambda^T \varphi_i(\beta) - \frac{1}{2} (\lambda^T \varphi_i(\beta))^2 \right] + o_p(1),$$

结合式(4.35), 经过简单计算可得

$$\begin{aligned} 2L(\beta) &= n \bar{\varphi}^T(\beta) S^{-1}(\beta) \bar{\varphi}(\beta) + o_p(1) \\ &= n \bar{\varphi}^T(\beta_0) S^{-1}(\beta_0) \bar{\varphi}(\beta_0) + n \bar{\varphi}(\beta_0) (S^{-1}(\beta) - S^{-1}(\beta_0)) \bar{\varphi}^T(\beta_0) \\ &\quad + n (\bar{\varphi}^T(\beta) - \bar{\varphi}^T(\beta_0)) S^{-1}(\beta) (\bar{\varphi}(\beta) - \bar{\varphi}(\beta_0)) \\ &\quad + 2n (\bar{\varphi}^T(\beta) - \bar{\varphi}^T(\beta_0)) S^{-1}(\beta) \bar{\varphi}(\beta_0) + o_p(1) \\ &= \mathcal{M}_1 + \mathcal{M}_2 + \mathcal{M}_3 + o_p(1). \end{aligned} \quad (4.36)$$

记 ∂D_n 为 D_n 的边界, 对 $\beta \in \partial D_n$, 令 $\beta = \beta_0 + c a_n \iota_\beta$, ι_β 表示单位向量。容易得出

$\mathcal{M}_2 = \mathcal{M}_3 = o_p(1)$, 则当 $n \rightarrow \infty$ 时, 有 $P\{2L(\beta) - 2L(\beta_0) > C\} \rightarrow 1$ 。另外当 n 足够大

时, 结合条件(C10), 对 $i \in \mathcal{H}$, 可得 $p_\tau(|\beta_i|) = p_\tau(|\beta_{i_0}|)$ 。则对足够大的 n , 有

$$\begin{aligned} L_p(\beta) - L_p(\beta_0) &= L(\beta) - L(\beta_0) + n \sum_{i=1}^p \{p_\tau(|\beta_i|) - p_\tau(|\beta_{i_0}|)\} \\ &\geq L(\beta) - L(\beta_0) + n \sum_{i \in \mathcal{H}} \{p_\tau(|\beta_i|) - p_\tau(|\beta_{i_0}|)\} \\ &\geq L(\beta) - L(\beta_0), \end{aligned}$$

因此, 依概率趋近于 1, 有 $L(\beta) > L(\beta_0)$ 。引理得证。

定理 4.1 的证明

根据 M 的定义, 有

$$M = \begin{pmatrix} Z_1^T g(U_1) \\ \vdots \\ Z_n^T g(U_n) \end{pmatrix} = D_u^Z \begin{pmatrix} g(u) \\ hg'(u) \end{pmatrix} + \begin{pmatrix} \frac{1}{2} Z_1^T g''(u)(U_1 - u)^2 \\ \vdots \\ \frac{1}{2} Z_n^T g''(u)(U_n - u)^2 \end{pmatrix} + o_p(h^2), \quad (4.37)$$

根据式(4.37)可得,

$$(D_u^\xi)^T \omega_u \hat{\Delta} M = (D_u^\xi)^T \omega_u \hat{\Delta} D_u^Z \begin{pmatrix} g(u) \\ hg'(u) \end{pmatrix} + \frac{h^2}{2} (D_u^\xi)^T \omega_u \hat{\Delta} \psi_u Z g''(u) + o_p(h^2), \quad (4.38)$$

其中 $\psi_u = \text{diag}\{((U_1 - u)/h)^2, ((U_2 - u)/h)^2, \dots, ((U_n - u)/h)^2\}$ 。

经过简单计算可得下式,

$$\begin{aligned} (D_u^\xi)^T \omega_u \hat{\Delta} D_u^Z \begin{pmatrix} g(u) \\ hg'(u) \end{pmatrix} &= \{(D_u^\xi)^T \omega_u \hat{\Delta} D_u^\xi - \Omega\} \begin{pmatrix} g(u) \\ hg'(u) \end{pmatrix} \\ &\quad + \{-(D_u^\xi)^T \omega_u \hat{\Delta} D_u^\eta + \Omega\} \begin{pmatrix} g(u) \\ hg'(u) \end{pmatrix}, \end{aligned} \quad (4.39)$$

$$\begin{aligned} &\{(D_u^\xi)^T \omega_u \hat{\Delta} D_u^\xi - \Omega\}^{-1} (D_u^\xi)^T \omega_u \hat{\Delta} \psi_u Z g''(u) \\ &= \frac{1}{\mu_2 - \mu_1^2} \begin{pmatrix} (\mu_2^2 - \mu_1 \mu_3) g''(u) \\ (\mu_3 - \mu_1 \mu_2) g''(u) \end{pmatrix} \{1 + o(1)\}. \end{aligned} \quad (4.40)$$

回顾 $\hat{g}(u)$ 的定义, 有

$$\begin{aligned} \hat{g}(u) &= (I_q \ 0_q) \{(D_u^\xi)^T \omega_u \hat{\Delta} D_u^\xi - \Omega\}^{-1} (D_u^\xi)^T \hat{\Delta} \omega_u (Y - W \hat{\beta}_p) \\ &= (I_q \ 0_q) \{(D_u^\xi)^T \omega_u \hat{\Delta} D_u^\xi - \Omega\}^{-1} (D_u^\xi)^T \hat{\Delta} \omega_u M \\ &\quad + (I_q \ 0_q) \{(D_u^\xi)^T \omega_u \hat{\Delta} D_u^\xi - \Omega\}^{-1} (D_u^\xi)^T \hat{\Delta} \omega_u W (\beta - \hat{\beta}_p) \\ &\quad + (I_q \ 0_q) \{(D_u^\xi)^T \omega_u \hat{\Delta} D_u^\xi - \Omega\}^{-1} (D_u^\xi)^T \hat{\Delta} \omega_u (\varepsilon - v\beta) \\ &:= \Lambda_1 + \Lambda_2 + \Lambda_3. \end{aligned} \quad (4.41)$$

结合式 (4.37)-(4.41)可得

$$\Lambda_1 = g(u) + \frac{1}{2} h^2 \frac{\mu_2^2 - \mu_1 \mu_3}{\mu_3 - \mu_1 \mu_2} g''(u)$$

$$+(I_q \ 0_q)\{(D_u^\xi)^T \omega_u \hat{\Delta} D_u^\xi - \Omega\}^{-1}\{-(D_u^\xi)^T \omega_u \hat{\Delta} D_u^\eta + \Omega\} \begin{pmatrix} g(u) \\ hg'(u) \end{pmatrix} + o_p(h^2).$$

结合引理 4.1 与引理 4.2 可得

$$\sqrt{nh}\Lambda_2 = \sqrt{nh}\Gamma^{-1}(u)\Phi(u)\{1 + O_p(c_n)\}O(n^{-1/2}) = o_p(1).$$

类似于 You^[65]关于 A4 至 A6 的证明, 有

$$\begin{aligned} & \sqrt{nh}\{(D_u^\xi)^T \omega_u \hat{\Delta} D_u^\xi - \Omega\}^{-1}\{(D_u^\xi)^T \hat{\Delta} \omega_u (\varepsilon - v\beta) \\ & + \{-(D_u^\xi)^T \omega_u \hat{\Delta} D_u^\eta + \Omega\} \begin{pmatrix} g(u) \\ hg'(u) \end{pmatrix} \} \xrightarrow{d} N(0, \nabla^*), \end{aligned}$$

且

$$\nabla^* = f^{-1}(u)\Sigma^* \otimes \frac{1}{(\mu_2 - \mu_1)^2} \begin{pmatrix} \mu_2^2 e_0 - 2\mu_1 \mu_2 e_1 + \mu_1^2 e_2 & (\mu_1^2 + \mu_2)e_1 - \mu_1 \mu_2 e_0 - \mu_1 e_2 \\ (\mu_1^2 + \mu_2)e_1 - \mu_1 \mu_2 e_0 - \mu_1 e_2 & e_2 - \mu_1(2e_1 + \mu_1 e_0) \end{pmatrix},$$

并结合定理 4.1 中 Σ^* 的定义, 结果得证。

定理 4.2 的证明

根据引理 4.7, 当 $\beta \in D_n$ 时, 对于 $j=1, 2, \dots, p$ 有

$$\frac{1}{n} \frac{\partial L_p(\beta)}{\partial \beta_j} = \frac{1}{n} \sum_{i=1}^n \lambda_{\beta_j}^T \frac{\partial \varphi_i(\beta)}{\partial \beta_j} [1 + o_p(1)] + p'_\tau(|\beta_j|) \text{sign}(\beta_j) := \Pi_1 + \Pi_2,$$

结合条件(C10), (C11), 引理 4.5 和引理 4.7, 有

$$\max_{j \notin \mathcal{H}} |\Pi_1| \leq \max_{j \notin \mathcal{H}} |\lambda_{\beta_j}^T| \max_{j \notin \mathcal{H}} \left| \frac{1}{n} \sum_{i=1}^n \frac{\partial \varphi_i(\beta)}{\partial \beta_j} \right| = O_p(a_n),$$

当 $\tau(n/p)^{1/2} \rightarrow \infty$ 时, 对于任意 $j \notin \mathcal{H}$, $p'_\tau(|\beta_j|) \text{sign}(\beta_j)$ 渐近地主导 $\frac{\partial \varphi_i(\beta)}{\partial \beta_j}$ 的符号。

因此, 对于任意 $j \notin \mathcal{H}$, 当 $n \rightarrow \infty$ 时, 依概率趋近于 1 有

$$\frac{\partial L_p(\beta)}{\partial \beta_j} < 0, \beta_j \in (-Ca_n, 0), \quad \frac{\partial L_p(\beta)}{\partial \beta_j} > 0, \beta_j \in (0, Ca_n),$$

故对于任意 $j \in \mathcal{H}^c$, 依概率趋近于 1 有 $\hat{\beta}_j = 0$, 结论(i)得证。

下证定理 4.2(ii)

取 Ψ_1 和 Ψ_2 满足 $\Psi_1 \beta = \beta_1$ 和 $\Psi_2 \beta = \beta_2$ 。由拉格朗日乘子法得到 $L_p(\beta)$ 的极小

值等价于极小化下式

$$\frac{1}{n} \sum_{i=1}^n \log\{1 + \lambda^T \varphi_i(\beta)\} + \sum_{j=1}^p p_\tau(|\beta_j|) + \mathcal{G}^T \Psi_2 \beta, \quad (4.42)$$

其中 $\mathcal{G}$ 表示 $(p-s)$ 维拉格朗日乘子。定义

$$\begin{aligned} \tilde{\Xi}_{1n}(\beta, \lambda, \mathcal{G}) &= \frac{1}{n} \sum_{i=1}^n \frac{\varphi_i(\beta)}{1 + \lambda^T \varphi_i(\beta)}, \\ \tilde{\Xi}_{2n}(\beta, \lambda, \mathcal{G}) &= \frac{1}{n} \sum_{i=1}^n \frac{[\partial \varphi_i(\beta) / \partial \beta]^T \lambda}{1 + \lambda^T \varphi_i(\beta)} + b(\beta) + \Psi_2^T \mathcal{G}, \\ \tilde{\Xi}_{3n}(\beta, \lambda, \mathcal{G}) &= \Psi_2 \beta, \end{aligned}$$

其中 $b(\beta) = \{p'_\tau(|\beta_1|) \text{sign}(\beta_1), p'_\tau(|\beta_2|) \text{sign}(\beta_2), \dots, p'_\tau(|\beta_p|) \text{sign}(\beta_p), 0, 0, \dots, 0\}^T$ 。

(4.42)式关于 $(\hat{\beta}, \hat{\lambda}, \hat{\mathcal{G}})$ 的最小值满足 $0 = \tilde{\Xi}_{jn}(\hat{\beta}, \hat{\lambda}, \hat{\mathcal{G}}) (j=1, 2, 3)$ 。因为 $\|\lambda\| = O_p(a_n)$,

对于 $\beta \in D_n$, 可以从 $0 = \tilde{\Xi}_{2n}(\hat{\beta}, \hat{\lambda}, \hat{\mathcal{G}})$ 知 $\|\mathcal{G}\| = O_p(a_n)$ 。将 $\tilde{\Xi}_{jn}(\hat{\beta}, \hat{\lambda}, \hat{\mathcal{G}}) (j=1, 2, 3)$ 在

$(\beta_0, 0, 0)$ 处泰勒展开, 首先给出以下偏导数

$$\begin{aligned} \frac{\partial \tilde{\Xi}_{1n}(\beta_0, 0, 0)}{\partial \lambda} &= -S(\beta_0), \frac{\partial \tilde{\Xi}_{1n}(\beta_0, 0, 0)}{\partial \beta} = \frac{1}{n} \sum_{i=1}^n \frac{\delta_i}{\pi(V_i, \hat{\theta})} \{\tilde{W}_i \tilde{W}_i^T - W^T Q_i^T \Sigma_\eta Q_i W + \Sigma_v\}, \\ \frac{\partial \tilde{\Xi}_{1n}(\beta_0, 0, 0)}{\partial \mathcal{G}} &= 0, \frac{\partial \tilde{\Xi}_{2n}(\beta_0, 0, 0)}{\partial \lambda} = \frac{1}{n} \sum_{i=1}^n \frac{\delta_i}{\pi(V_i, \hat{\theta})} \{\tilde{W}_i \tilde{W}_i^T - W^T Q_i^T \Sigma_\eta Q_i W + \Sigma_v\}, \\ \frac{\partial \tilde{\Xi}_{2n}(\beta_0, 0, 0)}{\partial \beta} &= b'(\beta), \frac{\partial \tilde{\Xi}_{2n}(\beta_0, 0, 0)}{\partial \mathcal{G}} = \Psi_2^T, \frac{\partial \tilde{\Xi}_{3n}(\beta_0, 0, 0)}{\partial \lambda} = 0, \\ \frac{\partial \tilde{\Xi}_{3n}(\beta_0, 0, 0)}{\partial \beta} &= \Psi_2, \frac{\partial \tilde{\Xi}_{3n}(\beta_0, 0, 0)}{\partial \mathcal{G}} = 0. \end{aligned}$$

通过泰勒展开可得

$$\begin{pmatrix} -\tilde{\Xi}_{1n}(\beta_0, 0, 0) \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -\Sigma_2 & \Sigma_1 & 0 \\ \Sigma_1 & 0 & \Psi_2^T \\ 0 & \Psi_2 & 0 \end{pmatrix} \begin{pmatrix} \hat{\lambda} \\ \hat{\beta} - \beta_0 \\ \hat{\mathcal{G}} \end{pmatrix} + \mathcal{R}_n, \quad (4.43)$$

其中 $\mathcal{R}_n = \sum_{l=1}^5 \mathcal{R}_n^{(l)}$, $\mathcal{R}_n^{(l)} = (\mathcal{R}_{1n}^{(l)T}, \mathcal{R}_{2n}^{(l)T}, 0)^T$, $\mathcal{R}_{jn}^{(l)T} \in R^p (j=1, 2)$, $K=1, 2, \dots, p$ 且第

K 个分量为

$$\mathcal{R}_{jn,K}^{(1)} = \frac{1}{2}(\hat{\varsigma} - \varsigma_0)^T \frac{\partial^2 \tilde{\Xi}_{jn,K}(\tilde{\varsigma})}{\partial \varsigma \partial \varsigma^T}(\hat{\varsigma} - \varsigma_0),$$

其中 $\varsigma = (\beta^T, \lambda^T)^T$, $\varsigma_0 = (\beta_0^T, 0^T)^T$, $\hat{\varsigma} = (\hat{\beta}^T, \hat{\lambda}^T)^T$, $\tilde{\varsigma} = (\tilde{\beta}^T, \tilde{\lambda}^T)^T$, 有 $\|\tilde{\varsigma} - \varsigma_0\| \leq \|\hat{\varsigma} - \varsigma_0\|$ 且

$$\begin{aligned} \mathcal{R}_n^{(2)} &= (0, b^T(\beta_0), 0)^T, \quad \mathcal{R}_n^{(3)} = [0, \{b'(\tilde{\beta})(\hat{\beta} - \beta_0)\}, 0]^T, \\ \mathcal{R}_n^{(4)} &= [\{(\Sigma_2 - S(\beta_0))\hat{\lambda}\}^T + \{[n^{-1} \sum_{i=1}^n \frac{\delta_i}{\pi(V_i, \hat{\theta})}(\tilde{W}_i \tilde{W}_i^T \\ &\quad - W^T Q_i^T \Sigma_\eta Q_i W \beta + \Sigma_v) - \Sigma_1](\hat{\beta} - \beta_0)\}^T, 0, 0]^T, \\ \mathcal{R}_n^{(5)} &= \left[0, \left\{ \left(n^{-1} \sum_{i=1}^n \frac{\delta_i}{\pi(V_i, \hat{\theta})}(\tilde{W}_i \tilde{W}_i^T - W^T Q_i^T \Sigma_\eta Q_i W \beta + \Sigma_v) - \Sigma_1 \right) \hat{\lambda} \right\}^T, 0 \right]^T. \end{aligned}$$

结合条件(C10)和(C11)可得, 对 $l=1, \dots, 5$, $\mathcal{R}_n^{(l)} = o_p(n^{-1/2})$ 。记 $\Upsilon_{11} = -\Sigma_2$,

$$\Upsilon_{12} = (\Sigma_1, 0), \quad \Upsilon_{21} = \Upsilon_{12}^T,$$

$$\Upsilon_{22} = \begin{pmatrix} 0 & \Psi_2^T \\ \Psi_2 & 0 \end{pmatrix}, \quad \Upsilon = \begin{pmatrix} \Upsilon_{11} & \Upsilon_{12} \\ \Upsilon_{21} & \Upsilon_{22} \end{pmatrix}.$$

定义 $\varpi = (\beta^T, \mathcal{G}^T)^T$, 则根据式 (4.43) 可得

$$\begin{pmatrix} \hat{\lambda} \\ \hat{\varpi} - \varpi_0 \end{pmatrix} = \Upsilon^{-1} \left\{ \begin{pmatrix} -\tilde{\Xi}_{1n}(\beta_0, 0, 0) \\ 0 \end{pmatrix} + \mathcal{R}_n \right\}. \quad (4.44)$$

由于矩阵 Υ 可分为四个分块, 则

$$\Upsilon^{-1} = \begin{pmatrix} \Upsilon_{11}^{-1} + \Upsilon_{11}^{-1} \Upsilon_{12} A^{-1} \Upsilon_{21} \Upsilon_{11}^{-1} & -\Upsilon_{11}^{-1} \Upsilon_{12} A^{-1} \\ -A^{-1} \Upsilon_{21} \Upsilon_{11}^{-1} & A^{-1} \end{pmatrix},$$

其中 $A = \Upsilon_{22} - \Upsilon_{21} \Upsilon_{11}^{-1} \Upsilon_{12}$, 则

$$\hat{\varpi} - \varpi_0 = A^{-1} \Upsilon_{21} \Upsilon_{11}^{-1} \tilde{\Xi}_{1n}(\beta_0, 0, 0) + o_p(n^{-1/2}).$$

同理, A 的逆矩阵为

$$A^{-1} = \begin{pmatrix} \Sigma - \Sigma \Psi_2^T (\Psi_2 \Sigma \Psi_2^T)^{-1} \Psi_2 \Sigma & -\Sigma \Psi_2^T (\Psi_2 \Sigma \Psi_2^T)^{-1} \\ -(\Psi_2 \Sigma \Psi_2^T)^{-1} \Psi_2 \Sigma & (\Psi_2 \Sigma \Psi_2^T)^{-1} \end{pmatrix},$$

因此可知 $\hat{\beta} - \beta_{10} = \{\Sigma - \Sigma \Psi_2^T (\Psi_2 \Sigma \Psi_2^T)^{-1} \Psi_2 \Sigma\} \{\Sigma_1^T \Sigma_2^{-1} \bar{\varphi}(\beta_0) + o_p(n^{-1/2})\}$ 。于是对 $\hat{\beta}_1$ 进

行展开, 可得

$$(\hat{\beta}_1 - \beta_{10}) = \{\Psi_1 \Sigma - \Psi_1 \Sigma \Psi_2^T (\Psi_2 \Sigma \Psi_2^T)^{-1} \Psi_2 \Sigma\} \{\Sigma_1^T \Sigma_2^{-1} \bar{\varphi}(\beta_0) + o_p(n^{-1/2})\},$$

其中 $\Psi_1 \Sigma - \Psi_1 \Sigma \Psi_2^T (\Psi_2 \Sigma \Psi_2^T)^{-1} \Psi_2 \Sigma = (\Sigma_d, 0)$, 则

$$\sqrt{n} \mathcal{A}_n \Sigma_d^{-1/2} (\hat{\beta}_1 - \beta_{10}) \xrightarrow{d} N(0, \mathcal{B}).$$

证毕。

定理 4.3 的证明

通过最小化式(4.12)可得 β 的估计值 $\hat{\beta}$, 式(4.12)通过泰勒展开可得

$$L_p(\hat{\beta}) = \Re_i - \Re_i^2 / 2 + \Re_i^3 / 3(1 + \ell_i)^4 + o_p(1),$$

其中 $\Re_i = \hat{\lambda} \varphi_i(\hat{\beta})$, 对于 ℓ_i 有 $|\ell_i| \leq |\Re_i|$, $o_p(1)$ 可由惩罚函数得出。通过式(4.34)可

知 $\max_{1 \leq i \leq n} |\Re_i| = o_p(1)$, 根据式(4.44)可得

$$\hat{\lambda} = [\Sigma_2^{-1} + \Sigma_2^{-1} \Sigma_1 \{\Sigma - \Sigma \Psi_2^T (\Psi_2 \Sigma \Psi_2^T)^{-1} \Psi_2 \Sigma\} \Sigma_1^T \Sigma_2^{-1}] \{\bar{\varphi}(\beta_0) + o_p(n^{-1/2})\}.$$

在零假设下, 由于 $M_n M_n^T = I_s$, 存在 $(p-s+d) \times p$ 矩阵 $\tilde{\Psi}_2$, 使得 $\tilde{\Psi}_2 \beta = 0$,

$\tilde{\Psi}_2 \tilde{\Psi}_2^T = I_{p-d+s}$ 。记 $\check{\beta} = \arg \min_{\beta} \{L_p(\beta) : \tilde{\Psi}_2 \beta = 0\}$, 且通过最小化式(4.42)可解得 $\check{\beta}$,

其中 Ψ_2 替换为 $\tilde{\Psi}_2$ 。因此可得,

$$\mathcal{L}_n = n \bar{\varphi}^T(\beta_0) \Sigma_2^{-1/2} (\aleph_1 - \aleph_2) \Sigma_2^{-1/2} \bar{\varphi}(\beta_0) + o_p(1),$$

其中 $\aleph_1 = \Sigma_2^{-1/2} \Sigma_1 \tilde{\Psi}_2^T (\tilde{\Psi}_2 \Sigma \tilde{\Psi}_2^T)^{-1} \tilde{\Psi}_2 \Sigma \Sigma_1^T \Sigma_2^{-1/2}$, $\aleph_2 = \Sigma_2^{-1/2} \Sigma_1 \Sigma \Psi_2^T (\Psi_2 \Sigma \Psi_2^T)^{-1} \Psi_2 \Sigma \Sigma_1^T \Sigma_2^{-1/2}$,

且 $\aleph_1$ 和 $\aleph_2$ 是两个幂等矩阵。 $\aleph_1 - \aleph_2$ 的秩为 s , $\aleph_1 - \aleph_2$ 可由 $\mathfrak{U}_n^T \mathfrak{U}_n$ 表示, 其中

$\mathfrak{U}_n \in R^{s \times p}$ 且 $\mathfrak{U} \mathfrak{U}^T = I_s$ 。记 $\mathcal{L}_{n1} = \sqrt{n} \mathfrak{U}_n \Sigma_2^{-1/2} \bar{\varphi}(\beta_0)$, 则

$$\mathcal{L}_{n1} = n^{-1/2} \mathfrak{U}_n \Sigma_2^{-1/2} \sum_{i=1}^n J_i \{1 + o_p(1)\} \xrightarrow{d} N(0, I_s),$$

因此, $\mathcal{L}_n = \mathcal{L}_{n1}^T \mathcal{L}_{n1} \xrightarrow{d} \chi_s^2$, 证毕。

4.6 本章小结

本章主要通过逆概率加权和局部纠偏的方法构造参数 β 的辅助随机向量, 并

在适当的条件下证明了其渐近正态性质。根据构造的参数 β 辅助随机向量获得参数 β 的经验对数似然比函数，同时建议惩罚经验似然(PEL)进行变量选择，在适当的条件下证明了所提出的惩罚经验估计具有Oracle特征且服从渐近标准卡方分布。最后进行数值模拟，可知基于惩罚经验似然方法的模型表现结果较好。非参数 $g(U)$ 的置信带在缺失率，测量误差较小的情况下要优于缺失率，测量误差较大的情况。最后将所述方法分析了香港日均因呼吸系统疾病住院人数与污染物和其他环境因素的影响关系，分析显示，在适当的条件下二氧化氮浓度对呼吸系统疾病住院人数有显著影响，二氧化硫浓度对呼吸系统疾病住院人数无显著影响。

第5章 总结与展望

5.1 总结

半参数模型目前受到了广泛的关注,由于社会的快速发展,对半参数模型的要求也越来越精细化。本文考虑响应变量随机缺失并且协变量存在测量误差的半参数模型,分别研究了在参数部分附加约束条件和高维协变量下的统计推断问题。

首先研究了响应变量随机缺失下部分线性测量误差模型的约束估计。利用反卷积的方法分别在完整观测数据和借补技术下得到了非参数部分的估计量,并在适当的条件下证明其一致性。同时采用局部纠偏的Profile最小二乘方法,在参数部分附加约束条件的基础上分别给出完整观测数据和借补技术下参数部分的纠偏约束估计量与借补纠偏约束估计量,然后基于一定的条件证明了其渐近性质。数值模拟结果显示参数部分的借补纠偏约束估计量要优于完整观测数据下的纠偏约束估计量。最后的实际例子也较好地体现了数值模拟结果的有效性。

其次考虑了响应变量随机缺失下高维的半参数变系数测量误差模型的惩罚经验似然估计。基于逆概率加权与局部纠偏技巧,给出了非参数函数的纠偏估计。然后将经验似然方法与惩罚函数相结合构建了惩罚经验似然函数,并在适当的条件下证明了建议的惩罚经验似然估计具有Oracle性质。这有效地解决了高维半参数变系数模型的参数估计与变量选择问题。数值模拟显示在缺失率、样本容量、参数维数相同的情况下,随着测量误差的减小,其参数部分估计值的均方误差越小。同时在样本容量、参数维数、测量误差相同的情况下,随着缺失率的减少,其参数部分估计值的均方误差越小。零系数平均值的估计结果也接近真实模型,说明惩罚经验似然方法表现较好。非参数函数的置信带在缺失率、测量误差较小的情况下有比较好的表现。最后的实际数据分析也进一步验证了惩罚经验似然方法的有效性。

5.2 展望

作者今后将对以下几个方面做进一步的研究：

1、本文可进一步研究在利用未知误差函数来表示误差可能存在形式下的半参数模型的统计推断问题。

2、本文主要考虑了响应变量是随机缺失的，而复杂数据不仅仅有缺失数据，还囊括了删失数据、纵向数据等。基于这些数据的研究将更好地拓宽半参数模型的适用范围，具有重要的研究意义。

3、本文主要考虑了半参数模型中的部分线性模型与半参数变系数模型，后续可以研究部分线性变系数单指标模型、可加模型等在附加约束条件下的估计问题与高维协变量下的惩罚经验似然估计。

参考文献

- [1] Engle R F , Granger C W J , Rice J , etc. Semiparametric estimates of the relation between weather and electricity sales[J]. Journal of the American Statistical Association, 1986, 81(394): 310-320.
- [2] Speckman P. Kernel smoothing in partial linear model[J]. Journal of the Royal Statistical Society. Series B: Methodological, 1988, 50(3): 413-436.
- [3] Liang H , Hrdle W , Carroll R J. Estimation in a semiparametric partially linear errors-in-variable model[J]. The Annals of Statistics, 1999, 27(5): 1519-1535.
- [4] Liang H. Asymptotic normality of parametric part in partially linear models with measurement error in the nonparametric part[J]. Journal of Statistical Planning and Inference, 2000, 86(1): 51-62.
- [5] Fan J J, Huang T. Profile likelihood inferences on semiparametric varying-coefficient partially linear models[J]. Bernoulli, 2005, 11(6): 1031-1057.
- [6] You J H, Chen G M. Estimation of a semiparametric varying-coefficient partially linear errors-in-variables model[J]. Journal of Multivariate Analysis, 2005, 97(2): 324-341.
- [7] 魏传华, 吴喜之. 部分线性变系数模型的 Profile Lagrange 乘子检验[J]. 系统科学与数学, 2008, 28(4): 416-424.
- [8] Wei C H. Statistical inference for restricted partially linear varying coefficient errors-in-variables models[J]. Journal of Statistical Planning and Inference, 2012, 142(8): 2464-2472.
- [9] Yang J S, Park B U. Efficient estimation for partially linear varying coefficient models when coefficient functions have different smoothing variables[J]. Journal of Multivariate Analysis, 2014, 126: 100-113.
- [10] Aman U, Wang T ,Yao W X. Semiparametric partially linear varying coefficient modal regression[J]. Journal of Econometrics, 2023, 235(2): 1001-1026.
- [11] Thomas D R, Grunkemeier G L. Confidence interval estimation of survival

- probabilities for censored data[J]. Journal of the American Statistical Association, 1975, 70(352): 865-871.
- [12] Owen A B. Empirical likelihood ratio confidence intervals for a single functional[J]. Biometrika, 1988, 75(2): 237-249.
- [13] Shi J, Lau T S. Empirical likelihood for partially linear models[J]. Journal of Multivariate Analysis, 2000, 72(1): 132-148.
- [14] Li G R, Xue L G. Empirical likelihood confidence region for the parameter in a partially linear errors-in-variables model[J]. Communications in Statistics - Theory and Methods, 2008, 37(10): 1552-1564.
- [15] Fan G L, Liang H Y, Wang J F. Empirical likelihood for heteroscedastic partially linear errors-in-variables model with α -mixing errors[J]. Statistical papers, 2013, 54(1): 85-112.
- [16] You J H, Zhou Y. Empirical likelihood for semiparametric varying-coefficient partially linear regression models[J]. Statistics and Probability Letters, 2005, 76(4): 412-422.
- [17] Fan G L, Xu H X, Liang H Y. Empirical likelihood inference for partially time-varying coefficient errors-in-variables models[J]. Electronic Journal of Statistics, 2012, 6: 1040-1058.
- [18] Yang Y P, Li G R, Peng H. Empirical likelihood of varying coefficient errors-in-variables models with longitudinal data[J]. Journal of Multivariate Analysis, 2014, 127: 1-18.
- [19] Fan G L, Xu H X, Huang Z S. Empirical likelihood for semivarying coefficient model with measurement error in the nonparametric part[J]. AStA Advances in Statistical Analysis, 2016, 100(1): 21-41.
- [20] Sun H H, Liu Q. Robust empirical likelihood inference for partially linear varying coefficient models with longitudinal data[J]. Journal of Statistical Computation and Simulation, 2023, 93(10): 1559-1579.
- [21] Wang Q H, Sun Z H. Estimation in partially linear models with missing responses at random[J]. Journal of Multivariate Analysis, 2007, 98(7): 1470-1493.
- [22] Xiao Y T, Li F X. Estimation in partially linear varying-coefficient errors-in-

- variables models with missing response variables[J]. Computational Statistics, 2020, 35(4): 1-22.
- [23] 范国良, 饶诗文, 王江峰. 缺失数据下变系数部分非线性测量误差模型的经验似然估计[J]. 系统科学与数学, 2021, 41(9): 2643-2659.
- [24] Yan Y X, Luo S H, Zhang C Y. Statistical inference for partially linear varying coefficient quantile models with missing responses[J]. Symmetry, 2022, 14(11): 2258-2269.
- [25] Zou Y Y, Wu C X, Fan G L, etc. Jackknife empirical likelihood of error variance for partially linear varying-coefficient model with missing covariates[J]. Communications in Statistics - Theory and Methods, 2023, 52(6): 1744-1766.
- [26] Zou H, Zhang H H. On the adaptive elastic-net with a diverging number of parameters[J]. The Annals of Statistics, 2009, 37(4): 1733-1751.
- [27] Zhu L P, Zhu L X. On distribution-weighted partial least squares with diverging number of highly correlated predictors[J]. Journal of the Royal Statistical Society. Series B (Statistical Methodology), 2009, 71(2): 525-548.
- [28] Wu Y C, Li L X. Asymptotic properties of sufficient dimension reduction with a diverging number of predictors[J]. Statistica Sinica, 2011, 2011(21): 707-730.
- [29] Wu T T, Li G, Tang C Y. Empirical likelihood for censored linear regression and variable selection[J]. Scandinavian journal of statistics, theory and applications, 2015, 42(3): 798-812.
- [30] Chen X, Mao L X. Penalized empirical likelihood for partially linear errors-in-variables models[J]. AStA Advances in Statistical Analysis, 2020, 104(4): 1-27.
- [31] He B Q, Lv S R. Penalized empirical likelihood for high-dimensional partially linear errors-in-function model with martingale difference errors[J]. Statistics, 2020, 54(6): 1214-1231.
- [32] Wang B H, Liang H Y. Quantile regression of ultra-high dimensional partially linear varying-coefficient model with missing observations[J]. Acta Mathematica Sinica, English Series, 2023, 39(9): 1701-1726.
- [33] Parzen E. On estimation of probability density function and mode[J]. The Annals of Mathematical Statistics, 1962, 33(3): 1065-1076.

-
- [34] Nadaraya E A. On estimating regression[J]. Theory of Probability and Its Applications, 1964, 9(1): 141–142.
 - [35] Watson G S. Smooth regression analysis[J]. Sankhyā the Indian Journal of Statistics, Series A, 1964, 26(4): 359-372.
 - [36] Cleveland W S. Robust locally weighted regression and smoothing scatterplots[J]. Journal of the American Statistical Association, 1979, 74(368), 829-836.
 - [37] Fan J Q, Gijbels I. Data-driven bandwidth selection in local polynomial fitting: Variable bandwidth and spatial adaptation[J]. Journal of the Royal Statistical Society, Series B (Methodological), 1995, 57(2): 371-394.
 - [38] Ruppert D. Empirical-bias bandwidths for local polynomial nonparametric regression and density estimation[J]. Journal of the American Statistical Association, 1997, 92(439): 1049-1062.
 - [39] Owen A B. Empirical likelihood[M]. New York: Chapman and Hall-CRC, 2001.
 - [40] Owen A B. Empirical likelihood ratio confidence regions[J]. The Annals of Statistics, 1990, 18(1): 90-120.
 - [41] Ma Y Y, Chiou J M, Wang N. Efficient semiparametric estimator for heteroscedastic partially linear models[J]. Biometrika, 2006, 93(1): 75-84.
 - [42] Sun Z H, Wang Q H, Dai P J. Model checking for partially linear models with missing responses at random[J]. Journal of Multivariate Analysis, 2009, 100(4): 636-651.
 - [43] Wang Q H. Statistical estimation in partial linear models with covariate data missing at random[J]. Annals of the Institute of Statistical Mathematics, 2009, 61(1): 47-84.
 - [44] Xue L G, Xue D. Empirical likelihood for semiparametric regression model with missing response data[J]. Journal of Multivariate Analysis, 2010, 102(4): 723-740.
 - [45] Feng S Y, Xue L G. Bias-corrected statistical inference for partially linear varying coefficient errors-in-variables models with restricted condition[J]. Annals of the Institute of Statistical Mathematics, 2014, 66(1): 121-140.
 - [46] Xia Y F, Zhao Z H, Niu Y J. Estimation of partially linear varying-coefficient EV model under restricted condition[J]. Journal of Statistical and Econometric

- Methods, 2014, 3(4): 57-70.
- [47] Fan J Q, Truong Y K. Nonparametric regression with errors in variables[J]. The Annals of Statistics, 1993, 21(4): 1900-1925.
- [48] Yang Y P, Xue L G, Cheng W H. Two-step estimators in partial linear models with missing response variables and error-prone covariates[J]. Journal of Systems Science and Complexity, 2011, 24(6): 1165-1182.
- [49] Cui H J. Estimation in partial linear EV models with replicated observations[J]. Science in China, Series A, 2004, 47(01) : 144-159.
- [50] Hu X M, Wang Z Z, Zhao Z W. Empirical likelihood for semiparametric varying-coefficient partially linear errors-in-variables models[J]. Statistics and Probability Letters, 2008, 79(8): 1044-1052.
- [51] Li G R, Lin L, Zhu L X. Empirical likelihood for a varying coefficient partially linear model with diverging number of parameters[J]. Journal of Multivariate Analysis, 2012, 105(1): 85-111.
- [52] Sun J, Sun Q H. An improved and efficient estimation method for varying-coefficient model with missing covariates[J]. Statistics and Probability Letters, 2015, 107: 296-303.
- [53] 陈盼盼, 冯三营, 薛留根. 缺失数据下半参数变系数部分线性模型的统计推断[J]. 数学物理学报, 2015, 35(2): 345-358.
- [54] Xu H X, Fan G L, Wu C X, etc. Statistical inference for varying-coefficient partially linear errors-in-variables models with missing data [J]. Communications in Statistics-Theory and Methods, 2019, 48(22): 5621-5636.
- [55] Fan J Q, Li R Z. Variable selection via nonconcave penalized likelihood and its oracle properties[J]. Publications of the American Statistical Association, 2001, 96(456): 1348-1360.
- [56] Tang C Y, Leng T C. Penalized high-dimensional empirical likelihood[J]. Biometrika, 2010, 97(4): 905-920.
- [57] Leng C L, Tang C Y. Penalized empirical likelihood and growing dimensional general estimating equations[J]. Biometrika, 2012, 99(3): 703-716.
- [58] Fan G L, Liang H Y, Shen Y. Penalized empirical likelihood for high-dimensional

- partially linear varying coefficient model with measurement errors[J]. *Journal of Multivariate Analysis*, 2016, 147: 183-201.
- [59] Yi G Y, Ma Y Y, Carroll R J. A functional generalized method of moments approach for longitudinal studies with missing responses and covariate measurement error[J]. *Biometrika*, 2012, 99(1): 151-165.
- [60] Wang H S, Li R Z, Tsai C L. Tuning parameter selectors for the smoothly clipped absolute deviation method[J]. *Biometrika*, 2007, 94(3): 553-568.
- [61] Xia Y C, Li W K. On the estimation and testing of functional-coefficient linear models[J]. *Statistica Sinica*, 1999, 9(3): 735-757.
- [62] Volkonskii V A , Rozanov Y A. Some limit theorems for random functions[J]. *Theory of Probability and Its Applications*, 1959, 4(2): 178-197.
- [63] Yang S C. Maximal moment inequality for partial sums of strong mixing sequences and application[J]. *Acta Mathematica Sinica, English Series*, 2007, 23(6): 1013-1024.
- [64] Hall P G, Heyde C C. *Martingale limit theory and its application*[M]. New York: Academic Press, 1980.
- [65] You H J, Zhou Y, Chen G M. Corrected local polynomial estimation in varying coefficient models with measurements errors[J]. *The Canadian Journal of Statistics*, 2006, 34(3): 391-410.

硕士期间发表的论文

- [1] 王龙, 何帮强. 缺失数据下部分线性测量误差模型的约束估计[J]. 安徽工程大学学报(已录用).
- [2] 何帮强, 王龙. 缺失数据下高维半参数变系数测量误差模型的惩罚经验似然[J]. 系统科学与数学(已录用).

致 谢

三年忙碌而又充实的研究生生涯即将进入尾声，值此之际，谨向三年里在学习和生活上给予我关心和帮助的老师、同学以及家人最真诚的感谢！

首先，我要特别感谢我的导师何帮强教授，非常感谢何老师平时十分重视对我专业知识的培养以及从论文选题到具体写作以及最后的修改定稿都给予了我具有建设性的指导意见。在生活上遇到的困难，何老师也给予了我无微不至的关心并且让我形成了良好的生活习惯。这将是我最人生中珍贵的财富之一。在此，衷心地祝愿何老师今后的工作和生活中万事如意！

其次，我要特别感谢姜培华老师、何俊老师、梁勇老师、张伟老师以及夏登峰老师在研一为我的研究生学习奠定了坚实的基础。我也要特别感谢张玥老师、沈明轩老师与白晓老师在研二专业技能的提升上提供了指导，同时也非常感谢院里的其他老师营造了浓厚的学习氛围，让我能够保持积极的学习态度去完成每一阶段的目标。在此，衷心地祝愿每一位老师事业蒸蒸日上，生活如意吉祥！

再次，我要特别感谢我的室友以及数学班的全体同学。在这三年里无论是学习还是生活中，大家能够相互交流、互帮互助、携手共进，留下了许许多多令人难忘的回忆，也给予了彼此在今后的工作和生活中非常多的祝福。在此，祝愿大家往后的工作和生活事事顺意！

最后我要感谢我的家人们，感谢他们对我无私的支持与爱，每当我遇到挫折时，他们的支持都会让我拥有莫大的勇气去迎接挑战，只要坚持终究会到达理想彼岸。我将带着家人的支持与爱在今后的生活中不断奋发向上！