

文章编号: 0583-1431(2018)04-0549-20

文献标识码: A

纵向数据的有效秩推断 基于修正的 Cholesky 分解

吕 晶

西南大学数学与统计学院 重庆 400715

E-mail: cqulv2015@126.com

郭朝会 杨 虎 李婷婷

重庆师范大学数学科学学院 重庆 401331

重庆大学数学与统计学院 重庆 401331

西南大学数学与统计学院 重庆 400715

E-mail: guochoahui2010@126.com; yh@cqu.edu.cn; tinalee@swu.edu.cn

摘要 本文基于修正的 Cholesky 分解提出新的方法估计纵向秩回归的组内协方差矩阵, 进而提出新的无偏估计函数改善不平衡纵向数据的估计效率. 在一些正则条件下, 建立了所提估计的渐近正态性. 进一步, 提出稳健的秩得分检验统计量对回归系数做假设检验. 模拟研究和实证分析表明所提方法能够获得高度有效的估计以及所提检验方法比存在的方法更好.

关键词 诱导光滑; 纵向数据; 修正的 Cholesky 分解; 秩得分检验

MR(2010) 主题分类 62E20

中图分类 O212.4

Efficient Rank Inference Based on the Modified Cholesky Decomposition with Longitudinal Data

Jing LV

School of Mathematics and Statistics, Southwest University, Chongqing 400715, P. R. China

E-mail: cqulv2015@126.com

Chao Hui GUO Hu YANG Ting Ting LI

College of Mathematics Science, Chongqing Normal University, Chongqing 401331, P. R. China

College of Mathematics and Statistics, Chongqing University, Chongqing 401331, P. R. China

School of Mathematics and Statistics, Southwest University, Chongqing 400715, P. R. China

E-mail: guochoahui2010@126.com; yh@cqu.edu.cn; tinalee@swu.edu.cn

收稿日期: 2017-02-28; 接受日期: 2017-09-07

基金项目: 国家社会科学基金(17CTJ015); 重庆市基础科学与前沿研究技术专项项目(cstc2017jcyjAX0182, cstc2016jcyjA0510); 中央高校基本科研业务费专项资金(XDK2017C051); 国家自然科学基金(11671059)

通讯作者: 郭朝会

Abstract This paper proposes a new approach to estimate the within-subject covariance of longitudinal rank regression based on the modified Cholesky decomposition matrix. Then new rank-based unbiased estimating functions are developed to improve estimation efficiency in the analysis of unbalanced longitudinal data. Under some regularity conditions, we establish the asymptotically normal distributions of the resulting estimators. Moreover, we propose a robust rank score test for hypothesis on the regression coefficients. Simulation studies and a real data analysis show that the proposed approach yields highly efficient estimators and the proposed rank score test is more powerful than the existing approaches.

Keywords induced smoothing; longitudinal data; modified Cholesky decomposition; rank score test

MR(2010) Subject Classification 62E20

Chinese Library Classification O212.4

1 引言

纵向数据频繁出现在许多科学领域, 例如生物医学、社会科学、环境研究。纵向数据的一个显著特征是同一个人在不同时间点所得观测值是相关的。Liang 和 Zeger^[11] 提出处理纵向数据的一般方法广义估计方程 (GEE)。此方法的优点是即使相关结构被误判所得估计仍然是相合的, 但是误判的相关结构会损失估计的效率。同时 GEE 方法对于异常值很敏感, 因为它本质上类似广义最小二乘。

Hettmansperger^[7] 指出基于秩回归的方法是分布自由的, 稳健的以及高度有效的。在独立数据下已有一些研究结果, 譬如, Wang 等人^[15] 提出了加权的 SCAD 秩回归估计, 所提方法可以同时进行变量选择和获得稳健估计。Wang 等人^[14] 关于变系数模型提出了局部秩估计并且研究了所提估计的大样本性质。Yang 等人^[19] 基于秩回归研究了高维线性模型的变量选择。秩回归的方法可以应用到纵向数据中, 并且相对于 GEE 来说更具有优势。Jung 等人^[9] 在独立工作结构下构造了广义的 Wilcoxon–Mann–Whitney 秩统计量。由于没有考虑组内相关结构。因此, 所得估计不是渐近有效的。为了提高估计的效率, 考虑组内相关结构是很有必要的。Wang 等人^[17] 基于组内和组间的秩构造了两类型的估计函数从而改进了^[9] 的方法。Wang 等人^[16] 考虑重复测量间的相关性提出了加权的 Wilcoxon 秩回归方法。基于 Wang 等人^[17] 的思想, Fu 等人^[5] 基于秩回归提出了诱导光滑的估计函数, 同时获得回归系数和它们的渐近协方差矩阵的估计。最近, Fu 等人^[4] 在可交换相关结构下考虑了基于秩的估计函数, 并且证明了它相对于大多数存在的方法具有更好的统计性质, 所提方法的缺陷是当组内相关结构不是可交换时会损失估计的效率。

因此, 建立协方差模型, 允许更一般的相关结构是一个非常有趣的话题。在建立协方差模型方面修正的 Cholesky 分解是一个非常有效的工具^[10, 20, 21]。然而, 以上所提参考文献仅仅考虑的是非稳健的回归模型。因此, 将修正的 Cholesky 分解应用到纵向秩回归模型还很缺乏。本文将利用修正的 Cholesky 分解提出稳健的均值协方差模型, 该方法很灵活且高度有效。然而, 基于秩回归的估计函数非凸, 不连续且不可微。本文将利用 Brown 等人^[1] 提出的诱导光滑的方法获得参数更快且更精确的估计。进一步, 假设检验在统计推断中是很重要的一个方面。尽管可以构造 Wald 检验统计量, 但是它在有限样本下表现不好且很难获得误差密度的精确估计。为了避免以

上问题, 本文基于有效的估计方程提出秩得分检验统计量.

本文第 2 节简单地介绍独立和可交换工作模型, 然后基于修正的 Cholesky 分解提出新的无偏估计函数和研究相应的渐近性质. 第 3 节提出更有效的秩得分检验统计量用于假设检验, 并且讨论它的渐近性质. 第 4 节通过随机模拟验证所提方法的有限样本性质. 第 5 节将所提方法应用到实际数据中. 第 6 节给出本文的结论. 定理需要满足的条件和主要结果的证明放在附录中.

2 基于秩回归的估计

考虑如下线性模型

$$Y_{ik} = \mathbf{X}_{ik}^T \boldsymbol{\beta} + \varepsilon_{ik}, \quad i = 1, \dots, n, \quad k = 1, \dots, m_i, \quad (2.1)$$

其中 n 是个体的数目, m_i 是第 i 个个体的重复测量次数, Y_{ik} 和 $\mathbf{X}_{ik} = (X_{ik1}, \dots, X_{ikp})^T$ 是响应变量和相应的协变量的观测值, $\boldsymbol{\beta}$ 是 p 维未知参数向量. $\varepsilon_i = (\varepsilon_{i1}, \dots, \varepsilon_{im_i})^T$ 为随机误差向量, 则 $\varepsilon_i, i = 1, \dots, n$ 是独立的. 但是对于同一个个体, 误差项 $(\varepsilon_{i1}, \dots, \varepsilon_{im_i})^T$ 可能是相关的. 假定 $\boldsymbol{\beta}_0$ 是 $\boldsymbol{\beta}$ 的真实值且总的观测次数为 $N = \sum_{i=1}^n m_i$. 注意在模型 (2.1) 中, 没有考虑截距项因为秩回归关于位置移动是不变的.

在独立工作结构下, Jung 和 Ying [9] 提出了如下的估计函数

$$U_I(\boldsymbol{\beta}) = n^{-2} \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^{m_i} \sum_{l=1}^{m_j} (\mathbf{X}_{ik} - \mathbf{X}_{jl}) \operatorname{sgn}(\varepsilon_{ik}(\boldsymbol{\beta}) - \varepsilon_{jl}(\boldsymbol{\beta})) \quad (2.2)$$

用于估计 $\boldsymbol{\beta}$, 其中 $\varepsilon_{ik}(\boldsymbol{\beta}) = Y_{ik} - \mathbf{X}_{ik}^T \boldsymbol{\beta}$. 令 $\hat{\boldsymbol{\beta}}_I$ 为 (2.2) 的解. 由于估计函数 (2.2) 是基于独立相关结构构造的, 因此 $\hat{\boldsymbol{\beta}}_I$ 会损失估计的效率.

为了提高估计的效率, Fu 和 Wang [4] 提出了如下最佳秩估计函数

$$\mathbf{U}_{\text{opt}}(\boldsymbol{\beta}) = \bar{\mathbf{D}}^T \boldsymbol{\Sigma}^{-1} \mathbf{S}(\boldsymbol{\beta}), \quad (2.3)$$

其中

$$\boldsymbol{\Sigma} = \operatorname{Cov}(\mathbf{S}(\boldsymbol{\beta})), \quad \mathbf{S}(\boldsymbol{\beta}) = (\mathbf{S}_1(\boldsymbol{\beta})^T, \dots, \mathbf{S}_n(\boldsymbol{\beta})^T)^T, \quad \mathbf{S}_i(\boldsymbol{\beta}) = (S_{i1}(\boldsymbol{\beta}), \dots, S_{im_i}(\boldsymbol{\beta}))^T,$$

$$S_{ik}(\boldsymbol{\beta}) = N^{-1} \sum_{j=1}^n \sum_{l=1}^{m_j} [I\{\varepsilon_{jl}(\boldsymbol{\beta}) < \varepsilon_{ik}(\boldsymbol{\beta})\} - 1/2], \quad \bar{\mathbf{D}} = \{\bar{\mathbf{D}}_{ik}, i = 1, \dots, n, k = 1, \dots, m_i\}$$

是 $\mathbf{S}(\boldsymbol{\beta})$ 的期望的导数且

$$\bar{\mathbf{D}}_{ik}(\boldsymbol{\beta}_0) = N^{-1} \sum_{j=1}^n \sum_{l=1}^{m_j} (\mathbf{X}_{jl} - \mathbf{X}_{ik}) f_{ikjl}(0),$$

其中 $f_{ikjl}(\cdot)$ 为 $\varepsilon_{ik}(\boldsymbol{\beta}_0) - \varepsilon_{jl}(\boldsymbol{\beta}_0)$ 的密度函数. 为了不引起叙述上的混淆, 本文将忽略 $\varepsilon_{ik}(\boldsymbol{\beta})$ 和 $S_{ik}(\boldsymbol{\beta})$ 中的 $\boldsymbol{\beta}$, 对于任意的 i 和 k . 文 [4] 进一步指出当误差分布满足一些条件时, 可以估计 $f_{ikjl}(0)$, 但是为了简化计算, 他们假定 $f_{ikjl}(0)$ 是一个常数尽管这样可能会损失一些效率. 因此, 在估计函数 (2.3) 中, 他们用 $\mathbf{D}_{ik} = \bar{\mathbf{X}} - \mathbf{X}_{ik}$ 代替 $\bar{\mathbf{D}}_{ik}$, 其中 $\bar{\mathbf{X}} = N^{-1} \sum_{j=1}^n \sum_{l=1}^{m_j} \mathbf{X}_{jl}$. 估计函数 (2.3) 包含未知的协方差矩阵 $\boldsymbol{\Sigma}$. 众所周知一个好的协方差矩阵估计可以提高 $\boldsymbol{\beta}$ 估计的效率. 然而, 文 [4] 指出当误差 ε_i 服从 AR(1) 或者其他相关结构时, 很难指定 \mathbf{S}_i 的相关结构且 \mathbf{S}_i 的相关结构很复杂. 为了克服以上困难, 文 [4] 在建立估计函数时采用可交换的相关结构来提高估计的效率. 具体地, 他们在可交换相关结构下提出了如下估计函数

$$\mathbf{U}_F(\boldsymbol{\beta}) = \sum_{i=1}^n \mathbf{D}_i^T \mathbf{V}_i^{-1} \mathbf{S}_i(\boldsymbol{\beta}) = \mathbf{0}, \quad (2.4)$$

其中 $\mathbf{D}_i = (\mathbf{D}_{i1}, \dots, \mathbf{D}_{im_i})^T$, $\mathbf{V}_i^{-1} = (\sigma_i^2 - \sigma_{ii})^{-1} [\mathbf{I}_{m_i} - \sigma_{ii}(\sigma_i^2 + (m_i - 1)\sigma_{ii})^{-1} \mathbf{J}_{m_i}]$, \mathbf{I}_{m_i} 是 $m_i \times m_i$ 的单位矩阵, \mathbf{J}_{m_i} 是元素全为 1 的 $m_i \times m_i$ 维矩阵. σ_i^2 和 σ_{ii} 定义于文 [4, 2.2 节]. 因为 σ_i^2 和 σ_{ii} 是未知的参数, 他们利用矩方法估计 σ_i^2 和 σ_{ii} . 可以参考 Fu 和 Wang [4] 的补充材料. 令 $\hat{\beta}_F$ 为 (2.4) 的解. 注意 \mathbf{V}_i 是在可交换相关结构下的工作协方差矩阵. 众所周知, 只有当 ε_i 具有可交换相关结构时, \mathbf{S}_i 和 ε_i 才具有相同的相关结构. 因此, 当 ε_i 不具有可交换相关结构时, $\hat{\beta}_F$ 会损失估计效率因为 \mathbf{S}_i 的相关结构被误判了.

本文将基于修正的 Cholesky 分解提出新的方法对 \mathbf{S}_i 的协方差矩阵进行建模. 新方法不需要事先指定工作相关结构, 因此该方法更一般且灵活. 进一步, 修正的 Cholesky 分解可以保证估计的协方差矩阵是正定的, 并且协方差矩阵中的参数具有很好的统计意义. 基于 Yao 和 Li [20], $\Sigma_i = \text{Cov}(\mathbf{S}_i(\beta_0))$ 可以分解为

$$\text{Cov}(\Phi_i \mathbf{S}_i) = \Phi_i \Sigma_i \Phi_i^T = \Lambda_i, \quad (2.5)$$

其中 Φ_i 是主对角元素全为 1 的下三角矩阵以及 $\Lambda_i = \text{diag}(d_{i1}^2, \dots, d_{im_i}^2)$ 是 $m_i \times m_i$ 对角矩阵. 进一步, 令 $\mathbf{e}_i = (e_{i1}, \dots, e_{im_i})^T = \Phi_i \mathbf{S}_i$ 和 $\phi_{k,l}^{(i)}$ 是 Φ_i 中第 (k, l) 个元素的相反数. 因此可得

$$\begin{aligned} S_{i1} &= e_{i1}, \\ S_{ik} &= \phi_{k,1}^{(i)} S_{i1} + \dots + \phi_{k,k-1}^{(i)} S_{i,k-1} + e_{ik}, \quad i = 1, \dots, n, \quad k = 2, \dots, m_i. \end{aligned} \quad (2.6)$$

注意 $e_{ik}, k = 1, \dots, m_i$ 是不相关的, 因为 Λ_i 是对角矩阵且 $\text{Var}(e_{ik}) = d_{ik}^2, k = 1, \dots, m_i$. 基于文 [12] 的思想, 本文分别对自回归系数 $\phi_{k,l}^{(i)}$ 和更新方差 d_{ik}^2 构造线性模型和非参模型

$$\phi_{k,l}^{(i)} = \mathbf{W}_{k,l}^{(i)T} \boldsymbol{\theta}, \quad d_{ik}^2 \triangleq d^2(t_{ik}) = \sigma(t_{ik}), \quad (2.7)$$

其中 $\boldsymbol{\theta}$ 和 $\sigma(\cdot)$ 分别是 $s \times 1$ 维的回归系数和一维未知光滑函数, 以及 $\mathbf{W}_{k,l}^{(i)}$ 是 $s \times 1$ 维协变量. 根据文 [21] 的建议, 本文取 $\mathbf{W}_{k,l}^{(i)} = (1, t_{ik} - t_{il}, \dots, (t_{ik} - t_{il})^{s-1})^T$.

接下来, 应用最小二乘方法估计未知的 $\boldsymbol{\theta}$ 和 $\sigma(\cdot)$. 首先, 可获得估计的残差 $\hat{\varepsilon}_{ik}(\hat{\beta}_I) = Y_{ik} - \mathbf{X}_{ik}^T \hat{\beta}_I, i = 1, \dots, n, k = 1, \dots, m_i$, 然后解如下最小二乘问题获得 $\hat{\boldsymbol{\theta}}$,

$$\begin{aligned} \hat{\boldsymbol{\theta}} &= \arg \min_{\boldsymbol{\theta}} \sum_{i=1}^n \sum_{k=2}^{m_i} (\hat{S}_{ik} - \hat{S}_{i1} \mathbf{W}_{k,1}^{(i)T} \boldsymbol{\theta} - \dots - \hat{S}_{i,k-1} \mathbf{W}_{k,k-1}^{(i)T} \boldsymbol{\theta})^2 \\ &= \left(\sum_{i=1}^n \sum_{k=2}^{m_i} \hat{\mathbf{v}}_{ik} \hat{\mathbf{v}}_{ik}^T \right)^{-1} \sum_{i=1}^n \sum_{k=2}^{m_i} \hat{\mathbf{v}}_{ik} \hat{S}_{ik}, \end{aligned}$$

其中

$$\hat{S}_{ik} = N^{-1} \sum_{j=1}^n \sum_{l=1}^{m_j} [I\{\hat{\varepsilon}_{jl}(\hat{\beta}_I) < \hat{\varepsilon}_{ik}(\hat{\beta}_I)\} - 0.5], \quad \hat{\mathbf{v}}_{ik} = \left(\sum_{l=1}^{k-1} \hat{S}_{il} \mathbf{W}_{k,l,1}^{(i)}, \dots, \sum_{l=1}^{k-1} \hat{S}_{il} \mathbf{W}_{k,l,s}^{(i)} \right)^T.$$

因此可得 $\hat{\phi}_{k,l}^{(i)} = \mathbf{W}_{k,l}^{(i)T} \hat{\boldsymbol{\theta}}$. 基于文 [3] 的思想, 可得 $d^2(t)$ 的估计, 即解如下最小二乘问题

$$(\hat{\eta}_1, \hat{\eta}_2) = \arg \min_{\eta_1, \eta_2} \sum_{i=1}^n \sum_{k=1}^{m_i} [\hat{e}_{ik}^2 - \eta_1 - \eta_2(t_{ik} - t)]^2 K_h(t_{ik} - t),$$

其中 $K_h(\cdot) = K(\cdot/h)/h$, h 是窗宽, $K(\cdot)$ 为核函数, $\hat{e}_{i1} = \hat{S}_{i1}$, $\hat{e}_{ik} = \hat{S}_{ik} - \sum_{l=1}^{k-1} \hat{\phi}_{k,l}^{(i)} \hat{S}_{il}, i = 1, \dots, n, k = 2, \dots, m_i$. 由此可得 $\hat{d}^2(t) = \hat{\eta}_1$. 接下来的定理给出了 $\hat{\boldsymbol{\theta}}$ 和 $\hat{d}^2(t)$ 的渐近性质.

定理 2.1 若附录中的条件 (C1)–(C7) 成立, 则当 $n \rightarrow \infty$,

$$\sqrt{N-n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{d} N(\mathbf{0}, \boldsymbol{\Lambda}^{-1} \boldsymbol{\Delta} \boldsymbol{\Lambda}^{-1}),$$

其中 \rightarrow^d 表示依分布收敛, $\Delta = \lim_{n \rightarrow \infty} \frac{1}{N-n} \sum_{i=1}^n \sum_{k=2}^{m_i} d_{ik}^2 E(\mathbf{v}_{ik} \mathbf{v}_{ik}^T)$ 以及 Λ 定义于附录中条件 (C7).

为了获得 $\hat{d}^2(t)$ 的渐近结果, 将 (2.6) 改写为如下异方差模型

$$\begin{aligned} S_{i1} &= d_{i1}\varsigma_{i1}, \\ S_{ik} &= \phi_{k,1}^{(i)}S_{i1} + \cdots + \phi_{k,k-1}^{(i)}S_{i,k-1} + d_{ik}\varsigma_{ik}, \quad i = 1, \dots, n, \quad k = 2, \dots, m_i, \end{aligned} \quad (2.8)$$

则 $E(\varsigma_{ik} | t_{ik}) = 0$, $\text{Var}(\varsigma_{ik} | t_{ik}) = 1$ 和 $\text{Var}(e_{ik}) = d_{ik}^2$, 其中 $e_{ik} = d_{ik}\varsigma_{ik}$ 和 $d_{ik} = d(t_{ik})$. 令 $\mu_j = \int u^j K(u)du$, $\nu_j = \int u^j K^2(u)du$ 和 $f_T(t)$ 表示 t_{ik} 的边缘密度函数.

定理 2.2 若定理 2.1 中的条件成立, 以及 $h \rightarrow 0$, $Nh \rightarrow \infty$, $n \rightarrow \infty$ 和 $\limsup_{n \rightarrow \infty} Nh^5 < \infty$, 则

$$\sqrt{Nh} \left\{ \hat{d}^2(t) - d^2(t) - \frac{1}{2}\mu_2 h^2 \ddot{d}^2(t) \right\} \xrightarrow{d} N(0, \Xi),$$

其中 $\Xi = \lim_{n \rightarrow \infty} \frac{\nu_0}{Nf_T(t)} \sum_{i=1}^n \sum_{k=1}^{m_i} E[(\varsigma_{ik}^2 - 1)^2 | t_{ik} = t] d^4(t)$ 和 $\ddot{d}(t) = d''(t)$ 表示 $d(t)$ 的二阶导数.

注 2.3 估计非参函数 $\sigma(\cdot)$ 时, 窗宽 h 扮演着非常重要的角色. 本文利用五折交叉验证获得最佳窗宽.

基于以上的讨论, 可得 $\hat{\Sigma}_i^{-1} = \hat{\Phi}_i^T \hat{\Lambda}_i^{-1} \hat{\Phi}_i$, 其中 $\hat{\Phi}_i$ 是主对角元素全为 1 的下三角矩阵且 $\hat{\Phi}_i$ 的第 (k, l) 个元素为 $-\hat{\phi}_{k,l}^{(i)}$, 以及 $\hat{\Lambda}_i = \text{diag}(\hat{d}_{i1}^2, \dots, \hat{d}_{im_i}^2)$. 类似于估计函数 (2.4), 可构造如下的有效估计函数

$$\mathbf{U}_R(\boldsymbol{\beta}) = \sum_{i=1}^n \mathbf{D}_i^T \hat{\Sigma}_i^{-1} \mathbf{S}_i(\boldsymbol{\beta}). \quad (2.9)$$

令 $\hat{\boldsymbol{\beta}}_R$ 为 (2.9) 的解. 然而所提估计函数 (2.9) 既不光滑也不单调. 为了克服计算困难, 我们利用文 [1] 提出的诱导光滑方法获得 $\boldsymbol{\beta}$ 的估计. 通过简单地计算, $\mathbf{U}_R(\boldsymbol{\beta})$ 的诱导光滑函数可写为

$$\tilde{\mathbf{U}}_R(\boldsymbol{\beta}) = \sum_{i=1}^n \mathbf{D}_i^T \hat{\Sigma}_i^{-1} \tilde{\mathbf{S}}_i(\boldsymbol{\beta}), \quad (2.10)$$

其中 $\tilde{\mathbf{S}}_i(\boldsymbol{\beta}) = (\tilde{S}_{i1}(\boldsymbol{\beta}), \dots, \tilde{S}_{im_i}(\boldsymbol{\beta}))^T$, $\tilde{S}_{ik}(\boldsymbol{\beta}) = N^{-1} \sum_{j=1}^n \sum_{l=1}^{m_j} [\Phi(\varepsilon_{ikjl}/r_{ikjl}) - 0.5]$, $(i, k) \neq (j, l)$, $\Phi(\cdot)$ 表示标准正态的卷积分布函数, $\varepsilon_{ikjl} = \varepsilon_{ik} - \varepsilon_{jl}$, $r_{ikjl}^2 = \mathbf{d}_{ikjl}^T \boldsymbol{\Omega} \mathbf{d}_{ikjl}$, $\mathbf{d}_{ikjl} = \mathbf{X}_{ik} - \mathbf{X}_{jl}$ 和 $\boldsymbol{\Omega}$ 是 $\tilde{\boldsymbol{\beta}}_R$ 的 $p \times p$ 维协方差矩阵, 其中 $\tilde{\boldsymbol{\beta}}_R$ 为秩回归估计函数 (2.10) 的解.

注 2.4 因为 $\tilde{\mathbf{U}}_R(\boldsymbol{\beta})$ 是 $\boldsymbol{\beta}$ 的光滑函数, 则 Newton-Raphson 迭代可以用来获得 $\tilde{\mathbf{U}}_R(\boldsymbol{\beta}) = \mathbf{0}$ 的根.

定理 2.5 若附录中的条件 (C1)–(C6) 成立, 则 $\hat{\boldsymbol{\beta}}_R \xrightarrow{p} \boldsymbol{\beta}_0$ 且

$$\sqrt{n}(\hat{\boldsymbol{\beta}}_R - \boldsymbol{\beta}_0) \xrightarrow{d} N(\mathbf{0}, \boldsymbol{\Gamma}^{-1} \boldsymbol{\Psi}(\boldsymbol{\Gamma}^T)^{-1}),$$

其中 \xrightarrow{p} 表示依概率收敛, $\boldsymbol{\Gamma}$ 和 $\boldsymbol{\Psi}$ 定义于附录中.

定理 2.6 若附录中的条件 (C1)–(C6) 成立, 则

$$n^{-1/2}[\tilde{\mathbf{U}}_R(\boldsymbol{\beta}) - \mathbf{U}_R(\boldsymbol{\beta})] = o_p(1).$$

定理 2.7 若附录中的条件 (C1)–(C6) 成立, 则 $\tilde{\boldsymbol{\beta}}_R \xrightarrow{p} \boldsymbol{\beta}_0$ 且

$$\sqrt{n}(\tilde{\boldsymbol{\beta}}_R - \boldsymbol{\beta}_0) \xrightarrow{d} N(\mathbf{0}, \boldsymbol{\Gamma}^{-1} \boldsymbol{\Psi}(\boldsymbol{\Gamma}^T)^{-1}).$$

注 2.8 定理 2.6 表明光滑估计函数和非光滑估计函数渐近等价. 定理 2.5 和 2.7 表明光滑估计与非光滑估计具有相同的渐近分布.

3 假设检验

对于线性回归模型, Gutenbrunner, Jurečková, Koenker 和 Portnoy^[6] 提出的秩得分检验是一种流行且强大的检验方法. 本节基于有效的估计函数 $\mathbf{U}_R(\boldsymbol{\beta})$ 提出秩得分检验统计量. 不失一般性, 把参数向量划分为 $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^T, \boldsymbol{\beta}_2^T)^T$, 其中 $\boldsymbol{\beta}_1 \in \mathbb{R}^{p_1}$ 和 $\boldsymbol{\beta}_2 \in \mathbb{R}^{p-p_1}$. 考虑原假设为 $H_0: \boldsymbol{\beta}_1 = \mathbf{0}_{p_1}$. 令 $\mathbf{X}_{(1)}$ 和 $\mathbf{X}_{(2)}$ 分别是对应于 $\boldsymbol{\beta}_1$ 和 $\boldsymbol{\beta}_2$ 的 $N \times p_1$ 和 $N \times (p - p_1)$ 维设计矩阵. 进一步, 定义 $\boldsymbol{\Pi} = (\mathbf{I} - \mathbf{X}_{(2)}(\mathbf{X}_{(2)}^T \mathbf{X}_{(2)})^{-1} \mathbf{X}_{(2)}^T) \mathbf{X}_{(1)}$ 以及 $\boldsymbol{\Pi}_{ik} \in \mathbb{R}^{p_1}$ 为 $\boldsymbol{\Pi}$ 的第 i 个个体的第 k 次观测值 $\mathbf{D}_{ik(1)} = \bar{\boldsymbol{\Pi}} - \boldsymbol{\Pi}_{ik}$, $\bar{\boldsymbol{\Pi}} = N^{-1} \sum_{j=1}^n \sum_{l=1}^{m_j} \boldsymbol{\Pi}_{jl}$; $\mathbf{D}_{i(1)} = (\mathbf{D}_{i1(1)}, \dots, \mathbf{D}_{im_i(1)})^T$. 令 $\bar{\boldsymbol{\beta}}$ 表示在原假设 H_0 下 (2.9) 的解, 于是定义如下秩得分检验统计量

$$\hat{T}_n = \hat{\mathbf{S}}_n^T \hat{\mathbf{V}}_n^{-1} \hat{\mathbf{S}}_n, \quad (3.1)$$

其中

$$\begin{aligned} \hat{\mathbf{S}}_n &= n^{-1/2} \sum_{i=1}^n \mathbf{D}_{i(1)}^T \hat{\boldsymbol{\Sigma}}_i^{-1} \mathbf{S}_i(\bar{\boldsymbol{\beta}}), \quad \hat{\mathbf{V}}_n = n^{-1} \sum_{i=1}^n \mathbf{D}_{i(1)}^T \hat{\boldsymbol{\Sigma}}_i^{-1} \mathbf{S}_i(\bar{\boldsymbol{\beta}}) (\mathbf{D}_{i(1)}^T \hat{\boldsymbol{\Sigma}}_i^{-1} \mathbf{S}_i(\bar{\boldsymbol{\beta}}))^T, \\ \mathbf{S}_i(\bar{\boldsymbol{\beta}}) &= (S_{i1}(\bar{\boldsymbol{\beta}}), \dots, S_{im_i}(\bar{\boldsymbol{\beta}}))^T, \quad \text{且} \quad S_{ik}(\bar{\boldsymbol{\beta}}) = N^{-1} \sum_{j=1}^n \sum_{l=1}^{m_j} [I\{\varepsilon_{jl}(\bar{\boldsymbol{\beta}}) < \varepsilon_{ik}(\bar{\boldsymbol{\beta}})\} - 1/2], \end{aligned}$$

及 $\varepsilon_{ik}(\bar{\boldsymbol{\beta}}) = Y_{ik} - \mathbf{X}_{ik}^T \bar{\boldsymbol{\beta}}$.

定理 3.1 若附录中的条件 (C1)–(C6) 和 (C8) 成立, 以及在原假设 $H_0: \boldsymbol{\beta}_1 = \mathbf{0}_{p_1}$ 下, 可得 $\hat{T}_n \xrightarrow{d} \chi^2(p_1)$, 当 $n \rightarrow \infty$.

4 模拟研究

为了研究所提方法的有限样本性质, 本文比较 $(\hat{\boldsymbol{\beta}}_R)$ 与一些存在的方法, 例如 $\hat{\boldsymbol{\beta}}_I$, 传统的 GEE 方法在指定的相关结构下, 如可交换相关结构 $(\hat{\boldsymbol{\beta}}_G^e)$ 和 AR(1) 相关结构 $(\hat{\boldsymbol{\beta}}_G^a)$ 以及 $\hat{\boldsymbol{\beta}}_F$.

例 4.1 数据来源于如下模型

$$Y_{ik} = X_{ik1}\beta_1 + X_{ik2}\beta_2 + \varepsilon_{ik},$$

其中 $\beta_1 = -0.5$, $\beta_2 = 1$, $\mathbf{X}_{ik} = (X_{ik1}, X_{ik2})^T$ 来源于相关系数为 0.5 的标准多元正态分布. 为了研究所提估计的表现行为, 考虑如下两种误差分布:

Case I 相关的正态误差, ε_i 服从多元正态分布 $N(\mathbf{0}, \boldsymbol{\Xi}_i)$, 其中 $\boldsymbol{\Xi}_i$ 将定义于下文.

Case II 相关的 t 分布, ε_i 服从自由度为 3 的多元 t 分布且协方差矩阵为 $\boldsymbol{\Xi}_i$.

类似于文 [12], ε_i 的协方差矩阵 $\boldsymbol{\Xi}_i$ 定义为 $\boldsymbol{\Xi}_i = \Delta_i^{-1} \mathbf{B}_i (\Delta_i^T)^{-1}$, 其中 \mathbf{B}_i 是 $m_i \times m_i$ 的对角矩阵且第 k 个对角元素为 $\sin(\pi \varsigma_{ik})/3 + 0.5$, $\varsigma_{ik} \sim U(0, 2)$, 以及 Δ_i 是对角元素全为 1 的下三角矩阵且第 (k, l) 个元素为 $-\delta_{k,l}^{(i)}$, $(k \neq l)$, $\delta_{k,l}^{(i)} = \mathbf{W}_{k,l}^{(i)T} \boldsymbol{\theta}^*$, 其中 $\mathbf{W}_{k,l}^{(i)} = (1, t_{ik} - t_{il})^T$, $\boldsymbol{\theta}^* = (0.2, 0.5)^T$ 和 $t_{ik} \sim U(0, 1)$. 对于重复测量次数 m_i , 考虑不平衡的情形. 令 m_i 为 2 到 m ($m = 10$ 或 20) 间的随机数, 此处的 m 可以看作是 m_i , $i = 1, \dots, n$ 的最大值. 考虑 $n = 50$ 和 100 , 重复模拟 200 次 (见表 1).

例 4.2 为了进行公平对比, 用文 [4] 类似方法构造协方差矩阵 $\boldsymbol{\Xi}_i = \mathbf{A}_i^{1/2} \mathbf{C}_i \mathbf{A}_i^{1/2}$, 其中 $\mathbf{A}_i = \text{diag}(\sigma_{i1}^2, \dots, \sigma_{im_i}^2)$ 和 \mathbf{C}_i 分别为相关系数 0.7 的可交换 (Exch) 和 AR(1) 相关结构, 及 $\sigma_{ij}^2 = 1 + 2t_{ij}^2$ 和 $t_{ij} \sim U(0, 2)$. 类同文 [10], 取 $m = 10$ 或 20 , 但每个元素有 20% 的概率缺失, 从而导致不平衡的纵向数据. 考虑 $n = 50$, 重复模拟 200 次. 其他的设置和例 4.1 一样 (见表 2).

n	m	Case I				Case II				
		β_1		β_2		β_1		β_2		
		bias	EFF	bias	EFF	bias	EFF	bias	EFF	
50	10	$\hat{\beta}_I$	-0.0469	1.000	-0.3787	1.000	0.2780	1.000	-0.3735	1.000
		$\hat{\beta}_G^e$	0.1138	1.129	0.2027	1.149	0.7369	0.538	-0.2283	0.605
		$\hat{\beta}_G^a$	0.2521	0.974	0.1944	0.947	0.7321	0.513	-0.3601	0.503
		$\hat{\beta}_F$	-0.1421	1.171	-0.2400	1.060	0.2895	1.151	-0.2756	1.224
		$\tilde{\beta}_R$	0.0676	1.576	-0.2396	1.592	0.0944	1.569	-0.0099	1.711
		$\hat{\beta}_I$	0.4049	1.000	0.3178	1.000	0.0086	1.000	0.1423	1.000
20	20	$\hat{\beta}_G^e$	-0.1741	0.660	0.2213	0.738	-0.5205	0.391	0.5733	0.393
		$\hat{\beta}_G^a$	-0.5534	0.634	0.2617	0.673	-0.2444	0.403	-0.0370	0.377
		$\hat{\beta}_F$	0.4257	1.389	0.1256	1.395	0.2083	1.138	0.0249	1.302
		$\tilde{\beta}_R$	0.0423	2.284	0.0361	2.722	0.1090	1.908	0.1227	1.874
		$\hat{\beta}_I$	0.2650	1.000	-0.5947	1.000	-0.1257	1.000	0.4308	1.000
		$\hat{\beta}_G^e$	0.0803	1.141	-0.1584	1.049	-0.2898	0.548	0.1529	0.523
100	10	$\hat{\beta}_G^a$	0.1426	0.966	-0.0229	0.879	-0.3589	0.560	0.1063	0.514
		$\hat{\beta}_F$	0.2876	1.195	-0.6085	1.241	-0.0937	1.090	0.4562	1.039
		$\tilde{\beta}_R$	0.0943	1.818	-0.5214	1.452	0.1219	1.427	0.0477	1.447
		$\hat{\beta}_I$	0.3137	1.000	0.2465	1.000	0.4180	1.000	-0.5289	1.000
		$\hat{\beta}_G^e$	0.2536	0.758	-0.2921	0.707	0.5236	0.350	-0.2008	0.295
		$\hat{\beta}_G^a$	0.0398	0.753	-0.1970	0.790	0.8834	0.301	-0.0498	0.322
20	20	$\hat{\beta}_F$	0.2578	1.313	0.2033	1.300	0.2220	1.210	-0.3094	1.365
		$\tilde{\beta}_R$	0.2961	2.213	-0.0479	2.345	-0.0028	1.915	-0.4137	2.458

表 1 例 4.1 的随机模拟结果

m	Exch				AR(1)					
	β_1		β_2		β_1		β_2			
	bias	EFF	bias	EFF	bias	EFF	bias	EFF		
Case I	10	$\hat{\beta}_I$	0.3062	1.000	-0.9067	1.000	-1.3634	1.000	0.3852	1.000
		$\hat{\beta}_G^e$	0.1973	2.089	-0.1955	2.009	-0.5783	1.261	-0.3449	1.154
		$\hat{\beta}_G^a$	0.3845	1.501	-0.6224	1.394	-0.2201	1.697	-0.1504	1.735
		$\hat{\beta}_F$	0.1758	2.581	-0.0024	2.889	-1.1924	1.380	0.1590	1.276
		$\tilde{\beta}_R$	0.3304	2.624	0.0207	2.882	-0.6764	1.462	-0.1072	1.459
		$\hat{\beta}_I$	-1.2801	1.000	0.7120	1.000	-0.1335	1.000	-0.0094	1.000
Case II	20	$\hat{\beta}_G^e$	0.1086	2.207	-0.1086	2.400	0.0011	1.009	-0.1683	1.014
		$\hat{\beta}_G^a$	0.0532	1.533	-0.0223	1.685	0.0431	1.813	-0.2634	1.762
		$\hat{\beta}_F$	-0.9554	2.476	0.4237	2.732	0.0365	1.287	-0.6172	1.109
		$\tilde{\beta}_R$	-0.5284	2.202	0.1091	3.092	-0.1734	1.312	-0.1539	1.286
		$\hat{\beta}_I$	0.2103	1.000	1.8830	1.000	0.5869	1.000	-0.6452	1.000
		$\hat{\beta}_G^e$	-0.0460	0.924	0.0932	0.973	-0.8176	0.643	1.5570	0.714
Case II	10	$\hat{\beta}_G^a$	-0.3303	0.686	0.0303	0.627	-0.4112	0.955	0.8063	1.121
		$\hat{\beta}_F$	0.2389	2.816	1.0736	2.465	-0.0407	1.493	-0.6503	1.465
		$\tilde{\beta}_R$	0.0503	2.846	0.6004	2.289	0.2275	1.752	-0.7470	1.810
		$\hat{\beta}_I$	0.4646	1.000	-0.2462	1.000	0.2327	1.000	-0.3240	1.000
		$\hat{\beta}_G^e$	-0.4362	1.166	0.2448	1.048	0.4940	0.533	-0.5063	0.515
		$\hat{\beta}_G^a$	-0.6014	0.806	0.2111	0.782	0.3539	0.865	-0.8785	0.829
20	20	$\hat{\beta}_F$	0.1362	2.714	-0.1463	2.489	0.5117	1.241	-1.0497	1.324
		$\tilde{\beta}_R$	0.4189	2.701	-0.4488	2.500	0.5815	1.518	-0.6667	1.491

表 2 例 4.2 的随机模拟结果

例 4.3 考虑混合分布. 随机误差项 ε_i 以 $1 - \delta\%$ 的概率服从多元正态分布 $N(\mathbf{0}, \boldsymbol{\Xi}_i^a)$ 以及剩下的 $\delta\%$ 个数据来源于自由度为 3 且协方差矩阵为 $\boldsymbol{\Xi}_i^b$ 的 t 分布. 考虑 $n = 100$, $m = 10$, 重复模拟 200 次. 其他的设置和例 4.2 中的一样. 具体地, 考虑如下两种混合分布 (见表 3):

Case III $\boldsymbol{\Xi}_i^a = \mathbf{A}_i^{1/2} \mathbf{C}_i^a \mathbf{A}_i^{1/2}$ 且 \mathbf{C}_i^a 为相关系数为 0.7 的可交换相关结构; $\boldsymbol{\Xi}_i^b = \mathbf{A}_i^{1/2} \mathbf{C}_i^b \mathbf{A}_i^{1/2}$ 且 \mathbf{C}_i^b 为相关系数为 0.7 的 AR(1) 相关结构.

Case IV $\boldsymbol{\Xi}_i^a = \mathbf{A}_i^{1/2} \mathbf{C}_i^a \mathbf{A}_i^{1/2}$ 且 \mathbf{C}_i^a 为相关系数为 0.7 的 AR(1) 相关结构; $\boldsymbol{\Xi}_i^b = \mathbf{A}_i^{1/2} \mathbf{C}_i^b \mathbf{A}_i^{1/2}$ 且 \mathbf{C}_i^b 为相关系数为 0.7 的可交换相关结构.

$\delta\%$	Case III				Case IV				
	β_1		β_2		β_1		β_2		
	bias	EFF	bias	EFF	bias	EFF	bias	EFF	
20%	$\hat{\beta}_I$	-0.9911	1.000	0.9022	1.000	1.2352	1.000	-0.5577	1.000
	$\hat{\beta}_G^e$	-0.2907	1.735	0.0388	1.863	0.3741	1.237	-0.1408	1.640
	$\hat{\beta}_G^a$	-0.1094	1.765	-0.1526	1.589	-0.4035	0.948	0.3787	1.477
	$\hat{\beta}_F$	-0.0256	3.146	0.2943	3.126	0.4213	1.956	-0.1400	2.239
	$\tilde{\beta}_R$	-0.1842	3.351	0.1004	3.106	0.8417	2.319	-0.4693	2.382
50%	$\hat{\beta}_I$	-0.2105	1.000	0.3689	1.000	-0.0428	1.000	0.0568	1.000
	$\hat{\beta}_G^e$	-0.4096	0.970	0.2962	0.913	0.4236	1.591	-0.0135	1.507
	$\hat{\beta}_G^a$	-0.3746	1.410	0.4220	1.326	-0.1584	1.538	0.2779	1.380
	$\hat{\beta}_F$	-0.5354	2.316	-0.1304	2.377	0.4308	2.715	0.0144	2.177
	$\tilde{\beta}_R$	-0.4208	2.450	-0.1165	2.399	0.6420	2.874	-0.4018	2.287

表 3 例 4.3 的随机模拟结果

首先考虑所提估计 $\tilde{\beta}_R$ 关于 $\mathbf{W}_{k,l}^{(i)}$ 的维数是否敏感. 本文考虑 $s = 2, 3, 4$ 及相应的结果 (图 1). 我们发现对于不同的 s , 所提估计表现得差不多. 因此, 在模拟结果中只列出了 $s = 2$ 的情况. 表 1-3 给出了 $\hat{\beta}_G^e$, $\hat{\beta}_G^a$, $\hat{\beta}_F$ 和 $\tilde{\beta}_R$ 关于 $\hat{\beta}_I$ 的渐近相对效以及这四种估计的偏差 ($\times 10^{-2}$). 根据这三张表, 我们可以得到如下结论.

第一, 在有限样本下所有估计的偏差都很小, 表明他们都是相合的.

第二, 在非正态分布下, 传统的 GEE 估计表现得最差, 这也说明 GEE 方法不是稳健的.

第三, 所提估计 $\tilde{\beta}_R$ 相对于其他两种秩估计 $\hat{\beta}_I$ 和 $\hat{\beta}_F$ 具有更高的效率. 因此, 利用修正的 Cholesky 分解处理秩回归模型的组内协方差矩阵是有效的.

第四, 当 ε_i 服从可交换相关结构时, $\tilde{\beta}_R$ 和 $\hat{\beta}_F$ 表现类似, 因为在这种情形下 \mathbf{S}_i 和 ε_i 具有相同的相关结构, 即 $\hat{\beta}_F$ 利用了正确的工作相关结构. 当误差 ε_i 不具有可交换相关结构时, 所提估计 $\tilde{\beta}_R$ 表现得比 $\hat{\beta}_F$ 好. 因此, 与存在的方法相比, 所提估计 $\tilde{\beta}_R$ 具有更强的竞争优势. 其主要原因在于所提方法允许协方差结构具有更一般的形式, 不需事先指定协方差的结构形式. 因此新方法比大多数存在的方法更灵活. 总之, 可以得出所提估计方法在纵向数据研究中具有更好的稳健性和有效性的结论.

由图 2 可知估计的密度函数曲线与正态密度函数曲线很接近, 以及 Q-Q 图中的散点基本上分布在一条线上. 这表明 $\hat{\theta}$ 渐近服从正态分布. 关于 $\hat{d}^2(t)$ 可以得到类似的结论, 为了节省空间这里将忽略其结果. 图 3 画出了估计的 $\hat{\sigma}(t)$ 曲线以及相应的 95% 置信区间, 可知 $\hat{\sigma}$ 随着时间变化. 随着样本量的增加, $\hat{\sigma}(t)$ 的 95% 置信区间越来越窄.

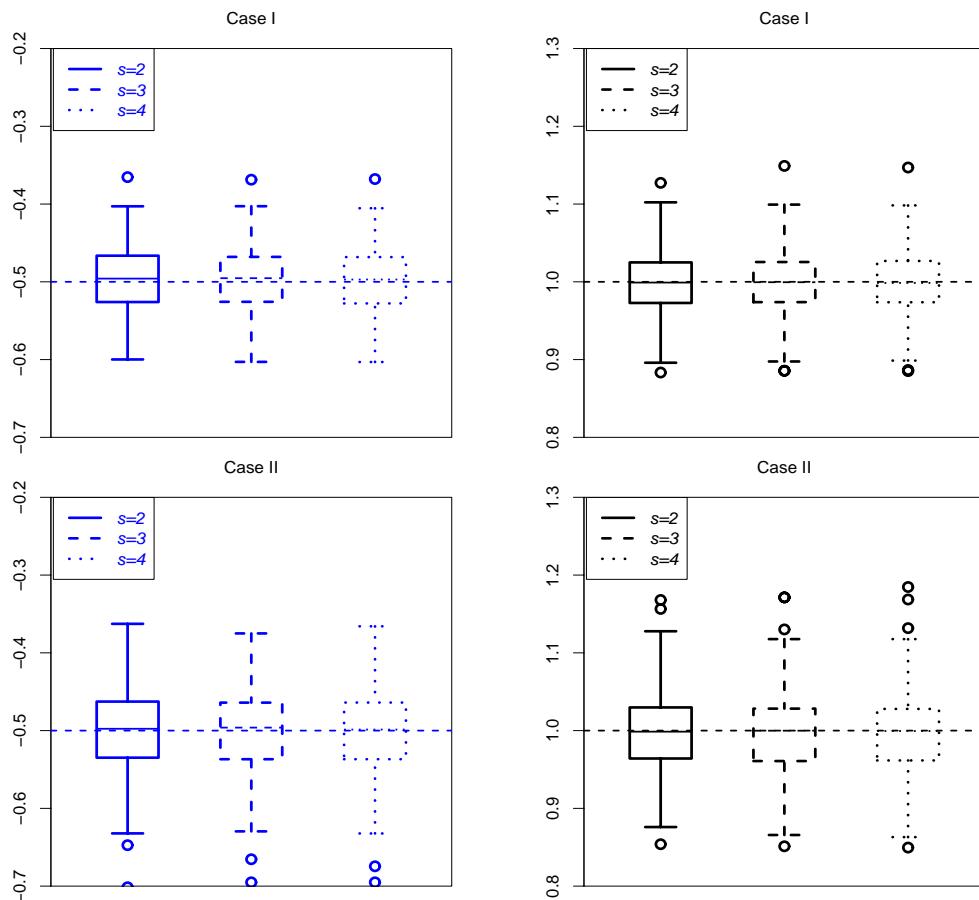


图 1 当 $n = 50$ 和 $m = 10$, 对于不同维数 s , 例 4.1 中 β_1 (左边) 和 β_2 (右边) 的箱线图

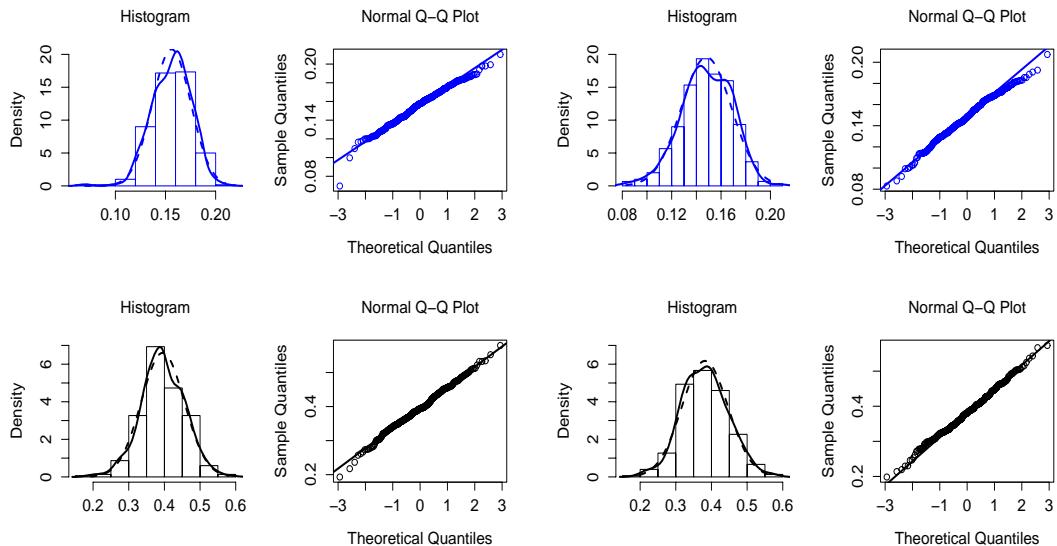
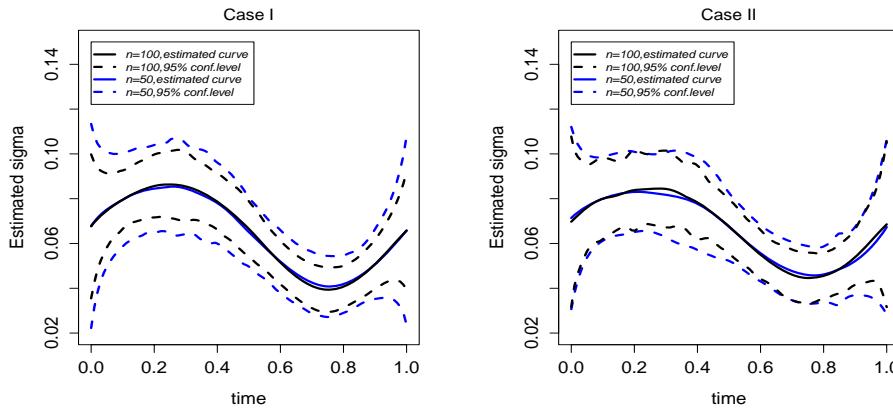


图 2 例 4.1 中 θ_1 (上排) 和 θ_2 (下排) 在 Case I (左平面) 和 Case II (右平面) 下 300 次估计值的直方图, 估计的密度函数曲线 (实线), 正态密度函数曲线 (虚线) 和 300 次估计值的 Q-Q 图当 $n=50$ 和 $m=10$

图 3 当 $m = 10$, 例 4.1 中 $\hat{\sigma}(t)$ 的曲线图 (实线) 以及相应的 95% 置信区间 (点线)

最后研究所提秩得分检验统计量 \hat{T}_n 的有限样本性质. 考虑原假设 $H_0: \beta_1 = 0$ 以及取 $\beta = (c, 1)^T$, 其中 c 以 0.05 个单位从 -0.2 增加到 0 . 此外, 取 $n = 50$ 和 $m = 10$ 重复模拟 400 次得到秩得分检验的势函数 (功效函数) 的估计, 即 $400^{-1} \sum_{i=1}^{400} I(\hat{T}_{n,i} > \chi_{1,\alpha}^2)$, 其中显著性水平为 $\alpha = 0.05$. 注意 $\hat{T}_{n,i}$ 是 \hat{T}_n 的第 i 次结果以及 $\chi_{1,\alpha}^2$ 是 χ_1^2 的 $100(1-\alpha)\%$ 分位点. 为了公平比较, 考虑三种秩得分检验, 在指定相关结构下, 如独立结构 (表示为 Ind) 和可交换相关结构 (表示为 Ex) 及本文提出的自适应方法 (表示为 Ad). 注意: Ind 的获得是将 (3.1) 中的 $\hat{\Sigma}_i$ 用 $m_i \times m_i$ 维的单位矩阵代替. Ex 的获得是将 (3.1) 中的 $\hat{\Sigma}_i$ 用 (2.4) 中 V_i 代替. 表 4, 5 列举了例 4.1 和 4.2 中三种秩得分检验的功效函数值. 当 $c = 0$ 时, 三种方法犯第一类错误的概率很接近显著性水平 0.05. 进一步, 发现本文所提检验的功效函数值大于其他两种检验的功效函数值当 $c \neq 0$ 时. 此外, 随着 c 的绝对值的增加功效函数值快速增加到 1. 这些结果表明本文所提秩得分检验统计量更好.

c	Case I			Case II		
	Ind	Ex	Ad	Ind	Ex	Ad
-0.00	0.040	0.040	0.055	0.070	0.075	0.068
-0.05	0.163	0.160	0.257	0.165	0.155	0.225
-0.10	0.420	0.473	0.650	0.330	0.370	0.485
-0.15	0.753	0.823	0.940	0.555	0.580	0.765
-0.20	0.930	0.950	0.985	0.815	0.855	0.925

表 4 在显著性水平 $\alpha = 0.05$ 下, 例 4.1 中三种检验方法的功效函数值

结构	c	Case I			Case II		
		Ind	Ex	Ad	Ind	Ex	Ad
Exch	-0.00	0.065	0.055	0.058	0.070	0.050	0.060
	-0.05	0.075	0.115	0.120	0.080	0.105	0.110
	-0.10	0.168	0.315	0.380	0.095	0.225	0.270
	-0.15	0.273	0.585	0.663	0.250	0.425	0.450
	-0.20	0.518	0.838	0.890	0.350	0.615	0.660
AR(1)	-0.00	0.050	0.048	0.060	0.043	0.048	0.057
	-0.05	0.073	0.093	0.092	0.055	0.090	0.090
	-0.10	0.185	0.233	0.250	0.108	0.170	0.180
	-0.15	0.283	0.375	0.425	0.208	0.308	0.333
	-0.20	0.468	0.588	0.655	0.325	0.430	0.508

表 5 在显著性水平 $\alpha = 0.05$ 下, 例 4.2 中三种检验方法的功效函数值

5 实证分析

我们将本文的方法应用到纵向孕酮数据中, 该数据的具体表述见文 [22]. 这项研究收集了 34 名女性的尿样. 每个妇女的重复观测次数从 11 到 28 变化且这组数据共有 492 个观测值. 已有一些学者利用线性模型研究这组数据 [2, 23]. 先前的文献旨在研究血清孕酮水平的对数与年龄 (Age), 身体质量指数 (BMI) 和时间 (t_{ij}) 的关系. 文 [2, 23] 利用 Huber 得分函数分析这组数据. 尽管有界的 Huber 得分函数可以减弱响应变量中异常值的影响从而得到稳健估计, 但是稳健的 Huber 得分函数相对于非稳健的估计会损失一些效率在没有异常值的情形下. 由于稳健的秩回归方法是分布自由的, 稳健的且高度有效的. 本文将利用秩回归研究血清孕酮水平的对数与 Age, BMI, t_{ij} 以及它们的交互项的关系. 具体地, 考虑如下模型

$$\begin{aligned} Y_{ij} = & \beta_0 + \beta_1 \text{Age}_i + \beta_2 \text{BMI}_i + \beta_3 t_{ij} + \beta_4 \text{Age}_i * t_{ij} + \beta_5 \text{BMI}_i * t_{ij} \\ & + \beta_6 t_{ij}^2 + \beta_7 \text{Age}_i * \text{BMI}_i + \varepsilon_{ij}. \end{aligned}$$

比较本文所提方法 ($\tilde{\beta}_R$) 与 $\hat{\beta}_I$, $\hat{\beta}_G^e$, $\hat{\beta}_G^a$ 和 $\hat{\beta}_F$. 表 6 列举了估计的系数和对应的标准差 (SE), 其中 SE 由 500 次 bootstrap 抽样获得.

	Age	BMI	Time	Age*Time	BMI*Time	Time*Time	Age*BMI
$\hat{\beta}_I$	Estimate	0.5465	0.5413	-1.6767	0.0211	0.0336	0.0048
	SE	5.1661	7.7701	1.4750	0.0332	0.0362	0.0239
$\hat{\beta}_G^e$	Estimate	-0.7967	-1.2043	-2.1248	0.0318	0.0305	0.0040
	SE	4.9974	7.5917	1.4805	0.0325	0.0355	0.0254
$\hat{\beta}_G^a$	Estimate	-0.7742	-1.2545	-3.2770	0.0500	0.0660	-0.0132
	SE	4.8613	7.3947	1.5693**	0.0334	0.0400*	0.0229
$\hat{\beta}_F$	Estimate	-0.9201	-1.3088	-1.6648	0.0235	0.0315	0.0025
	SE	5.1902	7.8760	1.5241	0.0343	0.0373	0.0243
$\tilde{\beta}_R$	Estimate	-3.1476	-4.4049	-3.8920	0.0513	0.0823	-0.0032
	SE	4.5837	7.1284	1.8600**	0.0321	0.0405**	0.0215

表 6 孕酮数据中所得估计的系数 ($\times 10^{-2}$) 以及相应的标准差 ($SE \times 10^{-2}$)

*代表估计在 $\alpha = 10\%$ 下显著; **代表估计在 $\alpha = 5\%$ 下显著.

由表 6 可得如下结论:

(1) $\hat{\beta}_G^a$ 和 $\tilde{\beta}_R$ 得到类似的模型, 它们把 Time 和 BMI*Time 看作重要变量, 其表明 Age 和 BMI 对孕激素水平没有显著影响. 这些结果和文 [2] 的一致.

(2) 与其他四种方法相比, 本文所提方法具有更小的标准差, 这表明本文提出的方法更有效. 此外估计的自回归系数为

$$\hat{\theta}_1 = 0.2180_{(0.0293)}, \quad \hat{\theta}_2 = -0.3822_{(0.0644)},$$

其中括号内的值为标准差, 表明估计的自回归系数是高度显著的. 图 4 画出了估计的更新方差 $\hat{\sigma}(t)$ 曲线、以及相应的 95% 置信区间, 表明 $\hat{\sigma}$ 随着时间变化.

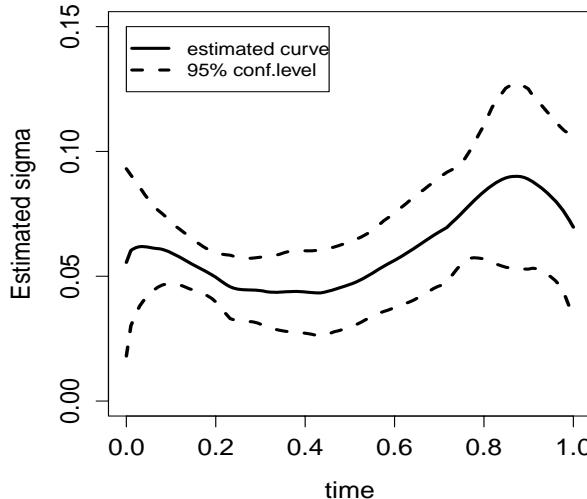


图 4 $\hat{\sigma}(t)$ 的曲线图 (实线) 以及相应的 95% 置信区间 (点线)

此外, 随机地把数据分为两部分: 训练集 (随机抽取 33 个妇女的观测值) 和测试集 (剩下的 1 个妇女的观测值). 利用绝对预测误差的均值 (MAPE) 来衡量预测效果, 其定义为

$$|Y_{ik} - \hat{Y}_{ik}|, \quad i \in D_T, \quad k = 1, \dots, m_i$$

的平均数, 其中 D_T 是测试集. 得到五种估计的 MAPE 如下:

$$\text{MAPE}(\hat{\beta}_I) = 0.2022, \quad \text{MAPE}(\hat{\beta}_G^e) = 0.2007, \quad \text{MAPE}(\hat{\beta}_G^a) = 0.2000,$$

$$\text{MAPE}(\hat{\beta}_F) = 0.2018; \quad \text{MAPE}(\tilde{\beta}_R) = 0.1969.$$

容易发现 $\tilde{\beta}_R$ 具有最小的 MAPE, 其表明本文所提方法具有最好的预测能力.

6 结论

本文利用修正的 Cholesky 分解处理纵向秩回归的组内相关性, 提出了更加有效的秩回归得分函数. 为了克服不光滑估计函数带来的计算困难, 用文 [1] 提出的诱导光滑获得回归参数的估计. 进一步, 基于有效的估计函数 (2.9) 提出新的秩得分检验, 并证明了检验统计量的渐近分布. 模拟研究和实证分析已经验证了所提方法的优越性.

A 附录

令 $\boldsymbol{\Gamma} = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \mathbf{D}_i^T \hat{\boldsymbol{\Sigma}}_i^{-1} \bar{\mathbf{D}}_i$ 和 $\boldsymbol{\Psi} = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \mathbf{D}_i^T \hat{\boldsymbol{\Sigma}}_i^{-1} \boldsymbol{\Sigma}_i \hat{\boldsymbol{\Sigma}}_i^{-1} \mathbf{D}_i$. 为了获得所提估计的渐近性质, 本文需要以下正则条件.

(C1) ε_{ik} 是连续的随机变量 $i = 1, \dots, n, k = 1, \dots, m_i$, 以及假定 $\varepsilon_{ik}(\boldsymbol{\beta}_0) - \varepsilon_{jl}(\boldsymbol{\beta}_0)$ 的分布的中位数为 0, 对于任意一对 (i, j) 和 (k, l) 包括 $i = j$, 但是 $(i, k) \neq (j, l)$.

(C2) 真实值 $\boldsymbol{\beta}_0$ 是参数空间 \mathcal{B} 的内点, 其中 \mathcal{B} 是 \mathbb{R}^p 的紧子集.

(C3) $f_{ikjl}(\cdot)$ 和 $F_{ikjl}(\cdot)$ 分别表示 $\varepsilon_{ik}(\boldsymbol{\beta}_0) - \varepsilon_{jl}(\boldsymbol{\beta}_0)$ 的密度函数和分布函数, 且 $f_{ikjl}(\cdot)$ 和它的一阶导数 $f'_{ikjl}(\cdot)$ 是在 0 处有界的函数对于 $i = 1, \dots, n, j = 1, \dots, n, k = 1, \dots, m_i, l = 1, \dots, m_j$.

(C4) $\max_{1 \leq i \leq n} m_i$ 有界, $\sup_i \|\mathbf{X}_i\| < +\infty$ 和 Γ 是一个正定矩阵, 其中 $\|\cdot\|$ 表示欧几里得范数.

(C5) $\boldsymbol{\Omega}$ 是一个正定矩阵以及 $\boldsymbol{\Omega} = O(1/n)$.

(C6) $\tilde{\mathbf{U}}_R(\boldsymbol{\beta})$ 的导数, $\frac{\partial \tilde{\mathbf{U}}_R(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}}$ 依概率 1 为正定矩阵.

(C7) 令 $\mathbf{v}_{ik} = (\sum_{l=1}^{k-1} S_{il} W_{k,l,1}^{(i)}, \dots, \sum_{l=1}^{k-1} S_{il} W_{k,l,s}^{(i)})^T$, 则 $(N-n)^{-1} \sum_{i=1}^n \sum_{k=2}^{m_i} \mathbf{v}_{ik} \mathbf{v}_{ik}^T \rightarrow_p \boldsymbol{\Lambda} > 0$.

(C8) 对于充分大的 n , $\mathbf{V}_n = n^{-1} \sum_{i=1}^n \mathbf{D}_{i(1)}^T \hat{\boldsymbol{\Sigma}}_i^{-1} \mathbf{S}_i(\boldsymbol{\beta}_0) \mathbf{S}_i^T(\boldsymbol{\beta}_0) \hat{\boldsymbol{\Sigma}}_i^{-1} \mathbf{D}_{i(1)}$ 的最小特征值远离 0 有界.

定理 2.1 的证明 根据 $\hat{\mathbf{v}}_{ik}$ 的定义可得

$$\begin{aligned} \hat{\mathbf{v}}_{ik} - \mathbf{v}_{ik} &= \left(\sum_{l=1}^{k-1} \hat{S}_{il} W_{k,l,1}^{(i)}, \dots, \sum_{l=1}^{k-1} \hat{S}_{il} W_{k,l,s}^{(i)} \right)^T - \left(\sum_{l=1}^{k-1} S_{il} W_{k,l,1}^{(i)}, \dots, \sum_{l=1}^{k-1} S_{il} W_{k,l,s}^{(i)} \right)^T \\ &= \left(\sum_{l=1}^{k-1} (\hat{S}_{il} - S_{il}) W_{k,l,1}^{(i)}, \dots, \sum_{l=1}^{k-1} (\hat{S}_{il} - S_{il}) W_{k,l,s}^{(i)} \right)^T, \end{aligned}$$

其中

$$\hat{S}_{ik} \triangleq \hat{S}_{ik}(\hat{\boldsymbol{\beta}}_I) = N^{-1} \sum_{j=1}^n \sum_{l=1}^{m_j} [I\{\hat{\varepsilon}_{jl}(\hat{\boldsymbol{\beta}}_I) < \hat{\varepsilon}_{ik}(\hat{\boldsymbol{\beta}}_I)\} - 0.5].$$

类似于文 [9] 可知, $\hat{\boldsymbol{\beta}}_I$ 是 $\boldsymbol{\beta}_0$ 的 \sqrt{n} 相合估计, 则 $\hat{S}_{il} - S_{il} = O_p(n^{-1/2})$ 和 $\hat{\mathbf{v}}_{ik} - \mathbf{v}_{ik} = O_p(n^{-1/2})$. 因此可得

$$\frac{1}{N-n} \sum_{i=1}^n \sum_{k=2}^{m_i} (\hat{\mathbf{v}}_{ik} - \mathbf{v}_{ik})(\hat{\mathbf{v}}_{ik} - \mathbf{v}_{ik})^T = O_p(n^{-1}). \quad (\text{A.1})$$

类似地

$$\begin{aligned} \frac{1}{N-n} \sum_{i=1}^n \sum_{k=2}^{m_i} (\hat{\mathbf{v}}_{ik} - \mathbf{v}_{ik}) \hat{\mathbf{v}}_{ik}^T &= \frac{1}{N-n} \sum_{i=1}^n \sum_{k=2}^{m_i} (\hat{\mathbf{v}}_{ik} - \mathbf{v}_{ik}) \mathbf{v}_{ik}^T \\ &\quad + \frac{1}{N-n} \sum_{i=1}^n \sum_{k=2}^{m_i} (\hat{\mathbf{v}}_{ik} - \mathbf{v}_{ik})(\hat{\mathbf{v}}_{ik} - \mathbf{v}_{ik})^T \\ &= \frac{1}{N-n} \sum_{i=1}^n \sum_{k=2}^{m_i} (\hat{\mathbf{v}}_{ik} - \mathbf{v}_{ik}) \mathbf{v}_{ik}^T + O_p(n^{-1}) \end{aligned}$$

和

$$\begin{aligned} \left| \frac{1}{N-n} \sum_{i=1}^n \sum_{k=2}^{m_i} (\hat{\mathbf{v}}_{ik} - \mathbf{v}_{ik}) \mathbf{v}_{ik}^T \right| &\leq o_p(1) \frac{1}{N-n} \sum_{i=1}^n \sum_{k=2}^{m_i} \left(\left| \sum_{l=1}^{k-1} S_{il} W_{k,l,1}^{(i)} \right|, \dots, \left| \sum_{l=1}^{k-1} S_{il} W_{k,l,s}^{(i)} \right| \right) \\ &= o_p(1), \end{aligned}$$

则

$$\frac{1}{N-n} \sum_{i=1}^n \sum_{k=2}^{m_i} (\hat{\mathbf{v}}_{ik} - \mathbf{v}_{ik}) \hat{\mathbf{v}}_{ik}^T = o_p(1).$$

利用如下事实

$$\begin{aligned} \sum_{i=1}^n \sum_{k=2}^{m_i} \hat{\mathbf{v}}_{ik} \hat{\mathbf{v}}_{ik}^T &= \sum_{i=1}^n \sum_{k=2}^{m_i} \{(\hat{\mathbf{v}}_{ik} - \mathbf{v}_{ik}) + \mathbf{v}_{ik}\} \{(\hat{\mathbf{v}}_{ik}^T - \mathbf{v}_{ik}^T) + \mathbf{v}_{ik}^T\} \\ &= \sum_{i=1}^n \sum_{k=2}^{m_i} \{(\hat{\mathbf{v}}_{ik} - \mathbf{v}_{ik})(\hat{\mathbf{v}}_{ik} - \mathbf{v}_{ik})^T + (\hat{\mathbf{v}}_{ik} - \mathbf{v}_{ik}) \hat{\mathbf{v}}_{ik}^T + \mathbf{v}_{ik} (\hat{\mathbf{v}}_{ik} - \mathbf{v}_{ik})^T + \mathbf{v}_{ik} \hat{\mathbf{v}}_{ik}^T\}. \end{aligned}$$

结合条件 (C7) 得

$$\frac{1}{N-n} \sum_{i=1}^n \sum_{k=2}^{m_i} \hat{\mathbf{v}}_{ik} \hat{\mathbf{v}}_{ik}^T = \frac{1}{N-n} \sum_{i=1}^n \sum_{k=2}^{m_i} \mathbf{v}_{ik} \mathbf{v}_{ik}^T + o_p(1) \xrightarrow{p} \boldsymbol{\Lambda}. \quad (\text{A.2})$$

当 $n \rightarrow \infty$. 此外

$$\begin{aligned} \sum_{i=1}^n \sum_{k=2}^{m_i} \hat{\mathbf{v}}_{ik} \hat{S}_{ik} &= \sum_{i=1}^n \sum_{k=2}^{m_i} \{(\hat{\mathbf{v}}_{ik} - \mathbf{v}_{ik})(\hat{S}_{ik} - S_{ik}) + (\hat{\mathbf{v}}_{ik} - \mathbf{v}_{ik})S_{ik} + \mathbf{v}_{ik}(\hat{S}_{ik} - S_{ik}) + \mathbf{v}_{ik}S_{ik}\} \\ &\triangleq \sqrt{N-n}(I_1 + I_2 + I_3 + I_4). \end{aligned}$$

类似于 (A.1) 的证明, 有 $I_1 = o_p(1)$, $I_2 = o_p(1)$ 和 $I_3 = o_p(1)$. 进一步,

$$I_4 = \frac{1}{\sqrt{N-n}} \sum_{i=1}^n \sum_{k=2}^{m_i} \mathbf{v}_{ik} S_{ik} = \frac{1}{\sqrt{N-n}} \sum_{i=1}^n \sum_{k=2}^{m_i} \mathbf{v}_{ik} \mathbf{v}_{ik}^T \boldsymbol{\theta} + \frac{1}{\sqrt{N-n}} \sum_{i=1}^n \sum_{k=2}^{m_i} \mathbf{v}_{ik} e_{ik}.$$

接下来只需证明

$$\frac{1}{\sqrt{N-n}} \sum_{i=1}^n \sum_{k=2}^{m_i} \mathbf{v}_{ik} e_{ik} \xrightarrow{d} N(\mathbf{0}, \boldsymbol{\Delta}). \quad (\text{A.3})$$

因此, 结合 (A.2) 和 (A.3), 利用 Slutsky 定理可得

$$\sqrt{N-n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{d} N(\mathbf{0}, \boldsymbol{\Lambda}^{-1} \boldsymbol{\Delta} \boldsymbol{\Lambda}^{-1}).$$

接下来证明 (A.3). 对于任意的 $s \times 1$ 维元素不全为 0 的常数向量 $\boldsymbol{\kappa} = (\kappa_1, \dots, \kappa_s)^T$. 令

$$\boldsymbol{\Psi} = (N-n)^{-1} \sum_{i=1}^n \sum_{k=2}^{m_i} (\boldsymbol{\kappa}^T \mathbf{v}_{ik}) e_{ik}.$$

很容易得到 $E(\boldsymbol{\Psi}) = 0$ 以及

$$\text{Var}(\sqrt{N-n} \boldsymbol{\Psi}) = \frac{1}{N-n} \sum_{i=1}^n E \left\{ \sum_{k=2}^{m_i} (\boldsymbol{\kappa}^T \mathbf{v}_{ik}) e_{ik} \right\}^2 = \frac{1}{N-n} \sum_{i=1}^n \sum_{k=2}^{m_i} d_{ik}^2 E \left(\sum_{l=1}^s \kappa_l v_{ikl} \right)^2.$$

令 $\xi_i = \sum_{k=2}^{m_i} (\boldsymbol{\kappa}^T \mathbf{v}_{ik}) e_{ik}$, 则

$$S^2 = \text{Var} \left(\sum_{i=1}^n \xi_i \right) = \sum_{i=1}^n \text{Var}(\xi_i) = \sum_{i=1}^n \sum_{k=2}^{m_i} d_{ik}^2 E(\boldsymbol{\kappa}^T \mathbf{v}_{ik})^2.$$

若李雅普诺夫 (Lyapunov) 条件

$$\frac{\sum_{i=1}^n E|\xi_i|^3}{S^3} \rightarrow 0 \quad (\text{A.4})$$

成立, 则 (A.3) 成立. 现在只需证明 (A.4) 成立. 利用条件 (C3), 有

$$\frac{\sum_{i=1}^n E|\xi_i|^3}{S^3} \leq \frac{CnE(\sum_{l=1}^s |\kappa_l| |v_{ikl}|)^3}{n^{3/2} [E\{\sum_{l=1}^s \kappa_l v_{ikl}\}]^{3/2}} = O_p(n^{-1/2}) = o_p(1).$$

因此, 完成了定理 2.1 的证明.

定理 2.2 的证明 类似于 Fan 和 Yao 文 [3, 定理 1], 可得

$$\hat{d}^2(t) - d^2(t) = \frac{1}{Nf_T(t)} \sum_{i=1}^n \sum_{k=1}^{m_i} K_h(t_{ik} - t) \{ \hat{\epsilon}_{ik}^2 - d^2(t) - \hat{d}^2(t)(t_{ik} - t) \}.$$

注意到

$$\begin{aligned}\hat{\epsilon}_{ik}^2 &= (\hat{S}_{ik} - \hat{\boldsymbol{v}}_{ik}^T \hat{\boldsymbol{\theta}})^2 = (\hat{S}_{ik} - \hat{\boldsymbol{v}}_{ik}^T \hat{\boldsymbol{\theta}} + S_{ik} - S_{ik})^2 \\ &= [\hat{S}_{ik} - S_{ik} - (\hat{\boldsymbol{v}}_{ik}^T \hat{\boldsymbol{\theta}} - \boldsymbol{v}_{ik}^T \boldsymbol{\theta}) + d_{ik} \varsigma_{ik}]^2 \\ &= d_{ik}^2 \varsigma_{ik}^2 + 2d_{ik} \varsigma_{ik} [\hat{S}_{ik} - S_{ik} - (\hat{\boldsymbol{v}}_{ik}^T \hat{\boldsymbol{\theta}} - \boldsymbol{v}_{ik}^T \boldsymbol{\theta})] + [\hat{S}_{ik} - S_{ik} - (\hat{\boldsymbol{v}}_{ik}^T \hat{\boldsymbol{\theta}} - \boldsymbol{v}_{ik}^T \boldsymbol{\theta})]^2.\end{aligned}$$

由此可得

$$\tilde{d}^2(t) - d^2(t) = I_1 + I_2 + I_3 + I_4 \{1 + o_p(1)\}, \quad (\text{A.5})$$

其中

$$\begin{aligned}I_1 &= \frac{1}{N f_T(t)} \sum_{i=1}^n \sum_{k=1}^{m_i} K_h(t_{ik} - t) \{d^2(t_{ik}) - d^2(t) - \ddot{d}^2(t)(t_{ik} - t)\}, \\ I_2 &= \frac{1}{N f_T(t)} \sum_{i=1}^n \sum_{k=1}^{m_i} K_h(t_{ik} - t) \{d^2(t_{ik})(\varsigma_{ik}^2 - 1)\}, \\ I_3 &= 2 \frac{1}{N f_T(t)} \sum_{i=1}^n \sum_{k=1}^{m_i} K_h(t_{ik} - t) d_{ik} \varsigma_{ik} [\hat{S}_{ik} - S_{ik} - (\hat{\boldsymbol{v}}_{ik}^T \hat{\boldsymbol{\theta}} - \boldsymbol{v}_{ik}^T \boldsymbol{\theta})], \\ I_4 &= \frac{1}{N f_T(t)} \sum_{i=1}^n \sum_{k=1}^{m_i} K_h(t_{ik} - t) [\hat{S}_{ik} - S_{ik} - (\hat{\boldsymbol{v}}_{ik}^T \hat{\boldsymbol{\theta}} - \boldsymbol{v}_{ik}^T \boldsymbol{\theta})]^2.\end{aligned}$$

很容易发现定理 2.2 可以直接由如下的 (a)–(d) 得到:

- (a) $I_1 = \frac{1}{2} \mu_2 h^2 \ddot{d}^2(t) + o_p(h^2);$
- (b) $\sqrt{Nh} I_2 \xrightarrow{d} N(0, \Xi);$
- (c) $I_3 = o_p(\frac{1}{\sqrt{Nh}});$
- (d) $I_4 = o_p(\frac{1}{\sqrt{Nh}}).$

(a) 直接根据泰勒展开得到. 利用中心极限定理, I_2 漐近服从正态分布其均值为 0 且方差为

$$\text{Var}(I_2) = \frac{\nu_0}{N^2 h f_T(t)} \sum_{i=1}^n \sum_{k=1}^{m_i} E[(\varsigma_{ik}^2 - 1)^2 | t_{ik} = t] d^4(t).$$

根据 I_3 的定义, 有

$$\begin{aligned}I_3 &= 2 \frac{1}{N f_T(t)} \sum_{i=1}^n \sum_{k=1}^{m_i} K_h(t_{ik} - t) d_{ik} \varsigma_{ik} [\hat{S}_{ik} - S_{ik}] \\ &\quad - 2 \frac{1}{N f_T(t)} \sum_{i=1}^n \sum_{k=1}^{m_i} K_h(t_{ik} - t) d_{ik} \varsigma_{ik} (\hat{\boldsymbol{v}}_{ik}^T \hat{\boldsymbol{\theta}} - \boldsymbol{v}_{ik}^T \boldsymbol{\theta}) \\ &\triangleq I_{31} + I_{32}.\end{aligned}$$

因为 $\hat{\beta}_I$ 是 β_0 的 \sqrt{n} 相合估计, 可得

$$\hat{S}_{ik} - S_{ik} = O_p(1/\sqrt{n}), \quad (\text{A.6})$$

利用 (A.6) 可得

$$(\hat{\boldsymbol{v}}_{ik}^T - \boldsymbol{v}_{ik}^T) \boldsymbol{\theta} = \left(\sum_{l=1}^{k-1} (\hat{S}_{il} - S_{il}) W_{k,l,1}^{(i)}, \dots, \sum_{l=1}^{k-1} (\hat{S}_{il} - S_{il}) W_{k,l,s}^{(i)} \right) \boldsymbol{\theta} = O_p(1/\sqrt{n}). \quad (\text{A.7})$$

进一步, 根据定理 2.1 以及 (A.6), (A.7), 有

$$\begin{aligned}\hat{\mathbf{v}}_{ik}^T \hat{\boldsymbol{\theta}} - \mathbf{v}_{ik}^T \boldsymbol{\theta} &= (\hat{\mathbf{v}}_{ik}^T - \mathbf{v}_{ik}^T) \boldsymbol{\theta} + (\hat{\mathbf{v}}_{ik}^T - \mathbf{v}_{ik}^T)(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) + \mathbf{v}_{ik}^T(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \\ &= O_p(n^{-1/2}) + O_p(n^{-1/2})O_p(n^{-1/2}) + O_p(n^{-1/2}) = O_p(n^{-1/2}).\end{aligned}\quad (\text{A.8})$$

另一个方面, 根据 $E(\varsigma_{ik} | t_{ik}) = 0$, $\text{Var}(\varsigma_{ik} | t_{ik}) = 1$, 结合 (A.6) 和 (A.8) 有

$$I_{31} = o_p\left(\frac{1}{\sqrt{Nh}}\right), \quad I_{32} = o_p\left(\frac{1}{\sqrt{Nh}}\right).$$

则 $I_3 = o_p(1/\sqrt{Nh})$. 类似于 I_3 的证明, 可得 $I_4 = o_p(1/\sqrt{Nh})$. 在条件 $h \rightarrow 0$, $Nh \rightarrow \infty$, $n \rightarrow \infty$ 和 $\limsup_{n \rightarrow \infty} Nh^5 < \infty$ 下, 可完成定理 2.2 的证明.

定理 2.5 的证明 令 $\mathbf{G}_i^T = \mathbf{D}_i^T \hat{\boldsymbol{\Sigma}}_i^{-1}$, 则 $\mathbf{U}_R(\boldsymbol{\beta}) = \sum_{i=1}^n \mathbf{G}_i^T \mathbf{S}_i$. 进一步假定

$$\mathbf{P}_i = \left(N^{-1} \sum_{j=1}^n \sum_{l=1}^{m_j} [p\{\varepsilon_{jl} < \varepsilon_{ik}\} - 1/2], \dots, N^{-1} \sum_{j=1}^n \sum_{l=1}^{m_j} [p\{\varepsilon_{jl} < \varepsilon_{im_i}\} - 1/2] \right)^T,$$

$\bar{\mathbf{U}}_R(\boldsymbol{\beta}) = \sum_{i=1}^n \mathbf{G}_i^T \mathbf{P}_i$, 可得

$$\begin{aligned}n^{-1} [\mathbf{U}_R(\boldsymbol{\beta}) - \bar{\mathbf{U}}_R(\boldsymbol{\beta})] &= n^{-1} \sum_{i=1}^n \mathbf{G}_i^T (\mathbf{S}_i - \mathbf{P}_i) \\ &= (nN)^{-1} \sum_{i=1}^n \sum_{k=1}^{m_i} \mathbf{g}_{ik} \sum_{j=1}^n \sum_{l=1}^{m_j} [I\{\varepsilon_{jl} < \varepsilon_{ik}\} - p\{\varepsilon_{jl} < \varepsilon_{ik}\}],\end{aligned}$$

其中 \mathbf{g}_{ik} 是 $p \times 1$ 的向量以及 $\mathbf{G}_i^T = (\mathbf{g}_{i1}, \dots, \mathbf{g}_{im_i})$. 利用条件 (C4) 和一致强大数定律^[13], 可得

$$\sup_{\boldsymbol{\beta} \in \mathcal{B}} \left| (nN)^{-1} \sum_{i=1}^n \sum_{k=1}^{m_i} \mathbf{g}_{ik} \sum_{j=1}^n \sum_{l=1}^{m_j} [I\{\varepsilon_{jl} < \varepsilon_{ik}\} - p\{\varepsilon_{jl} < \varepsilon_{ik}\}] \right| \rightarrow o(n^{-1/2}) \text{ a.s.},$$

则 $\sup_{\boldsymbol{\beta} \in \mathcal{B}} \|n^{-1} \{\mathbf{U}_R(\boldsymbol{\beta}) - \bar{\mathbf{U}}_R(\boldsymbol{\beta})\}\| = o(n^{-1/2})$ a.s.. 现在考虑 $\bar{\mathbf{U}}_R(\boldsymbol{\beta})$,

$$n^{-1} \mathbf{D}_R(\boldsymbol{\beta}_0) = n^{-1} \frac{\partial \bar{\mathbf{U}}_R(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} \Big|_{\boldsymbol{\beta}=\boldsymbol{\beta}_0} = n^{-1} \sum_{i=1}^n \mathbf{D}_i^T \hat{\boldsymbol{\Sigma}}_i^{-1} \bar{\mathbf{D}}_i.$$

利用条件 (C6) 可知 $n^{-1} \mathbf{D}_R(\boldsymbol{\beta}_0)$ 是一个正定矩阵 $N^{-1} \sum_{j=1}^n \sum_{l=1}^{m_j} p\{\varepsilon_{jl}(\boldsymbol{\beta}_0) < \varepsilon_{ik}(\boldsymbol{\beta}_0)\} = 0.5$. 由条件 (C1), 则 $\boldsymbol{\beta}_0$ 是估计方程 $\bar{\mathbf{U}}_R(\boldsymbol{\beta}) = \mathbf{0}$ 的唯一解. 因为 $\hat{\boldsymbol{\beta}}_R$ 是估计方程 $\mathbf{U}_R(\boldsymbol{\beta}) = \mathbf{0}$ 的解, 则 $\hat{\boldsymbol{\beta}}_R \rightarrow \boldsymbol{\beta}_0$ 当 $n \rightarrow +\infty$. 对于任意的 $\boldsymbol{\beta}$ 满足 $\|\boldsymbol{\beta} - \boldsymbol{\beta}_0\| < cn^{-1/3}$,

$$\begin{aligned}\mathbf{U}_R(\boldsymbol{\beta}) - \mathbf{U}_R(\boldsymbol{\beta}_0) &= \sum_{i=1}^n \mathbf{G}_i^T(\boldsymbol{\beta}) \mathbf{S}_i(\boldsymbol{\beta}) - \sum_{i=1}^n \mathbf{G}_i^T(\boldsymbol{\beta}_0) \mathbf{S}_i(\boldsymbol{\beta}_0) \\ &= \sum_{i=1}^n \mathbf{G}_i^T(\boldsymbol{\beta}) \{\mathbf{S}_i(\boldsymbol{\beta}) - \mathbf{S}_i(\boldsymbol{\beta}_0)\} + \sum_{i=1}^n \{\mathbf{G}_i(\boldsymbol{\beta}) - \mathbf{G}_i(\boldsymbol{\beta}_0)\}^T \mathbf{S}_i(\boldsymbol{\beta}_0) \\ &= I_1 + I_2.\end{aligned}$$

第一项可以记为

$$\begin{aligned}I_1 &= \sum_{i=1}^n \mathbf{G}_i^T(\boldsymbol{\beta}) \mathbf{P}_i(\boldsymbol{\beta}) + \sum_{i=1}^n \mathbf{G}_i^T(\boldsymbol{\beta}) \{\mathbf{S}_i(\boldsymbol{\beta}) - \mathbf{S}_i(\boldsymbol{\beta}_0) - \mathbf{P}_i(\boldsymbol{\beta})\} \\ &= \sum_{i=1}^n \mathbf{G}_i^T(\boldsymbol{\beta}) \mathbf{P}_i(\boldsymbol{\beta}) + \sum_{i=1}^n \sum_{k=1}^{m_i} \mathbf{g}_{ik} N^{-1} \sum_{j=1}^n \sum_{l=1}^{m_j} [I\{\varepsilon_{jl}(\boldsymbol{\beta}) < \varepsilon_{ik}(\boldsymbol{\beta})\} - I\{\varepsilon_{jl}(\boldsymbol{\beta}_0) < \varepsilon_{ik}(\boldsymbol{\beta}_0)\}] \\ &\quad - p\{\varepsilon_{jl}(\boldsymbol{\beta}) < \varepsilon_{ik}(\boldsymbol{\beta})\} + 1/2] \\ &= I_{11} + I_{12}.\end{aligned}$$

根据文 [8, 引理 3], 可得 $\sup |I_{12}| = o_p(\sqrt{n})$. 因此第一项为

$$I_1 = \sum_{i=1}^n \mathbf{G}_i^T(\boldsymbol{\beta}) \mathbf{P}_i(\boldsymbol{\beta}) + o_p(\sqrt{n}) = \bar{U}_R(\boldsymbol{\beta}) + o_p(\sqrt{n}).$$

由文 [13] 的强大数定律, 第二项为

$$I_2 = \sum_{i=1}^n \sum_{k=1}^{m_i} \{\mathbf{g}_{ik}(\boldsymbol{\beta}) - \mathbf{g}_{ik}(\boldsymbol{\beta}_0)\} N^{-1} \sum_{j=1}^n \sum_{l=1}^{m_j} [I\{\varepsilon_{jl}(\boldsymbol{\beta}_0) < \varepsilon_{ik}(\boldsymbol{\beta}_0)\} - 1/2] = o_p(\sqrt{n}).$$

因此, $\mathbf{U}_R(\boldsymbol{\beta}) - \mathbf{U}_R(\boldsymbol{\beta}_0) = \bar{U}_R(\boldsymbol{\beta}) + o_p(\sqrt{n})$. 将 $\bar{U}_R(\boldsymbol{\beta})$ 进行泰勒展开结合 $\bar{U}_R(\boldsymbol{\beta}_0) = 0$, 可得

$$\mathbf{U}_R(\boldsymbol{\beta}) - \mathbf{U}_R(\boldsymbol{\beta}_0) = \frac{\partial \bar{U}_R(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} \Big|_{\boldsymbol{\beta}=\boldsymbol{\beta}_0} (\boldsymbol{\beta} - \boldsymbol{\beta}_0) + o_p(\sqrt{n}) = \mathbf{D}_R(\boldsymbol{\beta}_0)(\boldsymbol{\beta} - \boldsymbol{\beta}_0) + o_p(\sqrt{n}).$$

因为 $\hat{\boldsymbol{\beta}}$ 在 $\boldsymbol{\beta}_0$ 的 $n^{-1/3}$ 领域内以及 $\mathbf{U}_R(\hat{\boldsymbol{\beta}}_R) = \mathbf{0}$, 有

$$\sqrt{n}(\hat{\boldsymbol{\beta}}_R - \boldsymbol{\beta}_0) = -(n^{-1}\mathbf{D}_R(\boldsymbol{\beta}_0))^{-1}n^{-1/2}\mathbf{U}_R(\boldsymbol{\beta}_0) + o_p(1).$$

根据大数定律, $n^{-1}\mathbf{D}_R(\boldsymbol{\beta}_0) \xrightarrow{P} \boldsymbol{\Gamma}$. \mathbf{S}_i 是独立的随机变量其均值为 0 和

$$\text{Var}(n^{-1/2}\mathbf{U}_R(\boldsymbol{\beta}_0)) = n^{-1} \sum_{i=1}^n \mathbf{D}_i^T \hat{\boldsymbol{\Sigma}}_i^{-1} \text{Cov}(\mathbf{S}_i(\boldsymbol{\beta}_0)) \hat{\boldsymbol{\Sigma}}_i^{-1} \mathbf{D}_i.$$

利用中心极限定理可得 $n^{-1/2}\mathbf{U}_R(\boldsymbol{\beta}_0) \xrightarrow{d} N(\mathbf{0}, \boldsymbol{\Psi})$. 进一步, 根据 Slutsky 定理可得

$$\sqrt{n}(\hat{\boldsymbol{\beta}}_R - \boldsymbol{\beta}_0) \xrightarrow{d} N(\mathbf{0}, \boldsymbol{\Gamma}^{-1}\boldsymbol{\Psi}(\boldsymbol{\Gamma}^T)^{-1}).$$

这就完成了定理 2.5 的证明.

定理 2.6 的证明 因为 $\tilde{S}_{ik} - S_{ik} = N^{-1} \sum_{j=1}^n \sum_{l=1}^{m_j} \text{sgn}(-\zeta_{ikjl}) \Phi(-|\zeta_{ikjl}|)$, 其中 $\text{sgn}(\cdot)$ 是符号函数, $\zeta_{ikjl} = \varepsilon_{ikjl}/r_{ikjl}$. 可得

$$n^{-1/2}[\tilde{\mathbf{U}}_R(\boldsymbol{\beta}) - \mathbf{U}_R(\boldsymbol{\beta})] = n^{-1/2}N^{-1} \sum_{i=1}^n \sum_{k=1}^{m_i} \sum_{j=1}^n \sum_{l=1}^{m_j} \mathbf{g}_{ik} \text{sgn}(-\zeta_{ikjl}) \Phi(-|\zeta_{ikjl}|),$$

因为

$$\begin{aligned} E(\tilde{S}_{ik} - S_{ik}) &= \int_{-\infty}^{+\infty} N^{-1} \sum_{j=1}^n \sum_{l=1}^{m_j} \Phi(-|\epsilon|/r_{ikjl}) \{2I(\epsilon \leq 0) - 1\} f_{ikjl}(\epsilon) d\epsilon \\ &= N^{-1} \int_{-\infty}^{+\infty} \sum_{j=1}^n \sum_{l=1}^{m_j} r_{ikjl} \Phi(-|t|) \{2I(t \leq 0) - 1\} \{f_{ikjl}(0) + f'_{ikjl}(\xi(t)) r_{ikjl} t\} dt, \end{aligned}$$

其中 $\xi(t)$ 介于 0 到 $r_{ikjl}t$. 因为 $\int_{-\infty}^{+\infty} \Phi(-|t|) \{2I(t \leq 0) - 1\} dt = 0$, 则

$$N^{-1} \int_{-\infty}^{+\infty} \sum_{j=1}^n \sum_{l=1}^{m_j} r_{ikjl} \Phi(-|t|) \{2I(t \leq 0) - 1\} f_{ikjl}(0) dt = 0.$$

由条件 (C3) 存在一个常数 C , 使得 $\sup_{ikjl} |f'_{ikjl}(\xi(t))| \leq C$ 和 $\int_{-\infty}^{+\infty} \Phi(-|t|) |t| dt = 1/2$. 因此

$$\begin{aligned} |E(\tilde{S}_{ik} - S_{ik})| &\leq N^{-1} \sum_{j=1}^n \sum_{l=1}^{m_j} r_{ikjl}^2 \int_{-\infty}^{+\infty} |t| \Phi(-|t|) |f'_{ikjl}(\xi(t))| dt \\ &\leq N^{-1} \sum_{j=1}^n \sum_{l=1}^{m_j} C r_{ikjl}^2 / 2. \end{aligned}$$

由条件 (C4) 和 (C5), 当 $n \rightarrow \infty$, 有

$$\|n^{-1/2}E\{\tilde{\mathbf{U}}_R(\boldsymbol{\beta}) - \mathbf{U}_R(\boldsymbol{\beta})\}\| \leq n^{-1/2} \sum_{i=1}^n \sum_{k=1}^{m_i} \sup_{i,k} |\mathbf{g}_{ik}| N^{-1} \sum_{j=1}^n \sum_{l=1}^{m_j} C r_{ikjl}^2 / 2 = o(1).$$

此外

$$n^{-1}\text{Var}[\tilde{\mathbf{U}}_R(\boldsymbol{\beta}) - \mathbf{U}_R(\boldsymbol{\beta})] = n^{-1} \sum_{i=1}^n \text{Var} \left\{ \sum_{k=1}^{m_i} \mathbf{g}_{ik} (\tilde{S}_{ik} - S_{ik}) \right\}.$$

根据 Cauchy-Schwartz 不等式

$$\begin{aligned} \frac{1}{n} \text{Var}[\tilde{\mathbf{U}}_R(\boldsymbol{\beta}) - \mathbf{U}_R(\boldsymbol{\beta})] &\leq \frac{1}{n} \sum_{i=1}^n \sum_{k=1}^{m_i} \mathbf{g}_{ik} \mathbf{g}_{ik}^T \text{Var}(\tilde{S}_{ik} - S_{ik}) \\ &+ \frac{1}{n} \sum_{i=1}^n \sum_{k=1}^{m_i} \sum_{k' \neq k} \mathbf{g}_{ik} \mathbf{g}_{ik'}^T \sqrt{\text{Var}(\tilde{S}_{ik} - S_{ik}) \text{Var}(\tilde{S}_{ik'} - S_{ik'})}, \end{aligned}$$

对于任意的 $k = 1, \dots, m_i$, 有

$$\begin{aligned} \text{Var}(\tilde{S}_{ik} - S_{ik}) &\leq E(\tilde{S}_{ik} - S_{ik})^2 \\ &= \int_{-\infty}^{+\infty} \left\{ N^{-1} \sum_{j=1}^n \sum_{l=1}^{m_j} \text{sgn}(-\epsilon/r_{ikjl}) \Phi(-|\epsilon/r_{ikjl}|) \right\} f_{ikjl}(\epsilon) d\epsilon \\ &\leq N^{-1} \sum_{j=1}^n \sum_{l=1}^{m_j} r_{ikjl} \int_{-\infty}^{+\infty} \Phi^2(-|t|) f_{ikjl}(r_{ikjl} t) dt \\ &= N^{-1} \sum_{j=1}^n \sum_{l=1}^{m_j} r_{ikjl} \int_{|t|>\Delta} \Phi^2(-|t|) f_{ikjl}(r_{ikjl} t) dt \\ &\quad + N^{-1} \sum_{j=1}^n \sum_{l=1}^{m_j} r_{ikjl} \int_{|t|\leq\Delta} \Phi^2(-|t|) f_{ikjl}(r_{ikjl} t) dt \\ &\leq N^{-1} \sum_{j=1}^n \sum_{l=1}^{m_j} r_{ikjl} \Phi^2(-\Delta) + 2N^{-1} \sum_{j=1}^n \sum_{l=1}^{m_j} r_{ikjl} \Delta f_{ikjl}(\xi), \end{aligned}$$

其中 Δ 是一个正数和 ξ 介于 $-r_{ikjl}\Delta$ 到 $r_{ikjl}\Delta$. 令 $\Delta = n^{1/3}$. 由条件 (C5) 可知 $r_{ikjl} = O(n^{-1/2})$, 则 $r_{ikjl}\Delta = O(n^{-1/6})$. 进一步可得

$$N^{-1} \sum_{j=1}^n \sum_{l=1}^{m_j} r_{ikjl} \Phi^2(-\Delta) \quad \text{和} \quad N^{-1} \sum_{j=1}^n \sum_{l=1}^{m_j} r_{ikjl} \Delta f_{ikjl}(\xi)$$

都趋于 0, 当 $n \rightarrow \infty$. 由条件 (C2) 和 (C4), 很容易得到 $n^{-1}\text{Var}[\tilde{\mathbf{U}}_R(\boldsymbol{\beta}) - \mathbf{U}_R(\boldsymbol{\beta})] = o(1)$. 因此, 对于任意的 $\boldsymbol{\beta}$, 当 $n \rightarrow \infty$ 时, 有 $n^{-1/2}[\tilde{\mathbf{U}}_R(\boldsymbol{\beta}) - \mathbf{U}_R(\boldsymbol{\beta})] \rightarrow 0$.

定理 2.7 的证明 根据定理 2.6 和 $\sup_{\boldsymbol{\beta} \in \mathcal{B}} \|n^{-1}[\tilde{\mathbf{U}}_R(\boldsymbol{\beta}) - \mathbf{U}_R(\boldsymbol{\beta})]\| = o(n^{-1/2})$ a.s. 以及三角不等式, 有

$$\sup_{\boldsymbol{\beta} \in \mathcal{B}} \|n^{-1}[\tilde{\mathbf{U}}_R(\boldsymbol{\beta}) - \bar{\mathbf{U}}_R(\boldsymbol{\beta})]\| = o(n^{-1/2}).$$

因为 $\boldsymbol{\beta}_0$ 是 $\bar{\mathbf{U}}_R(\boldsymbol{\beta}) = 0$ 的唯一解. 再根据 $\tilde{\boldsymbol{\beta}}_R$ 的定义, 当 $n \rightarrow \infty$ 时, 可得 $\tilde{\boldsymbol{\beta}}_R \rightarrow \boldsymbol{\beta}_0$. 为了证明

$\tilde{\beta}_R$ 的渐近正态性, 首先证明 $n^{-1}\{\tilde{\mathbf{D}}_R(\beta_0) - \mathbf{D}_R(\beta_0)\} \xrightarrow{p} 0$, 其中

$$\begin{aligned}\tilde{\mathbf{D}}_R(\beta) &= \frac{\partial \tilde{\mathbf{U}}_R(\beta)}{\partial \beta} = \sum_{i=1}^n \mathbf{D}_i^T \hat{\Sigma}_i^{-1} \tilde{\mathbf{D}}_i, \quad \tilde{\mathbf{D}}_i = (\tilde{\mathbf{D}}_{i1}, \dots, \tilde{\mathbf{D}}_{im_i})^T, \\ \tilde{\mathbf{D}}_{ik} &= \frac{\partial \tilde{S}_{ik}(\beta)}{\partial \beta} = N^{-1} \sum_{j=1}^n \sum_{l=1}^{m_j} r_{ikjl}^{-1} \phi(\varepsilon_{ikjl}/r_{ikjl})(\mathbf{X}_{jl} - \mathbf{X}_{ik}).\end{aligned}$$

令 \mathbf{g}_{ik} 是 $p \times 1$ 的向量和 $\mathbf{G}_i^T = (\mathbf{g}_{i1}, \dots, \mathbf{g}_{im_i})$, 可得

$$\begin{aligned}E\{\tilde{\mathbf{D}}_R(\beta_0)\} - \mathbf{D}_R(\beta_0) &= \sum_{i=1}^m \sum_{j=1}^{n_i} \mathbf{g}_{ik} \left\{ N^{-1} \sum_{j=1}^n \sum_{l=1}^{m_j} [r_{ikjl}^{-1} E\phi(\varepsilon_{ikjl}/r_{ikjl}) - f_{ikjl}(0)] \right\} (\mathbf{X}_{jl} - \mathbf{X}_{ik}).\end{aligned}$$

因为

$$\begin{aligned}|r_{ikjl}^{-1} E\phi(\varepsilon_{ikjl}/r_{ikjl}) - f_{ikjl}(0)| &= \left| r_{ikjl}^{-1} \int_{-\infty}^{+\infty} \phi(\epsilon/r_{ikjl}) f_{ikjl}(\epsilon) d\epsilon - f_{ikjl}(0) \right| \\ &= \left| \int_{-\infty}^{+\infty} \phi(t) f_{ikjl}(0) + r_{ikjl} t f'_{ikjl}(\zeta_t) dt - f_{ikjl}(0) \right| \\ &= \left| r_{ikjl} \int_{-\infty}^{+\infty} \phi(t) t f'_{ikjl}(\zeta_t) dt \right| \\ &\leq r_{ikjl} \int_{-\infty}^{+\infty} |\phi(t) t f'_{ikjl}(\zeta_t)| dt,\end{aligned}$$

其中 ζ_t 介于 0 到 $r_{ikjl} t$. 根据条件 (C3), $f'_{ikjl}(\cdot)$ 一致有界, 因此存在一个常数 C , 使得 $|f'_{ikjl}(\zeta_t)| \leq C$, 再结合条件 (C5), 有

$$|r_{ikjl}^{-1} E\phi(\varepsilon_{ikjl}/r_{ikjl}) - f_{ikjl}(0)| \leq \sqrt{\frac{2}{\pi}} r_{ikjl} C \rightarrow 0.$$

则 $|n^{-1}\{\tilde{\mathbf{D}}_R(\beta_0) - \mathbf{D}_R(\beta_0)\}| \rightarrow 0$. 由强大数定律有 $n^{-1}\tilde{\mathbf{D}}_R(\beta_0) \rightarrow E\{n^{-1}\tilde{\mathbf{D}}_R(\beta_0)\}$. 利用三角不等式, 则

$$|n^{-1}\{\tilde{\mathbf{D}}_R(\beta_0) - \mathbf{D}_R(\beta_0)\}| \leq |n^{-1}\{\tilde{\mathbf{D}}_R(\beta_0) - E\tilde{\mathbf{D}}_R(\beta_0)\}| + |n^{-1}\{E\tilde{\mathbf{D}}_R(\beta_0) - \mathbf{D}_R(\beta_0)\}| \rightarrow 0.$$

将 $\tilde{\mathbf{U}}_R(\beta)$ 在 β_0 处泰勒展开, 有

$$\tilde{\mathbf{U}}_R(\beta) = \tilde{\mathbf{U}}_R(\beta_0) + \tilde{\mathbf{D}}_R(\beta^*)(\beta - \beta_0),$$

其中 β^* 介于 β 到 β_0 . 令 $\beta = \tilde{\beta}_R$. 因为 $\tilde{\mathbf{U}}_R(\tilde{\beta}_R) = \mathbf{0}$ 和 $\tilde{\beta}_R \rightarrow \beta_0$, 可得 $\beta^* \rightarrow \beta_0$ 和 $\tilde{\mathbf{D}}_R(\beta^*) \rightarrow \tilde{\mathbf{D}}_R(\beta_0)$. 根据定理 2.6 和 $n^{-1}\{\tilde{\mathbf{D}}_R(\beta_0) - \mathbf{D}_R(\beta_0)\} = o_p(1)$, 有

$$\sqrt{n}(\tilde{\beta}_R - \beta_0) = -(n^{-1}\mathbf{D}_R(\beta_0))^{-1} n^{-1/2} \mathbf{U}_R(\beta_0) + o_p(1).$$

应用定理 2.5, 很容易完成定理 2.7 的证明.

定理 3.1 的证明 $S_n = n^{-1/2} \sum_{i=1}^n \mathbf{D}_{i(1)}^T \hat{\Sigma}_i^{-1} \mathbf{S}_i(\beta_0)$ 表示 n 个独立随机变量的和, 其均值为 0 方差为

$$\text{Cov}(\mathbf{S}_n) = n^{-1} \sum_{i=1}^n \text{Cov}(\mathbf{D}_{i(1)}^T \hat{\Sigma}_i^{-1} \mathbf{S}_i(\beta_0)) = n^{-1} \sum_{i=1}^n \mathbf{D}_{i(1)}^T \hat{\Sigma}_i^{-1} \Sigma_i \hat{\Sigma}_i^{-1} \mathbf{D}_{i(1)} \triangleq \mathbf{V}_n.$$

根据条件 (C8) 以及中心极限定理

$$\mathbf{V}_n^{-1/2} \mathbf{S}_n \xrightarrow{d} N(\mathbf{0}, I_{p_1}). \quad (\text{A.9})$$

由于 $\bar{\beta}$ 是 β_0 的 \sqrt{n} 相合估计, 即 $\|\bar{\beta} - \beta_0\| = O_p(n^{-1/2})$. 因此, 类似于文 [18, 引理 8.4 和 8.5] 中的证明, 可得

$$n^{-1/2} \left\| \sum_{i=1}^n \mathbf{D}_{i(1)}^T \hat{\Sigma}_i^{-1} \{ \mathbf{S}_i(\bar{\beta}) - \mathbf{S}_i(\beta_0) \} \right\| = o_p(1). \quad (\text{A.10})$$

类似于文 [14, 定理 2] 中的证明, 可以证明在原假设 H_0 下, $\hat{V}_n - V_n = o_p(1)$, 结合 (A.9) 和 (A.10) 可以完成定理 3.1 的证明.

参 考 文 献

- [1] Brown B. M., Wang Y. G., Standard errors and covariance matrices for smoothed rank estimators, *Biometrika*, 2005, **92**: 149–158.
- [2] Fan Y., Qin G., Zhu Z., Variable selection in robust regression models for longitudinal data, *J. Multivariate Anal.*, 2012, **109**: 156–167.
- [3] Fan J., Yao Q., Efficient estimation of conditional variance functions in stochastic regression, *Biometrika*, 1998, **85**: 645–660.
- [4] Fu L., Wang Y. G., Efficient Estimation for Rank-Based Regression with Clustered Data, *Biometrics*, 2012, **68**: 1074–1082.
- [5] Fu L., Wang Y. G., Bai Z., Rank regression for analysis of clustered data: a natural induced smoothing approach, *Comput. Statist. Data Anal.*, 2010, **54**: 1036–1050.
- [6] Gutenbrunner C., Jurečková J., Koenker R., et al., Tests of linear hypotheses based on regression rank scores, *J. Nonpara. Statist.*, 1993, **2**: 307–331.
- [7] Hettmansperger T. P., Statistical Inference Based on Ranks, John Wiley and Sons, New York, 1984.
- [8] Jung S. H., Quasi-likelihood for median regression models, *J. Amer. Statist. Assoc.*, 1996, **91**: 251–257.
- [9] Jung S. H., Ying Z., Rank-based regression with repeated measurements data, *Biometrika*, 2003, **90**: 732–740.
- [10] Leng C., Zhang W., Pan J., Semiparametric mean-covariance regression analysis for longitudinal data, *J. Amer. Statist. Assoc.*, 2010, **105**: 181–193.
- [11] Liang K. Y., Zeger S. L., Longitudinal data analysis using generalized linear models, *Biometrika*, 1986, **73**: 13–22.
- [12] Liu S., Li G., Varying-coefficient mean-covariance regression analysis for longitudinal data, *J. Statist. Plann. Inference*, 2015, **160**: 89–106.
- [13] Pollard D., Empirical Processes: Theories and Applications, Institut. Statist. Mathe., Hayward, California, 1990.
- [14] Wang L., Kai B., Li R., Local rank inference for varying coefficient models, *J. Amer. Statist. Assoc.*, 2009, **104**: 1631–1645.
- [15] Wang L., Li R., Weighted Wilcoxon-type smoothly clipped absolute deviation method, *Biometrics*, 2009, **65**: 564–571.
- [16] Wang Y. G., Zhao Y. D., Weighted rank regression for clustered data analysis, *Biometrics*, 2008, **64**: 39–45.
- [17] Wang Y. G., Zhu M., Rank-based regression for analysis of repeated measures, *Biometrika*, 2006, **93**: 459–464.
- [18] Wei Y., He X., Conditional growth charts (with discussion), *Ann. Statist.*, 2006, **34**: 2069–2097.
- [19] Yang H., Guo C., Lv J., SCAD penalized rank regression with a diverging number of parameters, *J. Multivariate Anal.*, 2015, **133**: 321–333.
- [20] Yao W., Li R., New local estimation procedure for a non-parametric regression function for longitudinal data, *J. Roy. Statist. Soc. Ser. B*, 2013, **75**: 123–138.
- [21] Ye H., Pan J., Modelling of covariance structures in generalised estimating equations for longitudinal data, *Biometrika*, 2006, **93**: 927–941.
- [22] Zhang D., Lin X., Raz J., et al., Semiparametric stochastic mixed models for longitudinal data, *J. Amer. Statist. Assoc.*, 1998, **93**: 710–719.
- [23] Zheng X., Fung W., Zhu Z., Robust estimation in joint mean-covariance regression model for longitudinal data, *Ann. Inst. Statist. Math.*, 2013, **65**: 617–638.