Advances in Applied Mathematics 应用数学进展, 2023, 12(9), 3991-3998
Published Online September 2023 in Hans. https://www.hanspub.org/journal/aam
https://doi.org/10.12677/aam.2023.129390

Hans汉斯

高维混合效应模型的统计推断：一种新的方法

赵宏媛

青岛大学数学与统计学院，山东 青岛

收稿日期：2023年8月12日；录用日期：2023年9月6日；发布日期：2023年9月13日

摘　要

线性混合效应模型广泛应用于分析聚类或重复测量数据。本文提出了一种拟似然结合弹性网的方法来估计高维线性混合模型中的未知参数，包括固定效应及随机效应的方差分量部分。在此基础上，也提出了相关的统计推断。所提出的方法适用于一般设置，其中随机效应的维度和簇可能很大。关于固定效应，我们提供的方法不依赖于方差分量的结构信息，即对方差分量中所涉及到的复杂未知参数使用代理矩阵进行简化。并且对所提出的方法在各种模拟设置中分别进行了固定效应的误差，假设检验的性能以及方差分量的估计误差的评估，均表现出较优的结果。

关键词

聚类数据，弹性网，纵向数据，随机效应，方差分量

Statistical Inference of High-Dimensional Mixed Effects Models: A New Approach

Hongyuan Zhao

School of Mathematics and Statistics, Qingdao University, Qingdao Shandong

Received: Aug. 12th, 2023; accepted: Sep. 6th, 2023; published: Sep. 13th, 2023

Abstract

The linear mixed-effects model is widely used for analyzing clustered or repeated measurements data. This paper proposes a pseudo-likelihood method combined with the elastic net to estimate unknown parameters in high-dimensional linear mixed models, including the variance components of both fixed and random effects. Furthermore, relevant statistical inferences are also presented. The proposed method is applicable to general settings where the dimension and clusters of random effects can be substantial. Regarding fixed effects, our approach does not rely on structural information about the variance components. Instead, it simplifies the complex unknown pa-

文章引用: 赵宏媛. 高维混合效应模型的统计推断: 一种新的方法[J]. 应用数学进展, 2023, 12(9): 3991-3998.
DOI: 10.12677/aam.2023.129390

rameters involved in the variance components using surrogate matrices. The performance of the proposed method is evaluated for fixed effects errors, hypothesis testing, and variance component estimation errors in various simulation settings, all of which demonstrate superior results.

Keywords

Clustered Data, Elastic Net, Longitudinal Data, Random Effects, Variance Components

1. 引言

集群数据通常出现在许多领域，如生物学、遗传学和经济学。线性混合效应模型提供了一个灵活的工具来分析这样的集群数据，其中包括重复测量数据，纵向数据和多级数据等[1] [2]。线性混合效应模型包含固定和随机效应。在许多基因和经济研究中，协变量的维数可以很大，可能远远大于样本的大小。许多学者提出了各种各样的统计模型和方法，来研究和分析高维数据。然而，大多数方法都局限于处理独立观察的数据，如线性模型和广义线性模型。对高维线性混合效应模型进行统计推断仍然是一个具有挑战性的问题。在这项工作中，我们考虑高维混合效应模型中未知参数的估计和推断。

当维度固定时，很多方法被提出用来估计固定效应和方差项，如 Gumedze 和 Dunne 在 2011 年应用极大似然估计和限制性极大似然估计于混合效应模型的参数估计[3]，但遗憾的是，它们不适用于高维情况。Zhang 和 Tong 在 2007 年提出了随机效应变系数模型的固定效应和方差参数的矩估计[4]。Peng 和 Lu (2012)考虑了固定维线性混合效应模型的矩估计[5]。他们都偏向于固定维设置。Ahmn，Zhang 和 Lu (2012)提出了另一种基于矩的方法[6]，该方法用于在固定维设置中估计和选择随机效应的方差分量。当各个集群大小相同时，此方法很适用。

对于固定维设置中方差分量的推断，可以采用似然比、得分和 Wald 检验等[7] [8] [9] [10] (Stram & Lee, 1994; Lin 1997; Verbeke & Molenberghs, 2003; Demidenko, 2004)。但是这些方法均基于极大似然估计以及限制性极大似然估计，故也存在缺陷。

在高维中，Fan 和 Li (2012) [11]研究了高维线性混合效应模型中，当聚类大小平衡时，固定效应和随机效应的选择无问题，选择变量的一致性要求关于固定效应和随机效应的最小信号强度条件。Bradic、Claeskens 和 Gueuning (2020) [12]考虑在具有固定聚类大小、固定数量的随机效应和亚高斯设计的高维线性混合效应模型中测试固定效应的单个系数。但是他们的理论分析都要求随机效应的协方差矩阵具有正定性。并且在高维情况下，对方差分量的估计仍是未知的。基于此，Li 等人[13]在 2022 年提出了一种针对于高维混合效应模型的拟似然参数估计方法，该方法基于 Lasso 并加入了拟似然来处理随机效应的影响。

而在本文中，我们开发了一种新的基于拟似然及弹性网的方法来推断高维线性混合效应模型中的未知参数。我们的方法适用于随机效应的数量可以是大的，集群大小可以是固定的或增长，平衡或不平衡的设置。该方法易于实现，每一步的优化都是解析的或凸的。基于真实协方差矩阵的代理，我们开发了一个惩罚准似然方法的固定效应估计。进一步开发了一个去偏估计的假设检验和建设的固定效应的置信区间。并且基于拟似然方法估计了方差分量。

本文的剩余部分组织如下：在第 2 节，我们介绍了混合效应模型且进行了相应的符号说明。第 3 节我们提出了基于弹性网的相关参数的拟似然估计方法。第 4 节我们展示了仿真实验结果，并将其进行了对比，最后，在第 5 节进行了总结和讨论。

2. 模型及符号

我们使用聚类数据的设置来呈现线性混合效应模型。对于重复测量数据，重复测量形成聚类。令 $i=1,\cdots,n$ 表示类群标签。对于第 i 个类，响应向量为 $y_i\in R^{m_i}$，固定效应的设计矩阵为 $X^i\in R^{m_i\times p}$，随机效应的设计矩阵为 $Z_i\in R^{m_i\times q}$。从而线性混合效应模型[14]形式如下所示：

$$y_i=X^i\beta^*+Z^i\gamma_i+\varepsilon_i,i=1,\cdots,n, \tag{1}$$

其中，$\beta^*\in R^p$ 是固定效应的系数向量，$\gamma_i\in R^q$ 是第 i 个类的随机效应的系数向量，而 ε_i 是第 i 个类的噪声向量。对于 $i=1,\cdots,n$，我们假设 γ_i 和 ε_i 独立同分布，且均值为 0，方差分别为 $\psi\in R^{q\times q}$ 和 $\sigma_e^2I_{m_i}$。定义 $N=\sum_{i=1}^n m_i$ 表示总体样本量。我们将 γ_i 和 ε_i 统称为随机部分。

本文，我们使用 i 表示第 i 个聚类，k 表示每个聚类中的第 k 个观测值。y,γ,ε 和 X 分别由 $y_i,\gamma_i,\varepsilon_i$ 和 X^i 得到。令 $Z\in R^{N\times nq}$ 是对角块矩阵，第 i 个块为 Z^i。令 $\Sigma_\theta^i=Z^i\psi Z^i+\sigma_e^2I_{m_i}$，且 $\Sigma_\theta\in R^{N\times N}$ 是第 i 个块为 Σ_θ^i 的对角块矩阵。有 $\Sigma_z^i=\left(Z^i\right)^{\mathrm{T}}Z^i/m_i$ 和 $\Sigma_{z,x}^i=\left(Z^i\right)^{\mathrm{T}}X^i/m_i,i=1,\cdots,n$。对于随机变量 $u\in R$，定义它的亚高斯范数为 $\|u\|_{\psi_2}=\sup_{1\le l}l^{-1/2}E^{1/l}\left[|u|^l\right]$。我们称 $\|u\|_{\psi_2,Z}=\sup_{1\le l}l^{-1/2}E^{1/l}\left[|u|^l\mid Z\right]$ 为 u 的条件亚高斯范数。对于随机向量 $U\in R^{n_0}$，定义其亚高斯范数为 $\|U\|_{\psi_2}=\sup_{\|v\|_2=1,v\in R^{n_0}}\|\langle U,v\rangle\|_{\psi_2}$，而相应的条件亚高斯范数为 $\|U\|_{\psi_2,Z}=\sup_{\|v\|_2=1,v\in R^{n_0}}\|\langle U,v\rangle\|_{\psi_2,Z}$。

令 $A\in R^{n_0\times n_0}$ 是对称矩阵，$A\succcurlyeq 0$ 表示矩阵 A 非负定，$A\succ 0$ 表示正定，$\Lambda_{\max}(A)$ 和 $\Lambda_{\min}(A)$ 分别表示矩阵 A 的最大和最小特征值，令 $\|A\|_2$ 定义 $\Lambda_{\max}(A)$，$\|A\|_1=\max_j\sum_{i=1}^{n_0}|a_{i,j}|$，$\|A\|_F=Tr\left(A^{\mathrm{T}}A\right)$，其中 $Tr(A)$ 表示 A 的迹。令 $c,c_0,c_1,\cdots,C,C_0,C_1,\cdots$ 表示在不同情况下可能变化的一些通用的正的常数。

3. 参数估计

引理 1 (Lasso 的收敛速度)：假设响应 y_i 是由模型(1)生成的，X 的每一行都是在以 Z 为条件下由协方差矩阵 $\Sigma_{x|z}$ 独立生成的。则对任意固定的 $j\in\{1,\cdots,p\}$，有

$$E\left[\left|\frac{1}{N}X_{\cdot,j}^{\mathrm{T}}\left(Z\gamma+\varepsilon\right)\right|^2\mid Z\right]=\frac{\left(\Sigma_{x|z}\right)_{j,j}\sigma_e^2}{N}+\frac{\left(\Sigma_{x|z}\right)_{j,j}\sum_{i=1}^n m_iTr\left(\Psi\Sigma_z^i\right)}{N^2}+\frac{\sum_{i=1}^n m_i^2\left\|\Psi^{1/2}E\left[\Sigma_{z,x}^i\mid Z\right]\right\|_2^2}{N^2}.$$

考虑一般的弹性网作用于模型(1)，我们有：

$$\beta^{(lm)}=\arg\min_{b\in R^p}\left\{\frac{1}{2N}\|y-Xb\|_2^2+\lambda_2^{(lm)}\|b\|_2+\lambda_1^{(lm)}\|b\|_1\right\}. \tag{2}$$

由引理 1，当 q 或者 m_i 增大且 X 和 Z 相关时，Lasso 的收敛速度对于聚类数据不是最优，从而，弹性网的收敛速度也非最优。因此，我们需要考虑引入新的估计方法。

基于此，我们考虑引入一种新的拟似然方法，将随机部分的方差记为：$\Sigma_{\theta^*}^i=Z^i\psi Z^i+\sigma_e^2I_{m_i}$，该式包含较难估计的未知参数。因此，考虑 $\Sigma_{\theta^*}^i$ 的代理矩阵：

$$\Sigma_a^i=aZ^iZ^i+I_{m_i},$$

其中，$a>0$ 是某预定常数。记 $\Sigma_a\in R^{N\times N}$ 为第 i 个块为 Σ_a^i 的对角块矩阵。

引理 2：若矩阵 Ψ 为正定矩阵，则 $\forall a>0$，

$$\min\left\{\frac{1}{\sigma_e^2},\frac{a}{\Lambda_{\max}(\Psi)}\right\}\Sigma_a^{-1}\preccurlyeq\Sigma_{\theta^*}^{-1}\preccurlyeq\max\left\{\frac{1}{\sigma_e^2},\frac{a}{\Lambda_{\min}(\Psi)}\right\}\Sigma_a^{-1},$$

因此，当 Ψ 的特征值有界且为正数时，Σ_a^{-1} 和 $\Sigma_{\theta^*}^{-1}$ 有相同的收敛速度。

接下来，提出具体的拟似然估计方法。使用 Σ_a 代替 Σ_{θ^*}，并且对响应向量 y 和固定效应矩阵 X 进行标准化处理，得到：

$$(X_a,y_a)=\left(\Sigma_a^{-1/2}X,\Sigma_a^{-1/2}y\right).$$

3.1. 固定效应的估计

首先基于标准化后的数据 (X_a,y_a) 对固定效应进行估计，对固定的 a，定义：

$$\hat{\beta}=\text{argmmin}_{\beta\in R^p}\left\{\frac{1}{2Tr\left(\Sigma_a^{-1}\right)}\left\|y_a-X_a\beta\right\|_2^2+\lambda_2\left\|\beta\right\|_2+\lambda_1\left\|\beta\right\|_1\right\},\tag{3}$$

其中，$\lambda_l,l=1,2$ 是调优参数，$Tr\left(\Sigma_a^{-1}\right)$ 为有效样本量，可以看作是 Lasso 和岭回归的线性组合。根据 Zou 等人(2017 年) [15]，我们定义 $\lambda=\lambda_1+\lambda_2$，$\alpha=\frac{\lambda_2}{\lambda_1}$，从而(3)式变为，

$$\hat{\beta}=\text{argmmin}_{\beta\in R^p}\left\{\frac{1}{2Tr\left(\Sigma_a^{-1}\right)}\left\|y_a-X_a\beta\right\|_2^2+\lambda\left(\alpha\left\|\beta\right\|_2+(1-\alpha)\left\|\beta\right\|_1\right)\right\}.\tag{4}$$

进一步，为了对 β^* 进行统计推断，定义如下的无偏估计，

$$\hat{\beta}_j^{(db)}=\hat{\beta}_j+\frac{\hat{\omega}_j^{\mathrm{T}}\left(y_a-X_a\hat{\beta}\right)}{\hat{\omega}_j^{\mathrm{T}}\left(X_a\right)_{\cdot,j}},\tag{5}$$

定义 $\hat{\omega}_j=\left(X_a\right)_{\cdot,j}-\left(X_a\right)_{\cdot,-j}\hat{k}_j$，其中，

$$\hat{k}_j=\arg\min_{\hat{k}_j\in R^{p-1}}\left\{\frac{1}{2Tr\left(\Sigma_a^{-1}\right)}\left\|\left(X_a\right)_{\cdot,j}-\left(X_a\right)_{\cdot,-j}\hat{k}_j\right\|_2^2+\lambda_{2j}\left\|k_j\right\|_2+\lambda_{1j}\left\|k_j\right\|_1\right\}.\tag{6}$$

从而，β_j^* 双边置信区间为：

$$\hat{\beta}_j^{(db)}\pm z_{\alpha/2}\sqrt{\hat{V}_j},$$

这里，z_τ 为标准正态分布的第 τ 分位数，而 $\hat{V}_j$ 为 $\hat{\beta}_j^{(db)}$ 方差的一个估计值，有，

$$\hat{V}_j=\frac{\sum_{i=1}^n\left[\left(\hat{\omega}_j^i\right)^{\mathrm{T}}\left(y_a^i-X_a^i\hat{\beta}\right)\right]^2}{\left(\hat{\omega}_j^{\mathrm{T}}\left(X_a\right)_{\cdot,j}\right)^2}.\tag{7}$$

3.2. 方差分量的估计

接下来，我们对模型中的方差分量进行估计。考虑到 Ψ 是一个对称矩阵，故有如下分解：

$$\Psi=\Psi_{\eta^*}=\sum_{j=1}^d\eta_j^*G_j,\tag{8}$$

其中，$G_1,\cdots,G_d$ 是线性无关的对称基矩阵，且满足：

$$\sum_{j=1}^{d} c_j G_j = 0, \text{ iff } c_1 = \cdots = c_d = 0. \tag{9}$$

用于估计方差分量的方法是高斯最大似然方法。将数据分为两部分，$I_1 \cup I_2 = [n]$，$I_1 \cap I_2 = \varnothing$，且 $|I_1| \approx |I_2| \approx n/2$。令 $\hat{\beta}^{(2)}$ 是数据 $\{X^i, Z^i, y_i\}_{i \in I_2}$ 估计得到的初始值，接下来在 $i \in I_1$ 中计算残差 $\hat{r}_i = y_i - X^i \hat{\beta}^{(2)}$，并且由下式得到 σ_e^2 的估计，

$$\hat{\sigma}_e^2 = \frac{1}{\sum_{i \in I_1} Tr\left(P_{Z^i}^{\perp}\right)} \sum_{i \in I_1} \left\| P_{Z^i}^{\perp} \hat{r}_i \right\|_2^2. \tag{10}$$

现在，我们估计 η^*，

$$\hat{\eta} = \arg\min_{\eta \in R^d} \sum_{i \in I_1} \left\| \left(\Sigma_a^i\right)^{-1/2} \left(\hat{r}_i \hat{r}_i^{\mathrm{T}} - Z^i \Psi_\eta \left(Z^i\right)^{\mathrm{T}} - \hat{\sigma}_e^2 I_{n_i} \right) \left(\Sigma_a^i\right)^{-1/2} \right\|_F^2, \tag{11}$$

其中 $K \geq K_2$ 是常数，且 $\hat{\sigma}_e^2$ 由(10)式得到。

公式(10)的基本原理是观测值 $P_{Z^i}^{\perp}\left(y_i - X^i \beta^*\right)$ 的协方差矩阵为 $\sigma_e^2 P_{Z^i}^{\perp}$，它只涉及目标参数 σ_e^2。将 β^* 用它的拟似然估计来代替，得到(10)。此估计量仅当 $\sum_{i \in I_1} Tr\left(P_{Z^i}^{\perp}\right) > 0$ 时才成立，即 $\sum_{i \in I_1} m_i \max\{0, 1 - q/m_i\} > 0$。

3.3. 去偏估计量的渐近性质

条件 1 (亚高斯随机变量)：随机噪声项 $\varepsilon_{i,k}, i = 1, \cdots, n; k = 1, \cdots, m$ 独立分布于均值为 0 方差为 $0 < \sigma_e^2 < K_0 < \infty$。$\varepsilon_{i,k}$ 的亚高斯范数以 K_0 为上界。随机效应 $\gamma_i \in R^q, i = 1, \cdots, n$ 独立分布于均值为 0，方差为 $\Psi \leq K_1 I_q$，其中 $K_1 > 0$ 是常数。对于 $i = 1, \cdots, n$，ε_i 和 γ_i 独立于彼此，且独立于 $\left(X^i, Z^i\right)$。$\Sigma_{\theta^*}^{-1/2}(Z\gamma + \varepsilon)$ 的亚高斯范数以 K_0 为界。

条件 2：在给定 Z 的条件下，X 的每一行都是独立的，具有零均值和协方差矩阵 $\Sigma_{X|Z}$ 满足 $0 < K_* \leq \Lambda_{\min}\left(\Sigma_{X|Z}\right) \leq \Lambda_{\max}\left(\Sigma_{X|Z}\right) \leq K^* < \infty$。在以 Z 为条件时，$X_{k,\cdot}^i$ 的条件亚高斯范数以 K_0 为上界。

定理 1 (无偏估计量的渐近性质)：假设条件 1 和条件 2 成立。令 $\lambda_l \wedge \lambda_{lj} \geq c_1 \sqrt{\log p / Tr\left(\Sigma_a^{-1}\right)}$，$l = 1, 2$，$j = 1, 2, \cdots, p-1$，$c_1$ 是一个足够大的常数。对于在(5)式中定义的 $\hat{\beta}_j^{(db)}$，若 $(s \log p)^2 \vee \log n \max_i m_i \ll Tr\left(\Sigma_a^{-1}\right) \Lambda_{\min}\left(\Sigma_a^{-1/2} \Sigma_{\theta^*} \Sigma_a^{-1/2}\right)$ 以及 $|H_j| \log p \ll Tr\left(\Sigma_a^{-1}\right)$，则，

$$V_j^{-1/2}\left(\hat{\beta}_j^{(db)} - \beta_j^*\right) = R_j + o_p(1), \tag{12}$$

其中，$R_j \xrightarrow{D} N(0,1)$，对

$$V_j = \frac{\hat{\omega}_j^T \Sigma_a^{-1/2} \Sigma_{\theta^*} \Sigma_a^{-1/2} \hat{\omega}_j}{\left\{\hat{\omega}_j^{\mathrm{T}} \left(X_a\right)_{\cdot,j}\right\}^2}.$$

这里，V_j 的大小满足，

$$V_j = \frac{\left(\Sigma_{X|Z}^{-1}\right)_{j,j} Tr\left(\Sigma_a^{-1} \Sigma_{\theta^*} \Sigma_a^{-1}\right)}{Tr^2\left(\Sigma_a^{-1}\right)} \left(1 + o_p(1)\right).$$

4. 仿真实验

在本节中，我们进行了仿真模拟来评估所提出的方法的 power 性能，并将其与相关的方法进行比较。我们考虑和 Li 等人(2022)相同的模拟设置：令 $N = 144$，$p = 300$。(X, Z) 的每一行都独立同分布产生于均值为 0，方差如 $\Sigma_x = I_p$，$\Sigma_z = I_q$ 以及 $\left(\Sigma_{X,Z}\right)_{k,j} = \rho^j, 1 \leq j, k \leq q$，$\left(\Sigma_{X,Z}\right)_{k,\cdot} = 0, k > q$。也就是说，如果 $j \leq p$，

则 X_j 和 Z 之间的相关性为非 0，如果 $j>q$，则相关性为 0。随机噪声项独立同分布于 $\varepsilon_i \sim N\left(0,0.25I_{m_i}\right)$。我们考虑 $q\in\{2,8,14\}$，响应 y 经由模型(1)生成，其中 $s=5$ 且 $\beta_{1:5}=(1,0.5,0.2,0.1,0.05)$ 和相等的簇大小，即 $m_1=m_2=\cdots=m_n=m$。每个设置都使用 300 次独立的 Monte Carlo 模拟进行重复。

4.1. 固定效应的统计推断

我们首先检查所提出的估计量的误差。考虑两种随机效应的协方差矩阵形式，一个是正定的 Ψ，我们记为“p.d.Ψ”，其中，$\Psi_{j,k}=0.56^{|j-k|},1\le j,k\le q$，以及奇异的 Ψ，记为“singularΨ”，其中 Ψ 是对角矩阵，且 $\Psi_{j,j}=0.56,1\le j\le q/2$，否则 $\Psi_{j,j}=0$。对于所提出的方法，首先通过交叉验证选择 a，且调优参数 λ_t 通过 $\hat{\sigma}_x\sqrt{2\log p/N}$ 计算，其中，$\hat{\sigma}_x$ 由 scaled-Lasso 和观测值 $\left((X_a)_{\cdot,j},(X_a)_{\cdot,-j}\right)$ 进行计算。

从表 1 可以看出，我们所提出的估计量具有较小的估计误差，且随着 m 和 q 的增加，误差变化不大，说明我们提出的方法具有较好的稳定性。同时，随着 β_j^* 的值递减，误差也呈现递减的规律。

Table 1. Standard error for p.d.Ψ and singularΨ when $\beta_j^*\in\{1,0.5,0.2,0\}$

表 1. $\beta_j^*\in\{1,0.5,0.2,0\}$ 时在正定 Ψ 以及奇异 Ψ 下的标准误

		p.d.Ψ				singularΨ			
q	m	SD (1)	SD (0.5)	SD (0.2)	SD (0)	SD (1)	SD (0.5)	SD (0.2)	SD (0)
2	4	0.057	0.056	0.054	0.052	0.048	0.046	0.045	0.042
	8	0.054	0.050	0.048	0.045	0.047	0.045	0.042	0.039
	12	0.054	0.048	0.045	0.041	0.049	0.044	0.041	0.037
8	4	0.162	0.158	0.154	0.149	0.104	0.098	0.096	0.091
	8	0.119	0.113	0.104	0.100	0.082	0.075	0.072	0.067
	12	0.086	0.081	0.076	0.069	0.066	0.062	0.057	0.051
14	4	0.231	0.225	0.220	0.221	0.138	0.135	0.129	0.128
	8	0.201	0.194	0.183	0.182	0.126	0.122	0.111	0.108
	12	0.165	0.157	0.144	0.136	0.114	0.103	0.095	0.089

接下来，我们检查基于 $\hat{\beta}_j^{(db)}$ 的假设检验的 power 性能。考虑两个随机效应的协方差矩阵，第一个是“p.d.Ψ”，另一个为“singularΨ”，其余参数设置和之前相同。

在表 2 中，我们展示了本文提出的方法和 Li 等人(2023)提出的方法的 I 类误差和 power。我们的运行时间和 Li 等人的时间相差不大。在理想情况下，原假设 H_0 下的拒绝率应该接近 5%，并且对{1, 0.5, 0.2}的拒绝率应该大于 5%。从中可以看到我们的方法和 Li 等人的方法在控制第 I 类误差方面都是有效的。然而，在保证第 I 类错误的情况下，我们的方法相对于 Li 等人的 power 性能较好。

4.2. 方差分量的统计推断

真实的固定效应和数据生成步骤与 4.1 节相同。我们使用全部数据来估计 σ_e^2 和 η^*。令 $\sigma_e^2=0.25$，我们考虑 $d=2$ 的对角矩阵 Ψ，基矩阵的设置如下所示：

$$G_1=\begin{pmatrix} I_{q/2} & 0 \\ 0 & 0 \end{pmatrix}, G_2=\begin{pmatrix} 0 & 0 \\ 0 & I_{q/2} \end{pmatrix},$$

对于正定矩阵 Ψ，$\eta^*=(0.56,0.56)^{\mathrm{T}}$，而奇异矩阵 Ψ，有 $\eta^*=(0.56,0)^{\mathrm{T}}$。表 3 显示了 $\sigma_e^2\left(\mathrm{mae}.\sigma_e^2\right)$、

$\eta_1(\text{mae}.\eta_1)$ 和 $\eta_2(\text{mae}.\eta_2)$ 的平均绝对误差。我们可以看到，估计的方差分量的平均绝对误差保持在较小的水平，且在 Ψ 是正定的情况下，估计的平均绝度误差小于奇异情况下的。

Table 2. The rejection rate for testing H_0: $\beta_j^*=0$ at 95% level for $\beta_j^*\in\{1,0.5,0.2,0\}$ with positive definite (p.d.) and singular when $\rho=0$

表 2. 测试 H_0：$\beta_j^*=0$ 在 95%水平上的 power，$\beta_j^*\in\{1,0.5,0.2,0\}$，且 Ψ 具有正定矩阵(p.d)以及 $\rho=0$ 时的奇异矩阵(singular)两种形式

			proposed				compare			
Ψ	q	m	1	0.5	0.2	0	1	0.5	0.2	0
p.d.Ψ	2	4	1	1	0.817	0.047	1	1	0.793	0.043
		8	1	1	0.917	0.04	1	1	0.887	0.033
		12	1	1	0.95	0.04	1	1	0.94	0.033
	8	4	1	0.74	0.23	0.07	0.997	0.71	0.16	0.047
		8	1	0.947	0.277	0.083	1	0.93	0.26	0.027
		12	1	1	0.553	0.057	1	0.993	0.44	0.017
	14	4	0.927	0.44	0.133	0.04	0.917	0.403	0.13	0.04
		8	0.98	0.527	0.157	0.057	0.977	0.5	0.143	0.047
		12	1	0.73	0.17	0.047	0.993	0.623	0.147	0.05
singularΨ	2	4	1	1	0.92	0.057	1	1	0.877	0.05
		8	1	1	0.94	0.063	1	1	0.92	0.04
		12	1	1	0.97	0.07	1	1	0.947	0.04
	8	4	1	0.986	0.367	0.057	1	0.983	0.33	0.037
		8	1	1	0.527	0.057	1	0.997	0.477	0.053
		12	1	1	0.72	0.05	1	1	0.623	0.047
	14	4	1	0.87	0.29	0.057	1	0.85	0.217	0.04
		8	1	0.94	0.3	0.073	1	0.92	0.247	0.057
		12	1	0.973	0.367	0.053	1	0.96	0.28	0.033

Table 3. Estimation of the variance components with the proposed method for positive definite and singular when $\rho=0$

表 3. 当 $\rho=0$ 时，利用所提出的方法对方差分量进行了正定和奇异估计

		p.d.Ψ			singularΨ		
m	q	$\text{mae}.\sigma_e^2$	$\text{mae}.\eta_1$	$\text{mae}.\eta_2$	$\text{mae}.\sigma_e^2$	$\text{mae}.\eta_1$	$\text{mae}.\eta_2$
4	2	0.094	0.094	0.078	0.266	0.266	0.087
8	2	0.122	0.122	0.065	0.279	0.279	0.075
	4	0.117	0.125	0.072	0.119	0.518	0.090
12	2	0.143	0.143	0.065	0.281	0.281	0.074
	6	0.144	0.098	0.076	0.145	0.403	0.087

5. 总结与讨论

本文我们基于高维线性混合模型的框架，考虑了未知参数的估计和推断问题。并且，提出了一种新的结合拟似然和弹性网的方法，该方法考虑使用拟似然消去未知参数的影响，并且通过弹性网进行变量选择及降维。在建模重复测量和纵向数据，特别是当集群大小是大的或异构的时候，具有普遍适用性。我们提出的估计过程计算效率高，不需要很强的对于随机效应和误差的分布假设。仿真实验表明，我们的方法在假设检验方面的性能优于之前的方法，并且，估计量的误差也处于合理的范围内。

参考文献

[1] Pinheiro, J. and Bates, D. (2006) Mixed-Effects Models in S and S-PLUS. Springer Science & Business Media, Berlin.

[2] Goldstein, H. (2011) Multilevel Statistical Models. John Wiley & Sons, New York. https://doi.org/10.1002/9780470973394

[3] Gumedze, F.N. and Dunne, T.T. (2011) Parameter Estimation and Inference in the Linear Mixed Model. *Linear Algebra and Its Applications*, **435**, 1920-1944. https://doi.org/10.1016/j.laa.2011.04.015

[4] Sun, Y., Zhang, W.Y. and Tong, H. (2007) Estimation of the Covariance Matrix of Random Effects in Longitudinal Studies. *Annals of Statistics*, **35**, 2795-2814. https://doi.org/10.1214/009053607000000523

[5] Müller, S., Scealy, J.L. and Welsh, A.H. (2013) Model Selection in Linear Mixed Models. *Statistical Science*, **28**, 135-167. https://doi.org/10.1214/12-STS410

[6] Ahn, M., Zhang, H.H. and Lu, W. (2012) Moment-Based Method for Random Effects Selection in Linear Mixed Models. *Statistica Sinica*, **22**, 1539-1562. https://doi.org/10.5705/ss.2011.054

[7] Stram, D.O. and Lee, J.W. (1994) Variance Components Testing in the Longitudinal Mixed Effects Model. *Biometrics*, **50**, 1171-1177. https://doi.org/10.2307/2533455

[8] Lin, X. (1997) Variance Component Testing in Generalised Linear Models with Random Effects. *Biometrika*, **84**, 309-326. https://doi.org/10.1093/biomet/84.2.309

[9] Verbeke, G. and Molenberghs, G. (2003) The Use of Score Tests for Inference on Variance Components. *Biometrics*, **59**, 254-262. https://doi.org/10.1111/1541-0420.00032

[10] Vonesh, E.F. (2006) Mixed Models: Theory and Applications. *Journal of the American Statistical Association*, **101**, 1724-1726. https://doi.org/10.1198/jasa.2006.s146

[11] Fan, Y. and Li, R. (2012) Variable Selection in Linear Mixed Effects Models. *Annals of Statistics*, **40**, 2043-2068. https://doi.org/10.1214/12-AOS1028

[12] Bradic, J., Claeskens, G. and Gueuning, T. (2020) Fixed Effects Testing in High-Dimensional Linear Mixed Models. *Journal of the American Statistical Association*, **115**, 1835-1850. https://doi.org/10.1080/01621459.2019.1660172

[13] Li, S., Cai, T.T. and Li, H. (2022) Inference for High-Dimensional Linear Mixed-Effects Models: A Quasi-Likelihood Approach. *Journal of the American Statistical Association*, **117**, 1835-1846. https://doi.org/10.1080/01621459.2021.1888740

[14] Laird, N.M. and Ware, J.H. (1982) Random-Effects Models for Longitudinal Data. *Biometrics*, **38**, 963-974. https://doi.org/10.2307/2529876

[15] Zou, H. and Hastie, T. (2005) Regularization and Variable Selection via the Elastic Net. *Journal of the Royal Statistical Society Series B*: *Statistical Methodology*, **67**, 301-320. https://doi.org/10.1111/j.1467-9868.2005.00503.x