Project 3 - TMA4315

martin.o.berild@ntnu.no
10014

yaolin.ge@ntnu.no
10026

November 16, 2020

Problem 1

a) It can be shown that the matrix formulation has two steps. First, the measurement model can be rewritten as follows:

$$y_{ij} = \mathbf{x}_{ij}^T \boldsymbol{\beta} + \mathbf{u}_{ij}^T \boldsymbol{\gamma_i} + \epsilon_{ij}$$

In this case,

$$\mathbf{x}_{ij} = \begin{bmatrix} 1 \\ x_{ij} \end{bmatrix}; \quad \mathbf{u}_{ij} = \begin{bmatrix} 1 \\ x_{ij} \end{bmatrix}; \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}; \quad \boldsymbol{\gamma}_i = \begin{bmatrix} \gamma_{0,i} \\ \gamma_{1,i} \end{bmatrix}$$

By collecting all individual- and cluster-specific responses y_{ij}, design vectors $\mathbf{x}_{ij}, \mathbf{u}_{ij}$ and errors $\epsilon_{ij}, j = 1, \cdots, n_i$ into vectors, then it becomes:

$$\mathbf{y}_i = \begin{pmatrix} y_{i1} \\ \vdots \\ y_{ij} \\ \vdots \\ y_{in_i} \end{pmatrix}; \quad \mathbf{X}_i = \begin{pmatrix} \boldsymbol{x}_{i1}^T \\ \vdots \\ \boldsymbol{x}_{ij}^T \\ \vdots \\ \boldsymbol{x}_{in_i}^T \end{pmatrix}; \quad \mathbf{U}_i = \begin{pmatrix} \boldsymbol{u}_{i1}^T \\ \vdots \\ \boldsymbol{u}_{ij}^T \\ \vdots \\ \boldsymbol{u}_{in_i}^T \end{pmatrix}; \quad \boldsymbol{\epsilon}_i = \begin{pmatrix} \epsilon_{i1} \\ \vdots \\ \epsilon_{ij} \\ \vdots \\ \epsilon_{in_i} \end{pmatrix}$$

Thus, the measurement model in matrix notation is

$$\boldsymbol{y}_i = \boldsymbol{X}_i \boldsymbol{\beta} + \boldsymbol{U_i} \boldsymbol{\gamma_i} + \boldsymbol{\epsilon_i}, \quad i = 1, \cdots, m$$

Given that

$$\boldsymbol{\gamma}_i \sim \mathcal{N}(\mathbf{0}, \boldsymbol{Q}); \quad \boldsymbol{\epsilon}_i \sim \mathcal{N}(\mathbf{0}, \sigma^2 \boldsymbol{I_{n_i}})$$

Thus, squeeze the above LMM model by defining the design matrices as follows:

$$\boldsymbol{y} = \begin{pmatrix} \boldsymbol{y}_1 \\ \vdots \\ \boldsymbol{y}_i \\ \vdots \\ \boldsymbol{y}_m \end{pmatrix}; \quad \boldsymbol{\epsilon} = \begin{pmatrix} \boldsymbol{\epsilon}_1 \\ \vdots \\ \boldsymbol{\epsilon}_i \\ \vdots \\ \boldsymbol{\epsilon}_m \end{pmatrix}; \quad \boldsymbol{\gamma} = \begin{pmatrix} \boldsymbol{\gamma}_1 \\ \vdots \\ \boldsymbol{\gamma}_i \\ \vdots \\ \boldsymbol{\gamma}_m \end{pmatrix};$$

Therefore, the LMM can be rewritten as

$$\boldsymbol{y} = \boldsymbol{X} \boldsymbol{\beta} + \boldsymbol{U} \boldsymbol{\gamma} + \boldsymbol{\epsilon}$$

```r
mylmm <- function(y, x, group, REML = TRUE){
    print("hello world")

}
```

Test function

```r
library(lme4)
```

```
## Loading required package: Matrix
```

```r
head(sleepstudy)
```

```
##   Reaction Days Subject
## 1 249.5600    0     308
## 2 258.7047    1     308
## 3 250.8006    2     308
## 4 321.4398    3     308
## 5 356.8519    4     308
## 6 414.6901    5     308
```

```r
mod <- lmer(Reaction ~ 1 + Days + (1 + Days|Subject), data = sleepstudy)
summary(mod)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: Reaction ~ 1 + Days + (1 + Days | Subject)
##    Data: sleepstudy
##
## REML criterion at convergence: 1743.6
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.9536 -0.4634  0.0231  0.4634  5.1793
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  Subject  (Intercept) 612.10   24.741
##           Days         35.07    5.922   0.07
##  Residual             654.94   25.592
## Number of obs: 180, groups:  Subject, 18
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)  251.405      6.825  36.838
## Days          10.467      1.546   6.771
##
## Correlation of Fixed Effects:
##      (Intr)
## Days -0.138
```

```r
mylmm()
```

```
## [1] "hello world"
```

The the model is (for cluster $i = 1, 2, \cdots, m$):

$$y_i = X_i\beta + U_i\gamma_i + \epsilon_i, \quad \gamma_i \sim \mathcal{N}(\prime, \mathcal{Q})$$

The marginal model is then

$$y_i \sim \mathcal{N}(X_i\beta, \quad U_iQU_I^T + \sigma^2 I_{n_i})$$

Conditional model is

$$<! - - y_i|\gamma_i \sim \mathcal{N}(X_i\beta + U_i\gamma_i, \quad \sigma^2 I_{n_i}) - - >$$

We are interested in computing the maximum likelihood and restricted maximum likelihood estimates of the parameters in the linear mixed model

$$y_{ij} = \beta_0 + \beta_1 x_{ij} + \gamma_{0,i} + \gamma_{1,i}x_{ij} + \epsilon_{ij},$$

where $\boldsymbol{\gamma}_i = (\gamma_{0,i}, \gamma_{1,i})$ are iid binomially distributed with zero mean and covariance matrix

$$\begin{bmatrix} \tau_0^2 & \tau_{01} \\ \tau_{01} & \tau_1^2 \end{bmatrix},$$

and ϵ_{ij} are idd Guassian distributed with zero mean and variance σ^2 for $i = 1, \ldots, m$ and $j = 1, \ldots, n$. Thus, we are interested in obtaining the estimates of $(\beta_0, \beta_1, \tau_0^2, \tau_1^2, \tau_{01}, \sigma^2)$.

Problem 2