STAT 574 HOMEWORK 1 SOLUTION

1 (Mixed Models)**.** Use the Weight v.s. Height example. For the following cases, write down the mixed-effect models, specify the unknown parameters, and determine the number of parameters to be fitted.

(a) The weight is linear in height, but the intercepts differ for different families.
(b) The weight is linear in height, but both the intercepts and the coefficients for height differ for different families.
(c) In addition to (b), we know the expected coefficient for height is 5.30.
(d) The weight is linear in height, but the intercepts differ for different families, and the height coefficients differ for different genders.

Solution.

(a) The model is

$$W_{ij} = \beta_0 + \beta_1 H_{ij} + b_i + \epsilon_{ij},$$

where $b_i \sim \mathcal{N}(0, \sigma^2 d)$ and $\epsilon_{ij} \sim \mathcal{N}(0, \sigma^2)$. The unknown parameters are β_0, β_1, σ^2 and d. The number of parameters is, therefore, 4.

(b) The model is

$$W_{ij} = \beta_0 + \beta_1 H_{ij} + b_{0i} + b_{1i} H_{ij} + \epsilon_{ij},$$

where $(b_{0i}, b_{1i})^T \sim \mathcal{N}(\mathbf{0}, \sigma^2 \boldsymbol{D})$ and $\epsilon_{ij} \sim \mathcal{N}(0, \sigma^2)$. The unknown parameters are β_0, β_1, σ^2 and $\boldsymbol{D}$, totaling 6 parameters (because $\boldsymbol{D} \in \mathbb{R}^{2\times 2}$ has to be symmetric).

(c) The model is

$$W_{ij} = \beta_0 + 5.3 H_{ij} + b_{0i} + b_{1i} H_{ij} + \epsilon_{ij}$$

or equivalently,

$$W_{ij} - 5.3 H_{ij} = \beta_0 + b_{0i} + b_{1i} H_{ij} + \epsilon_{ij},$$

where $(b_{0i}, b_{1i})^T \sim \mathcal{N}(\mathbf{0}, \sigma^2 \boldsymbol{D})$ and $\epsilon_{ij} \sim \mathcal{N}(0, \sigma^2)$. The unknown parameters are β_0, σ^2 and $\boldsymbol{D}$, totaling 5 parameters.

(d) The model is

$$W_{ijk} = \beta_0 + \beta_1 H_{ijk} + b_{0i} + b_{1j} H_{ijk} + \epsilon_{ijk},$$

where index (i, j, k) means the k-th person in i-th family and j-th gender, and $b_{0i} \sim \mathcal{N}(0, \sigma^2 d_0)$, $b_{1i} \sim \mathcal{N}(0, \sigma^2 d_1)$ and $\epsilon_{ijk} \sim \mathcal{N}(0, \sigma^2)$. The unknown parameters are β_0, β_1, σ^2, d_0 and d_1, totaling 5 parameters.

2 (Generalized Least Squares)**.** Consider a linear model with dependent observations that

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon},$$

where $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$.

(a) Write down the log-likelihood function and calculate the MLE $\hat{\boldsymbol{\beta}}^{(MLE)}$.
(b) Consider the following weighted residual sum-of-squares.

$$\mathrm{RSS}(w_1, \ldots, w_n) = \sum_{i=1}^{n} w_i (y_i - \boldsymbol{\beta}^T \boldsymbol{x}_i)^2,$$

where $\boldsymbol{x}_i$ is the covariates for unit i, y_i is the corresponding observation, and $w_i > 0$ is the weight assigned to unit i.
Find $\hat{\boldsymbol{\beta}}^{(WLS)}$ that minimizes $\mathrm{RSS}(w_1, \ldots, w_n)$.
(c) Figure out when you can find the weights $w_1, \ldots, w_n$ such that $\hat{\boldsymbol{\beta}}^{(MLE)} = \hat{\boldsymbol{\beta}}^{(WLS)}$.

Solution.

(a) The log-likelihood function is

$$\ell(\boldsymbol{\beta}) = -\frac{1}{2}\left[\log|\boldsymbol{\Sigma}| + (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})\right]$$

By maximizing the log-likelihood function, we have

$$\hat{\boldsymbol{\beta}}^{(MLE)} = (\boldsymbol{X}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{y}$$

(b) Taking the partial derivative of RSS, we have

$$\frac{\partial \text{RSS}}{\partial \boldsymbol{\beta}} = 2\sum_{i=1}^{n} w_i \boldsymbol{\beta}^T \boldsymbol{x}_i \boldsymbol{x}_i^T - 2\sum_{i=1}^{n} w_i y_i \boldsymbol{x}_i^T$$

By setting the partial derivative to zero, we have

$$\hat{\boldsymbol{\beta}}^{(WLS)} = \left(\sum_{i=1}^{n} w_i \boldsymbol{x}_i \boldsymbol{x}_i^T\right)^{-1} \sum_{i=1}^{n} w_i y_i \boldsymbol{x}_i$$

(c) Let $W = \text{diag}(w_1, w_2, \ldots, w_n)$. We notice that

$$\hat{\boldsymbol{\beta}}^{(WLS)} = \left(\boldsymbol{X}^T\boldsymbol{W}\boldsymbol{X}\right)^{-1}\boldsymbol{X}^T\boldsymbol{W}\boldsymbol{y}.$$

Therefore, when $\boldsymbol{\Sigma}^{-1} = \boldsymbol{W}$, that is when $\boldsymbol{\Sigma}$ is diagonal and $w_i \propto 1/\Sigma_{ii}$, where Σ_{ii} is the i-th diagonal of $\boldsymbol{\Sigma}$, MLE and WLS are equal.

3 (Restricted Maximum Likelihood)**.** Consider n i.i.d. observations $x_1, \ldots, x_n \sim \mathcal{N}(\mu, \sigma^2)$.

(a) Find the MLE $\hat{\mu}$ and $\hat{\sigma}^2$.

(b) Construct $(n-1)$ new observations in the following way

$$y_1 = x_1 - x_n,\ y_2 = x_2 - x_n,\ \ldots\ , y_{n-1} = x_{n-1} - x_n.$$

Show that $(y_1, \ldots, y_{n-1})$ is independent of μ.

(c) Determine the distribution of $(y_1, \ldots, y_{n_1})$.

(d) Use above as the restricted likelihood for $\hat{\sigma}^2$, and find out $\hat{\sigma}^2_R$ that maximizes the restricted likelihood.

(e) Compare $\hat{\sigma}^2$ and $\hat{\sigma}^2_R$.

Solution.

(a) The log-likelihood function is

$$\ell(\mu, \sigma^2) = -\frac{1}{2}\left[n\log\sigma^2 + \sigma^{-2}\sum_{i=1}^{n}(x_i - \mu)^2\right].$$

The MLE is therefore

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x}, \quad \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{\mu})^2.$$

(b) $(y_1, \ldots, y_{n-1}, \hat{\mu})$ is a linear combination of $(x_1, \ldots, x_n)$, and is therefore multivariate Gaussian. It suffices to show the covariance is zero. Since

$$\text{Cov}(y_i, \bar{x}) = \text{Cov}(x_i - x_n, x_i/n + x_n/n) = \text{Var}(x_i)/n - \text{Var}(x_n)/n = 0,$$

$(y_1, \ldots, y_{n-1})$ and $\bar{\mu}$ are independent.

(c) Since $(y_1, \ldots, y_{n-1}$ is multivariate Gaussian, it suffices to determine its mean and variance. Notice that

$$\mathbb{E}[y_i] = \mathbb{E}[x_i] - \mathbb{E}[x_n] = 0$$
$$\text{Var}(y_i) = \text{Var}(x_i) + \text{Var}(x_n) = 2\sigma^2$$
$$\text{Cov}(y_i, y_j) = \text{Var}(x_n) = \sigma^2, \quad \text{for } i \neq j.$$

Hence, we have

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{n-1} \end{bmatrix} \sim \mathcal{N}\left(\mathbf{0}, \sigma^2 \mathbf{\Sigma}\right),$$

where

$$\mathbf{\Sigma} = \begin{bmatrix} 2 & 1 & \cdots & 1 \\ 1 & 2 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 2 \end{bmatrix} \in \mathbb{R}^{(n-1)\times(n-1)}$$

(d) Notice that, $\Sigma = \mathbb{I} + \mathbf{1}\mathbf{1}^T$, and by Woodbury identity, we have

$$\mathbf{\Sigma}^{-1} = \mathbb{I} - \mathbf{1}(1 + \mathbf{1}^T\mathbf{1})^{-1}\mathbf{1}^T = \mathbb{I} - n^{-1}\mathbf{1}\mathbf{1}^T.$$

Hence, the restricted log-likelihood function is

$$\ell_R(\sigma^2) = -\frac{1}{2}\left[\log|\sigma^2\mathbf{\Sigma}| + \boldsymbol{y}^T(\sigma^2\mathbf{\Sigma})^{-1}\boldsymbol{y}\right]$$

The maximum is attained at

$$\hat{\sigma}_R^2 = \frac{1}{n-1}\boldsymbol{y}^T\mathbf{\Sigma}^{-1}\boldsymbol{y} = \frac{1}{n-1}\sum_{i=1}^n x_i^2 - \frac{1}{n(n-1)}\left(\sum_{i=1}^n x_i\right)^2 = \frac{n}{n-1}\hat{\sigma}^2.$$

(e) The restricted MLE is proportional to the MLE.

4 (Programming)**.** Download the Weight v.s. Height dataset in `Family.txt` from

https://github.com/eugenedemidenko/mixedmodels

The file is in `Data/MixedModels/Chapter02/` folder. For the four models in Question 1, fit with the dataset and report your final models (specifying the parameters and the distributions).

Solution.

(a) The model is

$$W_{ij} = -206.83 + 5.35H_{ij} + b_i + \epsilon_{ij},$$

where $b_i \sim \mathcal{N}(0, 14.07^2)$ and $\epsilon_{ij} \sim \mathcal{N}(0, 24.71^2)$.

(b) The model is

$$W_{ij} = -248.83 + 6.01H_{ij} + b_{0i} + b_{1i}H_{ij} + \epsilon_{ij},$$

where

$$\begin{pmatrix} b_{0i} \\ b_{1i} \end{pmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} 237.07^2 & -0.998 \times 237.07 \times 3.63 \\ -0.998 \times 237.07 \times 3.63 & 3.63^2 \end{bmatrix}\right)$$

and $\epsilon_{ij} \sim \mathcal{N}(0, 19.81^2)$.

(c) The model is

$$W_{ij} = -201.67 + 5.3H_{ij} + b_{0i} + b_{1i}H_{ij} + \epsilon_{ij}$$

$$\begin{pmatrix} b_{0i} \\ b_{1i} \end{pmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} 223.16^2 & -0.998 \times 223.16 \times 3.42 \\ -0.998 \times 223.16 \times 3.42 & 3.42^2 \end{bmatrix}\right)$$

and $\epsilon_{ij} \sim \mathcal{N}(0, 20.04^2)$.

(d) The model is

$$W_{ijk} = -74.52 + 3.39H_{ijk} + b_{0i} + b_{1j}H_{ijk} + \epsilon_{ijk},$$

where $b_{0i} \sim \mathcal{N}(0, 13.12^2)$, $b_{1i} \sim \mathcal{N}(0, 0.22^2)$ and $\epsilon_{ijk} \sim \mathcal{N}(0, 24.07^2)$.

The code for this problem is pasted below:

```
data = read.table("./Data/MixedModels/Chapter02/Family.txt", header=T, stringsAsFactors=F)

library(nlme)
fit.a = lme(fixed=Weight~Height, random=~1|FamilyID, data=data)
print(fit.a)

fit.b = lme(fixed=Weight~Height, random=~1+Height|FamilyID, data=data)
print(fit.b)

data$ynew = data$Weight - 5.3 * data$Height
fit.c = lme(fixed=ynew~1, random=~1+Height|FamilyID, data=data)
print(fit.c)

library(lme4)
fit.d = lmer(Weight~1+Height+(1|FamilyID)+(0+Height|Sex), data=data)
print(fit.d)
```