HW7

Ran Zhang

4/15/2019

Problem 1

Problem 1a Load the data;

```
library(lattice)
library(latticeExtra)
```

```
## Loading required package: RColorBrewer
```

Read the data;

```
pollution <- read.table("airpollution_all.txt")
colnames(pollution) <- c("id", "Height", "age", "height_base", "age_base", "logFEV1")
head(pollution)
```

```
##   id Height     age height_base age_base logFEV1
## 1  1   1.20  9.3415         1.2   9.3415 0.21511
## 2  1   1.28 10.3929         1.2   9.3415 0.37156
## 3  1   1.33 11.4524         1.2   9.3415 0.48858
## 4  1   1.42 12.4600         1.2   9.3415 0.75142
## 5  1   1.48 13.4182         1.2   9.3415 0.83291
## 6  1   1.50 15.4743         1.2   9.3415 0.89200
```

Fit the model with random intercept;

```
library(nlme)
fit_1a <- lme(logFEV1 ~ age + Height, random = ~1|id, data=pollution, method = "ML")
summary(fit_1a)
```

```
## Linear mixed-effects model fit by maximum likelihood
##  Data: pollution
##         AIC       BIC   logLik
##   -4499.257 -4471.267 2254.628
##
## Random effects:
##  Formula: ~1 | id
##         (Intercept)   Residual
## StdDev:   0.1053993 0.06365121
##
## Fixed effects: logFEV1 ~ age + Height
##                  Value   Std.Error   DF   t-value p-value
## (Intercept) -1.8584680 0.030724848 1692 -60.48746       0
## age          0.0197738 0.001312301 1692  15.06807       0
## Height       1.6186612 0.030135691 1692  53.71243       0
##  Correlation:
##        (Intr) age
## age     0.834
## Height -0.961 -0.935
##
## Standardized Within-Group Residuals:
##         Min          Q1         Med          Q3         Max
## -5.87502973 -0.52029092  0.07057182  0.59894669  2.83013620
##
## Number of Observations: 1994
## Number of Groups: 300
```

Problem 1b

Fit the model with random intercept and slope of Height;

Model:

```
fit_1b <- lme(logFEV1 ~ age + Height, random = list(id = pdDiag (form = ~1+Height|id)),
 data=pollution, method = "ML")
summary(fit_1b)
```

```
## Linear mixed-effects model fit by maximum likelihood
##  Data: pollution
##         AIC       BIC   logLik
##   -4497.257 -4463.669 2254.628
##
## Random effects:
##  Formula: ~1 + Height | id | id
##  Structure: Diagonal
##         (Intercept) 1 + Height | idTRUE   Residual
## StdDev:  0.07452858          0.07452854 0.06365121
##
## Fixed effects: logFEV1 ~ age + Height
##                  Value   Std.Error   DF   t-value p-value
## (Intercept) -1.8584680 0.030724848 1692 -60.48746       0
## age          0.0197738 0.001312301 1692  15.06807       0
## Height       1.6186612 0.030135691 1692  53.71243       0
##  Correlation:
##        (Intr) age
## age     0.834
## Height -0.961 -0.935
##
## Standardized Within-Group Residuals:
##         Min          Q1         Med          Q3         Max
## -5.87502970 -0.52029092  0.07057182  0.59894669  2.83013618
##
## Number of Observations: 1994
## Number of Groups: 300
```

```
getVarCov(fit_1b, type="random.effects")
```

```
## Random effects variance covariance matrix
##                      (Intercept) 1 + Height | idTRUE
## (Intercept)            0.0055545           0.0000000
## 1 + Height | idTRUE    0.0000000           0.0055545
##   Standard Deviations: 0.074529 0.074529
```

Problem 1c Fit the model with random slope and Height; Random effects are dependent with unstructured covariance matrix;

Model:

```
fit_1c <- lme(logFEV1 ~ age + Height, random = ~1+Height|id, data=pollution, method = "M
L")
summary(fit_1c)
```

```
## Linear mixed-effects model fit by maximum likelihood
##  Data: pollution
##         AIC       BIC  logLik
##   -4603.339 -4564.154 2308.67
##
## Random effects:
##  Formula: ~1 + Height | id
##  Structure: General positive-definite, Log-Cholesky parametrization
##             StdDev     Corr
## (Intercept) 0.29064436 (Intr)
## Height      0.19500111 -0.936
## Residual    0.05818073
##
## Fixed effects: logFEV1 ~ age + Height
##                  Value  Std.Error   DF   t-value p-value
## (Intercept) -1.9031573 0.03496215 1692 -54.43479       0
## age          0.0187633 0.00124892 1692  15.02362       0
## Height       1.6574059 0.03186760 1692  52.00912       0
##  Correlation:
##        (Intr) age
## age     0.709
## Height -0.962 -0.849
##
## Standardized Within-Group Residuals:
##         Min          Q1         Med          Q3         Max
## -6.49339816 -0.49660646  0.08054351  0.56585585  2.90405007
##
## Number of Observations: 1994
## Number of Groups: 300
```

Problem 1d 1. Fixed effects;

```
fixed.effects(fit_1a)
```

```
## (Intercept)         age      Height
## -1.85846803  0.01977384  1.61866125
```

```
fixed.effects(fit_1b)
```

```
## (Intercept)         age      Height
## -1.85846803  0.01977384  1.61866125
```

```
fixed.effects(fit_1c)
```

```
## (Intercept)         age      Height
## -1.90315726  0.01876327  1.65740586
```

The estimates of the regression parameters a bit of changed in the first three models.

2. Covariance structures of the random effects;

```
getVarCov(fit_1a)
```

```
## Random effects variance covariance matrix
##             (Intercept)
## (Intercept)    0.011109
##   Standard Deviations: 0.1054
```

```
getVarCov(fit_1b)
```

```
## Random effects variance covariance matrix
##                     (Intercept) 1 + Height | idTRUE
## (Intercept)           0.0055545           0.0000000
## 1 + Height | idTRUE   0.0000000           0.0055545
##   Standard Deviations: 0.074529 0.074529
```

```
getVarCov(fit_1c)
```

```
## Random effects variance covariance matrix
##             (Intercept)    Height
## (Intercept)    0.084474 -0.053043
## Height        -0.053043  0.038025
##   Standard Deviations: 0.29064 0.195
```

As the first model only has one random effect, the covariance structure is the varaince of the intercept; The second model indicates there is no relationship between the random effcts; The third model indicates that the there are some negative relationship between these two random effects.

3. Table of AIC/BIC;

```
matrix1 <- matrix(c(-4499.257, -4471.267, 2254.628, -4520.109, -4486.522, 2266.054, -460
3.339, -4564.154, 2308.67), nrow=3, byrow = T)
colnames(matrix1) <- c("AIC", "BIC", "logLik")
rownames(matrix1) <- c("model1", "model2", "model3")
matrix1
```

```
##              AIC       BIC   logLik
## model1 -4499.257 -4471.267 2254.628
## model2 -4520.109 -4486.522 2266.054
## model3 -4603.339 -4564.154 2308.670
```

As the third model has the lowest AIC & BIC model, the third models is what I mostly prefer.

Problem 1e 1.

2. The variance of Yij;

```
getVarCov(fit_1c, type = "marginal", individual= 1)
```

```
## id 1
## Marginal variance covariance matrix
##           1         2         3         4         5         6         7
## 1 0.0153120 0.0113340 0.010964 0.010297 0.0098518 0.0097035 0.0095553
## 2 0.0113340 0.0143700 0.010766 0.010373 0.0101110 0.0100230 0.0099357
## 3 0.0109640 0.0107660 0.014028 0.010420 0.0102720 0.0102230 0.0101730
## 4 0.0102970 0.0103730 0.010420 0.013891 0.0105630 0.0105820 0.0106010
## 5 0.0098518 0.0101110 0.010272 0.010563 0.0141420 0.0108220 0.0108870
## 6 0.0097035 0.0100230 0.010223 0.010582 0.0108220 0.0142870 0.0109820
## 7 0.0095553 0.0099357 0.010173 0.010601 0.0108870 0.0109820 0.0144620
##   Standard Deviations: 0.12374 0.11987 0.11844 0.11786 0.11892 0.11953 0.12026
```

The within-subject covariane matrix;

```
getVarCov(fit_1c, type = "conditional", individual= 1)
```

```
## id 1
## Conditional variance covariance matrix
##          1        2        3        4        5        6        7
## 1 0.003385 0.000000 0.000000 0.000000 0.000000 0.000000 0.000000
## 2 0.000000 0.003385 0.000000 0.000000 0.000000 0.000000 0.000000
## 3 0.000000 0.000000 0.003385 0.000000 0.000000 0.000000 0.000000
## 4 0.000000 0.000000 0.000000 0.003385 0.000000 0.000000 0.000000
## 5 0.000000 0.000000 0.000000 0.000000 0.003385 0.000000 0.000000
## 6 0.000000 0.000000 0.000000 0.000000 0.000000 0.003385 0.000000
## 7 0.000000 0.000000 0.000000 0.000000 0.000000 0.000000 0.003385
##   Standard Deviations: 0.058181 0.058181 0.058181 0.058181 0.058181 0.058181 0.058181
```

3. The estimate of the mean of logFEV1; -1.903+0.0187*12+1.657*1.4 = 0.642