Statistical Modeling 2

Weekly assignments 1

1. Output from Task1.R

```
# a)
Restricted MLE for phi = 0.3966289

# b)
Likelihood ratio test statistic = 108.6823
p-value = 2.511577e-24
Model H1 should be selected

# c)
Empirical blup for new obs. Y = 86.58952

# d)
From this part I got following error
Error in solve.default(t(X) %*% solve(V) %*% X) :
  system is computationally singular: reciprocal condition
number = 2.34382e-17
when I was calculating the variance of the prediction error.
Unfortunately I didn't had time to fix it.
```

2. Output from Task2.R

```
# a)
RMLE for sigma2.z = 2.767964

# b)
Forest area with largest random effect:
 Clark Fork
 2.356447

# c)
ML prediction for mu = 27.79635
```

3.

a)

3. a)

i) Show that $\phi V\underline{i} = \underline{w}$

$$V\underline{i} = \frac{1}{1-\phi^2} \begin{pmatrix} 1 & \phi & \phi^2 & \cdots & \phi^{t-2} & \phi^{t-1} \\ \phi & 1 & \phi & \cdots & \phi^{t-3} & \phi^{t-2} \\ \vdots & & \ddots & & & \\ \phi^{t-2} & \phi^{t-3} & \phi^{t-4} & \cdots & 1 & \phi \\ \phi^{t-1} & \phi^{t-2} & \phi^{t-3} & \cdots & \phi & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \phi^{t-1} \\ \phi^{t-2} \\ \phi^{t-3} \\ \vdots \\ \phi \\ 1 \end{pmatrix} \frac{1}{1-\phi^2}$$

$$\phi V\underline{i} = \frac{\phi}{1-\phi^2} \begin{pmatrix} \phi^{t-1} \\ \phi^{t-2} \\ \vdots \\ \phi \\ 1 \end{pmatrix} = \frac{1}{1-\phi^2} \begin{pmatrix} \phi^{t} \\ \phi^{t-1} \\ \vdots \\ \phi^2 \\ \phi \end{pmatrix} = w$$

ii) Show that BLUP of y_{t+1} is $\hat{y}_{t+1} = (1-\phi)\hat{\beta}_0 + \phi y_t$

$$\phi V\underline{i} = \underline{w} \iff \phi \underline{i}' = \underline{w}'V^{-1}$$

$$\begin{aligned} \hat{y}_{t+1} &= x_{t+1}'\hat{\underline{\beta}} + \underline{w}'V^{-1}(y - X\hat{\underline{\beta}}) \\ &= 1 \cdot \hat{\beta}_0 + w'V^{-1}(y - 1\hat{\beta}_0) \\ &= \hat{\beta}_0 + \phi \underline{i}'(y - \underline{1}\hat{\beta}_0) \\ &= \hat{\beta}_0 - \phi i \underline{1}\hat{\beta}_0 + \phi i' y \\ &= (1-\phi)\hat{\beta}_0 + \phi y_t \qquad \square \end{aligned}$$

b)

Matrix X:

Model matrix that consist all explanatory variable values. In the case of M_2 this matrix has 542 rows (observations) and 3 columns (intercept, x_{it} and x_{it}^2)

Matrix Z:

Second set of dummy explanatory variables corresponding to random effects. Matrix Z diagonally consist of blocks of Z_i where each Z_i corresponds to equivalent variable combination in X_i. In the case of M_2 we have unique random effects corresponding to variable Tree.ID, and Forest.ID and hence the columns of block Z_i consist Tree.ID-values and measured diameters for that tree. Overall we have 66 different tree species and therefore 66 different Z_i submatrices. The Forest.ID-values are not fixed for specific tree only so they doesn't belong to block Z_i but somewhere else in Z matrix (?). Overall, the size of matrix Z is 542x141, where 542 is number of samples and 141 comes from 66 Tree.ID-values times 2 random effects corresponding to those, plus 9 different Forest.ID-values.

Matrix G:

We assume that $b \sim N(0, G)$, where G's diagonal elements are matrices of form F_i, where
$F_i = cov(b_i) = \begin{pmatrix} \sigma_{b_0}^2 & \sigma_{b_0 b_1} \\ \sigma_{b_0 b_1} & \sigma_{b_1}^2 \end{pmatrix}$

Matrix V:

The marginal distribution of response variables y is assumed to be $y \sim N(X\beta, V)$, where $V = \sigma^2 I + ZGZ'$. Moreover, for each sampling unit i, $y_i \sim N(X_i\beta, V_i)$, where $V_i = \sigma^2 I + Z_i F_i Z_i'$

Vector β:

This vector consist the unknown fixed parameters for explanatory variables, say $\beta = (\beta_0, \beta_1, \beta_2)^T$.

Vector b:

This vector consist the unknown random effect parameters, say $b = (b_1, b_2, ..., b_{66})^T$, where $b_i = (b_{i0}, b_{i1})^T$

c)

(c) Show that the sample mean

$$\hat{\beta}_0 = (\underline{1}'\underline{1})^{-1}\underline{1}'\underline{y} = \frac{1}{n}\underline{1}'\underline{y} = \bar{y}$$

is the BLUP of y_f

linear predictor $g'y$, where $g' = \frac{1}{n}1'$

Show: $g'(X \vdots VM) = (X_f' \vdots w'M)$, where $X_f = 1$, $W' = \rho 1'$, $M = I - \frac{1}{n}11'$

$g'X = \frac{1}{n}1'1 = \frac{1}{n}n = 1 = x_f$

$$\begin{aligned}
g'VM &= \frac{1}{n}1'\left[(1-\rho)I + \rho 11'\right]\left[I - \frac{1}{n}11'\right] \\
&= \left[\frac{1}{n}1'(1-\rho)I + \frac{1}{n}1'\rho 11'\right]\left[I - \frac{1}{n}11'\right] \\
&= \frac{1}{n}1'(1-\rho)I\left[I - \frac{1}{n}11'\right] + \frac{1}{n}1'\rho 11'\left[I - \frac{1}{n}11'\right] \\
* &= 0 + \frac{1}{n}1'\rho 11'\left[I - \frac{1}{n}11'\right] \\
&= \frac{1}{n}1'1\rho 1'\left[I - \frac{1}{n}11'\right] \\
&= \rho 1'\left[I - \frac{1}{n}11'\right] \\
&= W'M
\end{aligned}$$

$\therefore g'(X \vdots VM) = (X_f' \vdots W'M)$

$$\begin{aligned}
* \quad & \frac{1}{n}1'(1-\rho)I\left[I - \frac{1}{n}11'\right] \\
&= (1-\rho)\left[\frac{1}{n}1'I\left[I - \frac{1}{n}11'\right]\right] \\
&= (1-\rho)\left[\frac{1}{n}1' - \frac{1}{n}1'\frac{1}{n}11'\right] \\
&= (1-\rho)\left[\frac{1}{n}1' - \frac{1}{n^2}1'11'\right] \\
&= (1-\rho)\left[\frac{1}{n}1' - \frac{1}{n^2}n1'\right] \\
&= (1-\rho)\left[\frac{1}{n}1' - \frac{1}{n}1'\right] = 0
\end{aligned}$$