GENE638 - Homework 3

Daniel Osorio - dcosorioh@tamu.edu
Department of Veterinary Integrative Biosciences
Texas A&M University

1. Given the Lactation Milfat Yield Data and assuming the model $y_{ijk} = \mu + H_i + L_j + C_k + \epsilon_{ijk}$ where herd (H), lactation (L), and cow (C) effects are fixed effects, and $\underline{y} = X\underline{\beta} + \epsilon$; $E\left[\underline{y}\right] = X\underline{\beta}$; $Var(\underline{y}) = Var(\underline{\epsilon}) = I\sigma_\epsilon^2 =$ V.

COW	HERD	LACTATION	MILKFAT (lb)
1	1	1	600
1	1	2	680
2	1	1	500
3	2	1	800
3	2	2	895
4	2	1	775
5	2	1	600
5	2	2	715

(a) What are $\underline{y}$, X and β for this design?

$$\underline{y} = \begin{bmatrix} 600 \\ 680 \\ 500 \\ 800 \\ 895 \\ 775 \\ 600 \\ 715 \end{bmatrix} \qquad X = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \qquad \beta = \begin{bmatrix} \mu \\ H_1 \\ H_2 \\ L_1 \\ L_2 \\ C_1 \\ C_2 \\ C_3 \\ C_4 \\ C_5 \end{bmatrix}$$

(b) Write the equations $X'X\beta = X'\underline{y}$ in matrix form (after multiplication of both sides by σ_ϵ^2)

$$\begin{bmatrix} 8 & 3 & 5 & 2 & 1 & 2 & 1 & 2 \\ 3 & 3 & 2 & 2 & 1 & 0 & 0 & 0 \\ 5 & 2 & 5 & 1 & 1 & 1 & 1 & 1 \\ 2 & 2 & 1 & 2 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 2 & 0 & 1 & 0 & 0 & 2 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 2 & 0 & 1 & 0 & 0 & 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} \mu \\ H_1 \\ H_2 \\ L_1 \\ L_2 \\ C_1 \\ C_2 \\ C_3 \\ C_4 \\ C_5 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 600 \\ 680 \\ 500 \\ 800 \\ 895 \\ 775 \\ 600 \\ 715 \end{bmatrix}$$

(c) Show that $\hat{\mu} = 705.8$; $\hat{H}_1 = -109.2$; $\hat{H}_2 = 0$; $\hat{L}_1 = -96.7$; $\hat{L}_2 = 0$; $\hat{C}_1 = 91.7$; $\hat{C}_2 = 0$; $\hat{C}_3 = 190$; $\hat{C}_4 = 165.8$; $\hat{C}_5 = 0$, provide a solution for the systems of equation in (b).

```
> yHat <- X %*% c(mu = 705.8, H1 = -109.2, H2 = 0, L1 = -96.7,
+          L2 = 0, C1 = 91.7, C2 = 0, C3 = 190, C4 = 165, C5 = 0)
> yHat
      [,1]
[1,] 591.6
[2,] 688.3
[3,] 499.9
[4,] 799.1
[5,] 895.8
[6,] 774.1
[7,] 609.1
[8,] 705.8
```

(d) Find $\underline{\hat{\epsilon}} = \underline{y} - X\beta$, which are the fitted errors when fixed effects model assumed.

```
> e <- y - yHat
> e
     [,1]
[1,]  8.4
[2,] -8.3
[3,]  0.1
[4,]  0.9
[5,] -0.8
[6,]  0.9
[7,] -9.1
[8,]  9.2
```

(e) Find $\underline{\hat{\epsilon}}'\underline{\hat{\epsilon}}$, the error sum of squares.

```
> t(e) %*% e
       [,1]
[1,] 309.17
```

2. Now, setting cow (C) effects as random effect, the model is: $\underline{y} = X\beta + Zu + \epsilon$, and:

$$E\begin{bmatrix} \underline{y} \\ \underline{u} \\ \underline{e} \end{bmatrix} = \begin{bmatrix} X\underline{\beta} \\ 0 \\ 0 \end{bmatrix}; Var\left(\begin{bmatrix} \underline{y} \\ \underline{u} \\ \underline{e} \end{bmatrix}\right) = \begin{bmatrix} ZGZ' + R & ZG & R \\ GZ' & G & 0 \\ R & 0 & R \end{bmatrix}$$

The difference here is that cows are not longer in $\underline{\beta}$. Generalized least-squares estimate of $\underline{\beta}$ and predictor $\underline{u}$ satisfy:

$$\begin{bmatrix} X'R^{-1}X & X'R^{-1}Z \\ Z'R^{-1}X & Z'R^{-1}Z + G^{-1} \end{bmatrix} \begin{bmatrix} \underline{\hat{\beta}} \\ \underline{\hat{u}} \end{bmatrix} = \begin{bmatrix} X'R^{-1}\underline{y} \\ Z'R^{-1}\underline{y} \end{bmatrix}$$

and if $G = I\sigma_g^2$, $R = I\sigma_\epsilon^2$, and $\lambda = \frac{\sigma_\epsilon^2}{\sigma_g^2}$, these may be written as:

$$\begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + I\lambda \end{bmatrix} \begin{bmatrix} \underline{\hat{\beta}} \\ \underline{\hat{u}} \end{bmatrix} = \begin{bmatrix} X'\underline{y} \\ Z'\underline{y} \end{bmatrix}$$

(a) What are X, Z, $\underline{\beta}$ and $\underline{u}$ for this design?

(b) If $\lambda = 0$, write these equations ans show that they are identical to the equations in 1(b) where cows were modeled as a fixed effect.

3. And now X in 1(a) can be partitioned as $[X : Z]$ for X and Z as defined in 2(a), so now:

$$\begin{aligned} \underline{y} &= X\underline{\beta} + Z\underline{u} + \epsilon \\ &= [X : Z] \begin{bmatrix} \hat{\underline{\beta}} \\ \hat{\underline{u}} \end{bmatrix} + \epsilon \\ &= W\underline{\theta} + \epsilon \end{aligned}$$

Which is of the same form as $\underline{y} = X\underline{\beta} + \underline{\epsilon}$. See therefore that separating fixed and random effects in the model as $\underline{y} = X\underline{\beta} + Z\underline{u} + \underline{\epsilon}$ is only convention. Elements of $\underline{u}$ are not treated as random effects for purpose of prediction until G^{-1} is included in the MME model.

(a) Assume $\lambda = \frac{\sigma_\epsilon^2}{\sigma_g^2} = 1.5$. Now write the completed equations from above

(b) Show that $\hat{\mu} = 819.4$; $\hat{H}_1 = -168.6$; $\hat{H}_2 = 0$; $\hat{L}_1 = -96.9$; $\hat{L}_2 = 0$; $\hat{C}_1 = 21.5$; $\hat{C}_2 = -21.5$; $\hat{C}_3 = 43.8$; $\hat{C}_5 = -64.8$ provide a solution to the systems of equations.

* Note here that $\hat{\mu}$ estimates $\mu + H_2 + L_2$; $\hat{H}_1$ estimates $H_1 - H_2$; $\hat{H}_2$ estimates 0; $\hat{L}_1$ estimates $L_1 - L_2$; $\hat{L}_2$ estimates 0; $\hat{C}_k$ predicts C_k

(c) Find $\hat{\underline{\epsilon}} - \underline{y} - X\hat{\underline{\beta}} - Z\hat{\underline{u}}$, which are the fitted errors when mixed model assumed. What do the residuals here and in 1(d) predict?

(d) Find $\hat{\underline{\epsilon}}'\hat{\underline{\epsilon}}$ and compare results to those in 1(e). What appears to be a consequence of modeling animals as fixed effects?