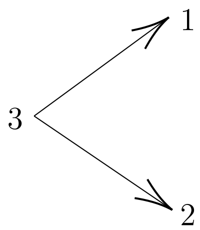
GENE638 - Homework 5

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

Data:

COW	YEAR	TICK COUNT
1	1	80
2	1	70
1	2	86
2	2	72
1	3	92
2	3	78

Pedrigree:



1. Construct the numerator relationship and its inverse

```
      A1   A2  A3  M1  M2
A1 1.00 0.25 0.5 0.5 0.0
A2 0.25 1.00 0.5 0.0 0.5
A3 0.50 0.50 1.0 0.0 0.0
M1 0.50 0.00 0.0 1.0 0.0
M2 0.00 0.50 0.0 0.0 1.0
```

```
   A1 A2   A3   M1   M2
A1  2  0 -1.0 -1.0  0.0
A2  0  2 -1.0  0.0 -1.0
A3 -1 -1  2.0  0.5  0.5
M1 -1  0  0.5  1.5  0.0
M2  0 -1  0.5  0.0  1.5
```

2. The predicted breeding value of the unidentified parent of animal 1 is $\frac{2}{3}(\hat{A}_1 - \frac{1}{2}\hat{A}_3)$ and that for the unidentified parent of animal 2 is $\frac{2}{3}(\hat{A}_2 - \frac{1}{2}\hat{A}_3)$. Identify the base generation animals and show that $1'A^{-1}\hat{u} = 0$ is consistent with the sum of breeding values of base generation being 0.

```
> X <- matrix(c(1,1,0,0,0,0,0,0,1,1,0,0,0,0,0,0,1,1), ncol = 3, byrow = FALSE)
> Z <- matrix(c(1,0,1,0,1,0,0,1,0,1,0,1,0,0,0,0,0,0,
+               0,0,0,0,0,0,0,0,0,0,0,0), ncol = 5)
> colnames(Z) <- c("A1","A2","A3","M1","M2")
> colnames(X) <- c("Y1", "Y2", "Y3")
> X1 <- rbind(
+   cbind(t(X) %*% X, t(X) %*% Z),
+   cbind(t(Z) %*% X, t(Z) %*% Z + Ainv)
+   )
> y <- c(80,70,86,72,92,78)
> Y1 <- c(t(X) %*% y,t(Z) %*% y)
> u <- round(solve(X1) %*% Y1,2)
> rep(1,5) %*% (Ainv %*% u[4:8])
```

```
              [,1]
[1,] -1.776357e-15
```

3. Write the data in general matrix terms $y = X\beta + Zu + e$ for the model $Y_{ij} = \text{Year}_i + \text{Animal}_j + e_{ij}$ where $\text{Year}_i = \mu + \text{Year}_i$. Include animal 3 in $\hat{u}$ as in (2)

```
> X

     Y1 Y2 Y3
[1,]  1  0  0
[2,]  1  0  0
[3,]  0  1  0
[4,]  0  1  0
[5,]  0  0  1
[6,]  0  0  1

> Z[,1:3]

     A1 A2 A3
[1,]  1  0  0
[2,]  0  1  0
[3,]  1  0  0
[4,]  0  1  0
[5,]  1  0  0
[6,]  0  1  0
```

4. Assuming $R = I\sigma_e^2$ and $G = A\sigma_a^2$ and $\lambda = \frac{\sigma_e^2}{\sigma_a^2} = 3$, write the MME for model in (3)

```
> lambda <- 3
> Z <- Z[,1:3]
> Ainv <- solve(A[1:3,1:3])
> Z

     A1 A2 A3
[1,]  1  0  0
[2,]  0  1  0
[3,]  1  0  0
[4,]  0  1  0
[5,]  1  0  0
[6,]  0  1  0

> X1 <- rbind(
+   cbind(t(X) %*% X, t(X) %*% Z),
+   cbind(t(Z) %*% X, t(Z) %*% Z + Ainv*lambda)
+   )
> y <- c(80,70,86,72,92,78)
> Y1 <- c(t(X) %*% y,t(Z) %*% y)
> X1

   Y1 Y2 Y3            A1            A2 A3
Y1  2  0  0  1.000000e+00  1.000000e+00  0
Y2  0  2  0  1.000000e+00  1.000000e+00  0
Y3  0  0  2  1.000000e+00  1.000000e+00  0
```

```
A1  1  1  1  7.000000e+00  1.249001e-16 -2
A2  1  1  1  4.440892e-17  7.000000e+00 -2
A3  0  0  0 -2.000000e+00 -2.000000e+00  5
```

```
> Y1
```

```
[1] 150 158 170 258 220   0
```

```
> round(solve(X1) %*% Y1,4)
```

```
      [,1]
Y1 75.0000
Y2 79.0000
Y3 85.0000
A1  2.7143
A2 -2.7143
A3  0.0000
```

5. From the row of the MME corresponding to $\hat{A}_3$ show that $\hat{A}_3 = \frac{2}{5}(\hat{A}_1 + \hat{A}_2)$

```
> X1[6,]
```

```
Y1 Y2 Y3 A1 A2 A3
 0  0  0 -2 -2  5
```

$$\begin{aligned} 0 &= -2\hat{A}_1 - 2\hat{A}_2 + 5\hat{A}_3 \\ -5\hat{A}_3 &= -2\hat{A}_1 - 2\hat{A}_2 \\ 5\hat{A}_3 &= 2\hat{A}_1 + 2\hat{A}_2 \\ \hat{A}_3 &= \frac{2\hat{A}_1 + 2\hat{A}_2}{5} \\ \hat{A}_3 &= \frac{2}{5}(\hat{A}_1 + \hat{A}_2) \end{aligned}$$

6. Absorb the year equations into the animal equations to obtain a system of equations involving only the unknown breeding values $\hat{A}_1$, $\hat{A}_2$, and $\hat{A}_3$

```
> M <- diag(6) - (X %*% solve(t(X) %*% X) %*% t(X))
> M
```

```
     [,1] [,2] [,3] [,4] [,5] [,6]
[1,]  0.5 -0.5  0.0  0.0  0.0  0.0
[2,] -0.5  0.5  0.0  0.0  0.0  0.0
[3,]  0.0  0.0  0.5 -0.5  0.0  0.0
[4,]  0.0  0.0 -0.5  0.5  0.0  0.0
[5,]  0.0  0.0  0.0  0.0  0.5 -0.5
[6,]  0.0  0.0  0.0  0.0 -0.5  0.5
```

```
> C22 <- t(Z) %*% M %*% Z + (Ainv * lambda)
> C22
```

```
     A1   A2 A3
A1  5.5 -1.5 -2
A2 -1.5  5.5 -2
A3 -2.0 -2.0  5
```

```
> b <- t(Z) %*% M %*% y
> b

   [,1]
A1   19
A2  -19
A3    0

> u <- round(solve(t(Z) %*% M %*% Z + (Ainv * lambda)) %*% t(Z) %*% M %*% y,4)
> u

      [,1]
A1  2.7143
A2 -2.7143
A3  0.0000
```

7. What are the effective numbers of observations on all three animals?

```
> diag(t(Z) %*% M %*% Z)

 A1  A2  A3
1.5 1.5 0.0
```

8. From the appropiate row of the absorbed MME, once again show that $\hat{A}_3 = \frac{2}{5}(\hat{A}_1 + \hat{A}_2)$

```
> C22[3,]

A1 A2 A3
-2 -2  5
```

$$
\begin{aligned}
0 &= -2\hat{A}_1 - 2\hat{A}_2 + 5\hat{A}_3 \\
-5\hat{A}_3 &= -2\hat{A}_1 - 2\hat{A}_2 \\
5\hat{A}_3 &= 2\hat{A}_1 + 2\hat{A}_2 \\
\hat{A}_3 &= \frac{2\hat{A}_1 + 2\hat{A}_2}{5} \\
\hat{A}_3 &= \frac{2}{5}(\hat{A}_1 + \hat{A}_2)
\end{aligned}
$$

9. Using ordinary Gauss-Seidel iteration, find two succesive aproximations to the predicted breeding values in (6)

```
> L <- D <- matrix(0,nrow = nrow(C22), ncol = ncol(C22))
> L[lower.tri(L)] <- C22[lower.tri(C22)]
> diag(D) <- diag(C22)
> L

     [,1] [,2] [,3]
[1,]  0.0    0    0
[2,] -1.5    0    0
[3,] -2.0   -2    0

> D
```

```
     [,1] [,2] [,3]
[1,]  5.5  0.0    0
[2,]  0.0  5.5    0
[3,]  0.0  0.0    5

> x0 <- solve(D) %*% b
> for(i in 1:2){
+   Xi <- solve(L + D) %*% (b - t(L) %*% x0)
+   print(Xi)
+   x0 <- Xi
+ }

           [,1]
[1,]  2.5123967
[2,] -2.7693464
[3,] -0.1027799
           [,1]
[1,]  2.6618947
[2,] -2.7659487
[3,] -0.0416216
```

10. Show that $\hat{A}_1 = 2.7143$, $\hat{A}_2 = -2.7143$ and $\hat{A}_3 = 0$, provide a solution to the absorbed equations in (6)

```
> t(Z) %*% M %*% Z + (Ainv * lambda)

     A1   A2 A3
A1  5.5 -1.5 -2
A2 -1.5  5.5 -2
A3 -2.0 -2.0  5

> u

      [,1]
A1  2.7143
A2 -2.7143
A3  0.0000

> (t(Z) %*% M %*% Z + (Ainv * lambda)) %*% u

       [,1]
A1  19.0001
A2 -19.0001
A3   0.0000

> t(Z) %*% M %*% y

   [,1]
A1   19
A2  -19
A3    0
```

11. Using the breeding values in (10), backsolve the solutions to $\hat{Y}_1$, $\hat{Y}_2$ and $\hat{Y}_3$

```
> beta <- solve(t(X) %*% X) %*% t(X)%*%(y - Z %*% u)
> beta

   [,1]
Y1   75
Y2   79
Y3   85
```

12. The inverse of the coefficient matrix in (6) for animals in the order $\hat{A}_3, \hat{A}_1, \hat{A}_2$, is

$$\begin{bmatrix} 0.33330 & 0.16667 & 0.16667 \\ 0.16667 & 0.27976 & 0.13690 \\ 0.16667 & 0.1369 & 0.27976 \end{bmatrix} = (Z'MZ + A^{-1}\lambda)^{-1}$$

Which also is a submatrix of the inverse

$$\begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + A^{-1}\lambda \end{bmatrix} \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}$$

Calculate the first aproximations:

$$\sigma_e^2 = \frac{y'y - \underline{\hat{\beta}'}X'y - \hat{u}'Z'\underline{y}}{N - p}$$

$$\sigma_a^2 = \frac{\hat{u}'A^{-1}\hat{u} + \sigma_e^2 tr\left[A^{-1}C_{22}\right]}{q}$$

$$\lambda = \frac{\sigma_e^2}{\sigma_a^2}$$

```
> N <- length(y)
> p <- Matrix::rankMatrix(X)
> q <- ncol(Z)
> sigma2E <-(t(y) %*% y - t(beta) %*% t(X) %*%
+              y - t(u) %*% t(Z) %*% y) / as.numeric(N - p)
> sigma2E

         [,1]
[1,] 47.61887

> sigma2A <- (t(u) %*% Ainv %*% u +
+               sigma2E * sum(diag(Ainv %*% solve(C22))))/q
> sigma2A

         [,1]
[1,] 20.15421

> sigma2E/sigma2A

         [,1]
[1,] 2.362725
```