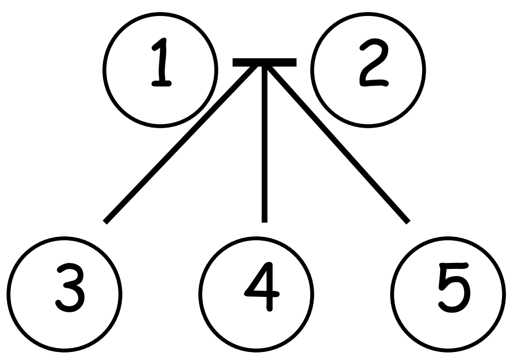
Mixed-Model Problem

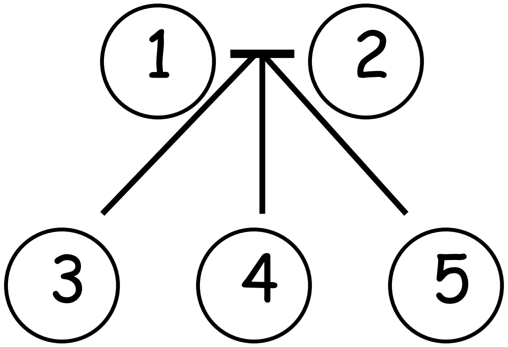
Consider the following pedigree



Here, 1 and 2 are unrelated, and not inbred

First, compute the A matrix for this pedigree

Computing the A matrix



All outbreds: $A_{ii} = 1$

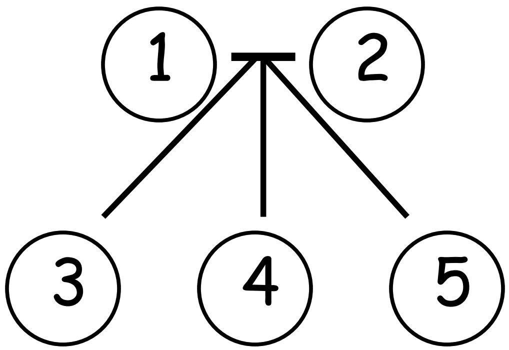
1,2: unrelated, $A_{ij} = 0$

1:3,4,5: Parent-offspring, $A_{ij} = 1/2$

2:3,4,5: Parent-offspring, $A_{ij} = 1/2$

3,4,5: full sibs, $A_{ij} = 1/2$

$$\begin{array}{cc} & \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\ \mathbf{A} = \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} & \begin{pmatrix} 1 & 0 & 1/2 & 1/2 & 1/2 \\ 0 & 1 & 1/2 & 1/2 & 1/2 \\ 1/2 & 1/2 & 1 & 1/2 & 1/2 \\ 1/2 & 1/2 & 1/2 & 1 & 1/2 \\ 1/2 & 1/2 & 1/2 & 1/2 & 1 \end{pmatrix} \end{array}$$



$$\mathbf{A} = \frac{1}{2}\begin{pmatrix} 2 & 0 & 1 & 1 & 1 \\ 0 & 2 & 1 & 1 & 1 \\ 1 & 1 & 2 & 1 & 1 \\ 1 & 1 & 1 & 2 & 1 \\ 1 & 1 & 1 & 1 & 2 \end{pmatrix}$$

Suppose only measure 1, 3-5, with $y_1 = 40$, $y_3 = 20$, $y_4 = 25$, $y_5 = 30$, only fixed effect is the mean μ.

However, we wish to estimate all five breeding values

$$\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{pmatrix}$$

What are the elements in the mixed model $\mathbf{y} = \mathbf{X}\beta + \mathbf{Z}\mathbf{a} + \mathbf{e}$,

$$\underbrace{\begin{pmatrix}40\\20\\25\\30\end{pmatrix}}_{\mathbf{y}} = \underbrace{\begin{pmatrix}1\\1\\1\\1\end{pmatrix}}_{\mathbf{X}} \underbrace{(\mu)}_{\beta} + \underbrace{\begin{pmatrix}1&0&0&0&0\\0&0&1&0&0\\0&0&0&1&0\\0&0&0&0&1\end{pmatrix}}_{\mathbf{Z}} \underbrace{\begin{pmatrix}a_1\\a_2\\a_3\\a_4\\a_5\end{pmatrix}}_{\mathbf{a}} + \underbrace{\begin{pmatrix}e_1\\e_3\\e_4\\e_5\end{pmatrix}}_{\mathbf{e}}$$

$\mathbf{y} = \mathbf{X}\beta + \mathbf{Z}\mathbf{a} + \mathbf{e}$

Only slightly tricky part is Z

$$\begin{pmatrix}y_1\\y_3\\y_4\\y_5\end{pmatrix} = \begin{pmatrix}a_1\\a_3\\a_4\\a_5\end{pmatrix} = \begin{matrix} & 1 & 2 & 3 & 4 & 5 \\ 1 & 1&0&0&0&0\\ 3 & 0&0&1&0&0\\ 4 & 0&0&0&1&0\\ 5 & 0&0&0&0&1\end{matrix} \begin{pmatrix}a_1\\a_2\\a_3\\a_4\\a_5\end{pmatrix}$$

$$\mathbf{y} = \begin{pmatrix} 40 \\ 20 \\ 25 \\ 30 \end{pmatrix}, \quad \mathbf{X} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{Z} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\mathbf{A} = \frac{1}{2} \begin{pmatrix} 2 & 0 & 1 & 1 & 1 \\ 0 & 2 & 1 & 1 & 1 \\ 1 & 1 & 2 & 1 & 1 \\ 1 & 1 & 1 & 2 & 1 \\ 1 & 1 & 1 & 1 & 2 \end{pmatrix}$$

Suppose $\sigma_A^2=30$, $\sigma_e^2=70$, giving $h^2 = 0.3$

Here, **a** ~ (0,G=30*A),

e ~ (0, 70*I),

V = ZGZ^T + R

= Z(30*A)Z^T + 70*I

1st compute V

Then solve for **a** and μ using

$$\hat{\beta} = \left(\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{V}^{-1}\mathbf{y}$$

$$\hat{\mathbf{a}} = \mathbf{G}\mathbf{Z}^T\mathbf{V}^{-1}\left(\mathbf{y} - \mathbf{X}\hat{\beta}\right)$$

Refresher: R matrix commands

- A^T, the transpose of A, **`t(A)`**
- A^{-1}, the inverse of A, **`solve(A)`**
- AB, the matrix product, **`A %*% B`**

Enter A

```
> A<-0.5*matrix(c(2,0,1,1,1,0,2,1,1,1,1,1,2,1,1,1,1,1,2,1,1,1,1,1,2),nrow=5)
> A
     [,1] [,2] [,3] [,4] [,5]
[1,]  1.0  0.0  0.5  0.5  0.5
[2,]  0.0  1.0  0.5  0.5  0.5
[3,]  0.5  0.5  1.0  0.5  0.5
[4,]  0.5  0.5  0.5  1.0  0.5
[5,]  0.5  0.5  0.5  0.5  1.0
```

Enter Z

```
> Z<-matrix(c(1,0,0,0,0,0,0,0,0,1,0,0,0,0,1,0,0,0,0,1),nrow=4)
> Z
     [,1] [,2] [,3] [,4] [,5]
[1,]    1    0    0    0    0
[2,]    0    0    1    0    0
[3,]    0    0    0    1    0
[4,]    0    0    0    0    1
```

Enter I

```
> I<-matrix(c(1,0,0,0,0,1,0,0,0,0,1,0,0,0,0,1),nrow=4)
> I
     [,1] [,2] [,3] [,4]
[1,]    1    0    0    0
[2,]    0    1    0    0
[3,]    0    0    1    0
[4,]    0    0    0    1
```

Now compute (and store) $V = Z(30^*A)Z^T + 70^*I$

```
> V<-Z%*%(30*A)%*%t(Z) + 70*I
> V
     [,1] [,2] [,3] [,4]
[1,]  100   15   15   15
[2,]   15  100   15   15
[3,]   15   15  100   15
[4,]   15   15   15  100
```

Enter y

```
> y<-matrix(c(40,20,25,30),nrow=4)
> y
     [,1]
[1,]   40
[2,]   20
[3,]   25
[4,]   30
```

Now compute (and store) bhat = $(X^TV^{-1}X)^{-1}X^TV^{-1}y$

```
bhat <- solve( t(X) %*% solve(V) %*% X) %*% t(X) %*% solve(V) %*% y
```

bhat (the weighted estimate of the mean) = 28.75

Now, solve for ahat = $GZ^TV^{-1}(y\text{-}X \text{ bhat})$, where $G = 30^*A$

```
ahat <- 30*A %*% t(Z) %*% solve(V) %*% (y- X %*% bhat)
```

```
> ahat
           [,1]
[1,]  1.9852941
[2,] -1.9852941
[3,] -1.5441176
[4,] -0.6617647
[5,]  0.2205882
>
```

Now, suppose that $y_2 = 20$. What are the new estimates of a_1 through a_5?

Note that **A** is the same, but you need to compute new **y**, **X**, and **Z** matrices, and that the covariance matrix 70***I** for the residuals is now 70*$\mathbf{I}_{5 \times 5}$, rather than 70*$\mathbf{I}_{4 \times 4}$ as was the case when we had only four y values.

$$\mathbf{y} = \begin{pmatrix} 40 \\ 20 \\ 20 \\ 25 \\ 30 \end{pmatrix}, \quad \mathbf{X} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{Z} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

```
> y<-matrix(c(40,20,20,25,30),nrow=5)
> X<-matrix(c(1,1,1,1,1),nrow=5)
> I<-matrix(c(1,0,0,0,0,0,1,0,0,0,0,0,1,0,0,0,0,0,1,0,0,0,0,0,1),nrow=5)
> Z<-I
> V<-Z%*%(30*A)%*%t(Z) + 70*I
> bhat<-solve( t(X) %*% solve(V) %*% X) %*% t(X) %*% solve(V) %*% y
> ahat <- 30*A %*% t(Z) %*% solve(V) %*% (y- X %*% bhat)
> ahat
           [,1]
[1,]  3.0000000
[2,] -3.0000000
[3,] -1.2770898
[4,] -0.3947368
[5,]  0.4876161
```

Addition play-at-home exercises

- See what the effect of keeping **y**, **X**, and **Z** constant, but varying *A* (the pedigree) is.
  - For example, suppose all unrelated
- Use Henderson's mixed model equations to solve for β, **a**.

More on the animal model

- Under the animal model
  - $\mathbf{y} = \mathbf{X}\beta + \mathbf{Za} + \mathbf{e}$
  - $\mathbf{a} \sim (0, \sigma_A^2 \mathbf{A})$, $e \sim (0, \sigma_e^2 \mathbf{I})$
  - $\text{BLUP}(\mathbf{a}) = \sigma_A^2 \mathbf{A}\mathbf{Z}^T\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\beta)$
  - Where $\mathbf{V} = \mathbf{ZGZ}^T + \mathbf{R} = \sigma_A^2 \mathbf{ZAZ}^T + \sigma_e^2 \mathbf{I}$
- Consider the simplest case of a single observation on one outbred individual, where the only fixed effect is the mean μ, which is assumed known
  - Here $\mathbf{Z} = \mathbf{A} = \mathbf{I} = (1)$,
  - $\mathbf{V} = \sigma_A^2 + \sigma_e^2$
  - $\sigma_A^2 \mathbf{AZ}^T\mathbf{V}^{-1} = \sigma_A^2 / (\sigma_A^2 + \sigma_e^2) = h^2$
  - $\text{BLUP}(a) = h^2(y - \mu)$

- More generally, with single observations on n unrelated individuals,
  - $A = Z = I_{n \times n}$
  - $\mathbf{V} = \sigma_A^2 \mathbf{ZAZ}^T + \sigma_e^2 \mathbf{I} = (\sigma_A^2 + \sigma_e^2)\,\mathbf{I}$
  - $\sigma_A^2\, \mathbf{AZ}^T\mathbf{V}^{-1} = h^2\, \mathbf{I}$
  - $\text{BLUP}(\mathbf{a}) = \sigma_A^2 \mathbf{AZ}^T\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\beta) = h^2(\mathbf{y} - \mu)$
- Hence, the predicted breeding value of individual i is just $\text{BLUP}(a_i) = h^2(y_i - \mu)$
- When at least some individuals are related and/or inbred (so that $A \neq I$) and/or missing or multiple records (so that $Z \neq I$), then the estimates of the BV differ from this simple form, but BLUP fully accounts for this