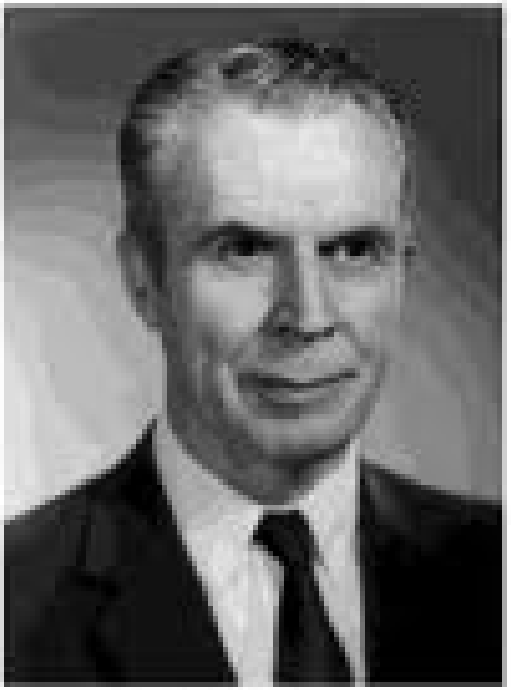
Prediction

LRS

CGIL

July-Aug 2012

C. R. Henderson



Sire Problem

Sire Progeny Means(numbers).

		Sires		
		1	2	3
Herds	1	128(5)	150(4)	0(0)
	2	90(5)	110(20)	120(2)

$$\bar{y}_{ij} = \mu + H_i + S_j + \bar{e}_{ij}$$

Model

$$\bar{y}_{ij} = \mu + H_i + S_j + \bar{e}_{ij}$$

$$\begin{aligned} E(\bar{y}_{ij}) &= \mu \\ Var(H_i) &= \sigma_h^2 = \frac{1}{6}\sigma_e^2 \\ Var(S_j) &= \sigma_s^2 = \frac{1}{15}\sigma_e^2 \\ Var(\bar{e}_{ij}) &= \sigma_e^2/n_{ij} \end{aligned}$$

Sires are not related. Which sire is better?

Matrix Notation

$$\mathbf{y} = \begin{pmatrix} 128 \\ 150 \\ 90 \\ 110 \\ 120 \end{pmatrix}, \quad \mathbf{Xb} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \mu, \quad \mathbf{Zu} = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} H_1 \\ H_2 \\ S_1 \\ S_2 \\ S_3 \end{pmatrix}$$

$$\mathbf{G} = \begin{pmatrix} \frac{1}{6} & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{6} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{15} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{15} & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{15} \end{pmatrix} \sigma_e^2, \quad \mathbf{R} = \begin{pmatrix} \frac{1}{5} & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{4} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{5} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{20} & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} \end{pmatrix} \sigma_e^2$$

Prediction Problem

The problem is to predict the function

$$\mathbf{K}'\mathbf{b} + \mathbf{M}'\mathbf{u},$$

provided that $\mathbf{K}'\mathbf{b}$ is an estimable function by a linear function of $\mathbf{y}$,

$$\mathbf{L}'\mathbf{y}$$

such that

$$\begin{aligned} E(\mathbf{L}'\mathbf{y}) &= \mathbf{L}'\mathbf{X}\mathbf{b} \\ E(\mathbf{K}'\mathbf{b} + \mathbf{M}'\mathbf{u}) &= \mathbf{K}'\mathbf{b} \\ \mathbf{L}'\mathbf{X} &= \mathbf{K}' \\ \mathbf{L}'\mathbf{X} - \mathbf{K}' &= \mathbf{0} \end{aligned}$$

and minimize

$$\begin{aligned} Var(\mathbf{K}'\mathbf{b} + \mathbf{M}'\mathbf{u} - \mathbf{L}'\mathbf{y}) &= Var(\mathbf{M}'\mathbf{u} - \mathbf{L}'\mathbf{y}) \\ &= \mathbf{M}'\mathbf{G}\mathbf{M} + \mathbf{L}'\mathbf{V}\mathbf{L} - \mathbf{M}'\mathbf{G}\mathbf{Z}'\mathbf{L} - \mathbf{L}'\mathbf{Z}\mathbf{G}\mathbf{M} \end{aligned}$$

Function to be minimized

$$\mathbf{F} = Var(PE) + (\mathbf{L}'\mathbf{X} - \mathbf{K}')\mathbf{\Phi}$$

$$\begin{aligned} \frac{\partial \mathbf{F}}{\partial \mathbf{L}} &= 2\mathbf{VL} - 2\mathbf{ZGM} + \mathbf{X\Phi} = \mathbf{0} \\ \frac{\partial \mathbf{F}}{\partial \mathbf{\Phi}} &= \mathbf{X}'\mathbf{L} - \mathbf{K} = \mathbf{0} \end{aligned}$$

GLS

$$\hat{\mathbf{b}} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y},$$

then

$$\mathbf{L}'\mathbf{y} = \mathbf{K}'\hat{\mathbf{b}} + \mathbf{M}'\mathbf{G}\mathbf{Z}'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\hat{\mathbf{b}})$$

is BLUP.

Mixed Model Equations, MME

Take the first and second partial derivatives of **F**,

$$\begin{pmatrix} \mathbf{V} & \mathbf{X} \\ \mathbf{X}' & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{L} \\ \theta \end{pmatrix} = \begin{pmatrix} \mathbf{ZGM} \\ \mathbf{K} \end{pmatrix}$$

Recall $\mathbf{V} = Var(\mathbf{y}) = \mathbf{ZGZ}' + \mathbf{R}$,

$$\mathbf{S} = \mathbf{G}(\mathbf{Z}'\mathbf{L} - \mathbf{M})$$

re-arrange as

$$\mathbf{M} = \mathbf{Z}'\mathbf{L} - \mathbf{G}^{-1}\mathbf{S}$$

$$\begin{pmatrix} \mathbf{R} & \mathbf{X} & \mathbf{Z} \\ \mathbf{X}' & \mathbf{0} & \mathbf{0} \\ \mathbf{Z}' & \mathbf{0} & -\mathbf{G}^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{L} \\ \theta \\ \mathbf{S} \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{K} \\ \mathbf{M} \end{pmatrix}$$

MME - BLUP

$$\begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}+\mathbf{G}^{-1} \end{pmatrix} \begin{pmatrix} \hat{\mathbf{b}} \\ \hat{\mathbf{u}} \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \end{pmatrix}$$

Easier to solve usually. Henderson (1949)

$$\begin{pmatrix} \hat{\mathbf{b}} \\ \hat{\mathbf{u}} \end{pmatrix} = \begin{pmatrix} \mathbf{C}_{xx} & \mathbf{C}_{xz} \\ \mathbf{C}_{zx} & \mathbf{C}_{zz} \end{pmatrix} \begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1} \\ \mathbf{Z}'\mathbf{R}^{-1} \end{pmatrix} \mathbf{y}$$

Predictors

$$Var \begin{pmatrix} \hat{\mathbf{b}} \\ \hat{\mathbf{u}} \end{pmatrix} = \begin{pmatrix} \mathbf{C}_{xx} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}-\mathbf{C}_{zz} \end{pmatrix}$$

Prediction Errors

$$Var \begin{pmatrix} \hat{\mathbf{b}} \\ \hat{\mathbf{u}}-\mathbf{u} \end{pmatrix} = \begin{pmatrix} \mathbf{C}_{xx} & \mathbf{C}_{xz} \\ \mathbf{C}_{zx} & \mathbf{C}_{zz} \end{pmatrix}$$

Example Problem

$$\mathbf{y} = \begin{pmatrix} 128 \\ 150 \\ 90 \\ 110 \\ 120 \end{pmatrix}, \quad \mathbf{Xb} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \mu, \quad \mathbf{Zu} = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} H_1 \\ H_2 \\ S_1 \\ S_2 \\ S_3 \end{pmatrix}$$

$$\mathbf{G} = \begin{pmatrix} \frac{1}{6} & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{6} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{15} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{15} & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{15} \end{pmatrix} \sigma_e^2, \quad \mathbf{R} = \begin{pmatrix} \frac{1}{5} & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{4} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{5} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{20} & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} \end{pmatrix} \sigma_e^2$$

MME

$$\begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} \end{pmatrix} =$$

$$\begin{pmatrix} 36 & 9 & 27 & 10 & 24 & 2 \\ 9 & 9+6 & 0 & 5 & 4 & 0 \\ 27 & 0 & 27+6 & 5 & 20 & 2 \\ 10 & 5 & 5 & 10+15 & 0 & 0 \\ 24 & 4 & 20 & 0 & 24+15 & 0 \\ 2 & 0 & 2 & 0 & 0 & 2+15 \end{pmatrix}$$

$$\begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \end{pmatrix} = \begin{pmatrix} 4130 \\ 1240 \\ 2890 \\ 1090 \\ 2800 \\ 240 \end{pmatrix}$$

Solve

$$\begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} \end{pmatrix}^{-1} = \mathbf{C}$$

$$\begin{pmatrix} .1480 & -.0709 & -.0958 & -.0258 & -.0347 & -.0064 \\ -.0709 & .1100 & .0567 & -.0058 & .0033 & .0017 \\ -.0958 & .0567 & .1100 & .0058 & -.0033 & -.0017 \\ -.0258 & -.0058 & .0058 & .0503 & .0139 & .0025 \\ -.0347 & .0033 & -.0033 & .0139 & .0483 & .0045 \\ -.0064 & .0017 & -.0017 & .0025 & .0045 & .0597 \end{pmatrix}$$

$$\begin{aligned} \mu &= 119.673140 \\ H_1 &= 11.519036 \qquad & H_2 &= -11.519036 \\ S_1 &= -4.269256 \qquad & S_2 &= 2.875621 \\ S_3 &= 1.393635 \end{aligned}$$

Exercise

```
y = matrix(data=c(25.3,17.8,20.1,26.1,
 19.2,24.6,23.7,24.5,26.0,28.4,22.9,25.8),
 ncol=1)
RI = diag(c(40,39,37,35,40,40,33,38,40,30,28,34))
FB=c(1,1,1,2,2,2,3,3,3,4,4,4)
X =desgn(FB,4)
LL=c(1,2,3,1,2,3,1,2,3,1,2,3)
Z = desgn(LL,3)
GI = diag(c(10,10,10))
plants=MME(X,Z,GI,RI,y)
plants$SOLNS
plants$SEP
```

MME Function

```
MME = function(X,Z,GI,RI,y){
sst = t(y) %*% RI %*% y
XX = t(X) %*% RI %*% X
XZ = t(X) %*% RI %*% Z
ZZ = t(Z) %*% RI %*% Z
Xy = t(X) %*% RI %*% y
Zy = t(Z) %*% RI %*% y
N = length(y)
Q = ginv(XX) %*% XX
rX = sum(diag(Q))
R1 = cbind(XX,XZ)
R2 = cbind(t(XZ),(ZZ+GI))
LHS = rbind(R1,R2)
RHS = rbind(Xy,Zy)
```

MME function (continued)

```
C = ginv(LHS)
SOLNS = C %*% RHS
SSR = t(SOLNS) %*% RHS
resv = as.numeric(sst - SSR)/(N - rX)
SEP = sqrt(diag(C*resv))
sep = matrix(data=SEP,ncol=1)
return(list(LHS=LHS,RHS=RHS,RESV=resv,
C=C,SEP=sep,SOLNS=SOLNS))  }
```

Example Problem

$$\mathbf{y} = \begin{pmatrix} 128 \\ 150 \\ 90 \\ 110 \\ 120 \end{pmatrix}, \quad \mathbf{Xb} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \mu, \quad \mathbf{Zu} = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} H_1 \\ H_2 \\ S_1 \\ S_2 \\ S_3 \end{pmatrix}$$

$$\mathbf{G} = \begin{pmatrix} \frac{1}{6} & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{6} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{15} & \frac{1}{30} & \frac{1}{30} \\ 0 & 0 & \frac{1}{30} & \frac{1}{15} & \frac{1}{30} \\ 0 & 0 & \frac{1}{30} & \frac{1}{30} & \frac{1}{15} \end{pmatrix} \sigma_e^2, \quad \mathbf{R} = \begin{pmatrix} \frac{1}{5} & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{4} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{5} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{20} & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} \end{pmatrix} \sigma_e^2$$

MME

$$\begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} \end{pmatrix} =$$

$$\begin{pmatrix} 36 & 9 & 27 & 10 & 24 & 2 \\ 9 & 9+6 & 0 & 5 & 4 & 0 \\ 27 & 0 & 27+6 & 5 & 20 & 2 \\ 10 & 5 & 5 & 10+22.5 & -7.5 & -7.5 \\ 24 & 4 & 20 & -7.5 & 24+22.5 & -7.5 \\ 2 & 0 & 2 & -7.5 & -7.5 & 2+22.5 \end{pmatrix}$$

$$\begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \end{pmatrix} = \begin{pmatrix} 4130 \\ 1240 \\ 2890 \\ 1090 \\ 2800 \\ 240 \end{pmatrix}$$

Solve

$$\begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} \end{pmatrix}^{-1} = \mathbf{C}$$

$$\begin{pmatrix} .1666 & -.0709 & -.0958 & -.0458 & -.0518 & -.0357 \\ -.0709 & .1096 & .0570 & -.0031 & .0023 & .0009 \\ -.0958 & .0570 & .1096 & .0031 & -.0023 & -.0009 \\ -.0458 & -.0031 & .0031 & .0615 & .0380 & .0339 \\ -.0518 & .0023 & -.0023 & .0380 & .0607 & .0346 \\ -.0357 & .0009 & -.0009 & .0339 & .0346 & .0648 \end{pmatrix}$$

$$\begin{aligned} \mu &= 119.720701 \\ H_1 &= 11.204318 & H_2 &= -11.204318 \\ S_1 &= -2.680175 & S_2 &= 1.962449 \\ S_3 &= 0.717726 \end{aligned}$$

Residual Variance

Estimate σ_e^2

$$\hat{\sigma}_e^2 = (\mathbf{y}'\mathbf{y} - \hat{\mathbf{b}}'\mathbf{X}'\mathbf{y} - \hat{\mathbf{u}}'\mathbf{Z}'\mathbf{y})/(N - r(\mathbf{X}))$$

Solutions to MME times their corresponding right hand sides.
N must be greater than $r(\mathbf{X})$.
Single traits only.

Page 94 Exercise.