Chapter 22
Animal Model, Single Records

C. R. Henderson

1984 - Guelph

We shall describe a number of different genetic models and present methods for BLUE, BLUP, and estimation of variance and covariance components. The simplest situation is one in which we have only one trait of concern, we assume an additive genetic model, and no animal has more than a single record on this trait. The scalar model, that is, the model for an individual record, is

$$y_i = \mathbf{x}_i'\boldsymbol{\beta} + \mathbf{z}_i'\mathbf{u} + a_i + e_i.$$

$\boldsymbol{\beta}$ represents fixed effects with $\mathbf{x}_i$ relating the record on the i^{th} animal to this vector.

$\mathbf{u}$ represents random effects other than breeding values and $\mathbf{z}_i$ relates this vector to y_i.

a_i is the additive genetic value of the i^{th} animal.

e_i is a random error associated with the individual record.

The vector representation of the entire set of records is

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{Z}_a\mathbf{a} + \mathbf{e}. \quad (1)$$

If $\mathbf{a}$ represents only those animals with records, $\mathbf{Z}_a = \mathbf{I}$. Otherwise it is an identity matrix with rows deleted that correspond to animals without records.

$$\begin{aligned}
Var(\mathbf{u}) &= \mathbf{G}. \\
Var(\mathbf{a}) &= \mathbf{A}\sigma_a^2. \\
Var(\mathbf{e}) &= \mathbf{R}, \text{ usually } \mathbf{I}\sigma_e^2. \\
Cov(\mathbf{u}, \mathbf{a}') &= \mathbf{0}, \\
Cov(\mathbf{u}, \mathbf{e}') &= \mathbf{0}, \\
Cov(\mathbf{a}, \mathbf{e}') &= \mathbf{0}.
\end{aligned}$$

If $\mathbf{Z}_a \neq \mathbf{I}$, the mixed model equations are

$$\begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}_a \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}_a \\ \mathbf{Z}_a'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}_a'\mathbf{R}^{-1}\mathbf{Z} & \mathbf{Z}_a'\mathbf{R}^{-1}\mathbf{Z}_a + \mathbf{A}^{-1}/\sigma_a^2 \end{pmatrix}
\begin{pmatrix} \boldsymbol{\beta}^o \\ \hat{\mathbf{u}} \\ \hat{\mathbf{a}} \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}_a'\mathbf{R}^{-1}\mathbf{y} \end{pmatrix}. \quad (2)$$

If $\mathbf{Z}_a = \mathbf{I}$, (22.2) simplifies to

$$\begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} & \mathbf{X}'\mathbf{R}^{-1} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}+\mathbf{G}^{-1} & \mathbf{Z}'\mathbf{R}^{-1} \\ \mathbf{R}^{-1}\mathbf{X} & \mathbf{R}^{-1}\mathbf{Z} & \mathbf{R}^{-1}+\mathbf{A}^{-1}/\sigma_a^2 \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \hat{\mathbf{u}} \\ \hat{\mathbf{a}} \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{R}^{-1}\mathbf{y} \end{pmatrix}. \quad (3)$$

If $\mathbf{R} = \mathbf{I}\sigma_e^2$ (22.3) simplifies further to

$$\begin{pmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z} & \mathbf{X}' \\ \mathbf{Z}'\mathbf{X} & \mathbf{Z}'\mathbf{Z}+\mathbf{G}^{-1}\sigma_e^2 & \mathbf{Z}' \\ \mathbf{X} & \mathbf{Z} & \mathbf{I}+\mathbf{A}^{-1}\sigma_e^2/\sigma_a^2 \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}^o \\ \hat{\mathbf{u}} \\ \hat{\mathbf{a}} \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{Z}'\mathbf{y} \\ \mathbf{y} \end{pmatrix}. \quad (4)$$

If the number of animals is large, one should, of course, use Henderson's method (1976) for computing $\mathbf{A}^{-1}$. Because this method requires using a "base" population of non-inbred, unrelated animals, some of these probably do not have records. Also we may wish to evaluate some progeny that have not yet made a record. Both of these circumstances will result in $\mathbf{Z}_a \neq \mathbf{I}$, but $\hat{\mathbf{a}}$ will contain predicted breeding values of these animals without records.

1 Example With Dam-Daughter Pairs

We illustrate the model above with 5 pairs of dams and daughters, the dams' records being made in period 1 and the daughters' in period 2. Ordering the records within periods and with record 1 being made by the dam of the individual making record 6, etc.

$$\begin{aligned} \mathbf{X}' &= \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}, \\ \mathbf{Z}_a &= \mathbf{I}, \\ \mathbf{y}' &= [5, 4, 3, 2, 6, 6, 7, 3, 5, 4]. \\ \mathbf{A} &= \begin{pmatrix} \mathbf{I}_5 & .5\mathbf{I}_5 \\ .5\mathbf{I}_5 & \mathbf{I}_5 \end{pmatrix}, \\ \mathbf{R} &= \mathbf{I}_{10}\sigma_e^2. \end{aligned}$$

The sires and dams are all unrelated. We write the mixed model equations with σ_e^2/σ_a^2 assumed to be 5. These equations are

$$\left(\begin{array}{cccccccccccc}
5 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 5 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\
1 & 0 & & \frac{23}{3} & \mathbf{I}_5 & & & \frac{-10}{3} & \mathbf{I}_5 & & & \\
1 & 0 & & & & & & & & & & \\
1 & 0 & & & & & & & & & & \\
1 & 0 & & & & & & & & & & \\
1 & 0 & & & & & & & & & & \\
0 & 1 & & & & & & & & & & \\
0 & 1 & & & & & & & & & & \\
0 & 1 & & \frac{-10}{3} & \mathbf{I}_5 & & & \frac{23}{3} & \mathbf{I}_5 & & & \\
0 & 1 & & & & & & & & & & \\
0 & 1 & & & & & & & & & &
\end{array}\right)
\begin{pmatrix} p_1 \\ p_2 \\ \mathbf{a} \end{pmatrix} =
\begin{pmatrix} 20 \\ 25 \\ 5 \\ 4 \\ 3 \\ 2 \\ 6 \\ 6 \\ 7 \\ 3 \\ 5 \\ 4 \end{pmatrix}. \qquad (5)$$

The inverse of the coefficient matrix is

$$\left(\begin{array}{cccccccccccc}
.24 & .02 & -.04 & -.04 & -.04 & -.04 & -.04 & -.02 & -.02 & -.02 & -.02 & -.02 \\
.02 & .24 & -.02 & -.02 & -.02 & -.02 & -.02 & -.04 & -.04 & -.04 & -.04 & -.04 \\
-.04 & -.02 & & & & & & & & & & \\
-.04 & -.02 & & & & & & & & & & \\
-.04 & -.02 & & & \mathbf{P} & & & & & \mathbf{Q} & & \\
-.04 & -.02 & & & & & & & & & & \\
-.04 & -.02 & & & & & & & & & & \\
-.02 & -.04 & & & & & & & & & & \\
-.02 & -.04 & & & & & & & & & & \\
-.02 & -.04 & & & \mathbf{Q} & & & & & \mathbf{P} & & \\
-.02 & -.04 & & & & & & & & & & \\
-.02 & -.04 & & & & & & & & & &
\end{array}\right) \qquad (6)$$

$\mathbf{P}$ is a 5×5 matrix with .16867 in diagonals and .00783 in all off-diagonals. $\mathbf{Q}$ is a 5×5 matrix with .07594 in diagonals and .00601 in off-diagonals. The solution is (4, 5, .23077, .13986, -.30070, -.32168, .25175, .23077, .32168, -.39161, -.13986, -.20298).

Let us estimate σ_e^2, σ_a^2 by MIVQUE using the prior on $\sigma_e^2/\sigma_a^2 = 5$ as we did in computing BLUP. The quadratics needed are

$$\hat{\mathbf{e}}'\hat{\mathbf{e}} \text{ and } \hat{\mathbf{a}}'\mathbf{A}^{-1}\hat{\mathbf{a}}.$$

$$\hat{\mathbf{e}} = \mathbf{y} - (\mathbf{X} \;\; \mathbf{Z}_a)\begin{pmatrix} \hat{\mathbf{p}} \\ \hat{\mathbf{a}} \end{pmatrix} = (.76923,\ -.13986,\ -.69930,\ -1.67832,\ 1.74825,$$
$$.76923, 1.67832,\ -1.60839,\ .13986, -.97902)'.$$

$$
\begin{aligned}
\hat{\mathbf{e}}'\hat{\mathbf{e}} &= 13.94689. \\
Var(\hat{\mathbf{e}}) &= (\mathbf{I} - \mathbf{WCW}')(\mathbf{I} - \mathbf{WCW}')\ \sigma_e^2 \\
&\quad + (\mathbf{I} - \mathbf{WCW}')\mathbf{A}(\mathbf{I} - \mathbf{WCW}')\ \sigma_a^2. \\
\mathbf{W} &= (\mathbf{X}\ \ \mathbf{Z}_a), \\
\mathbf{C} &= \text{matrix of (22.6)}. \\
E(\hat{\mathbf{e}}'\hat{\mathbf{e}}) &= tr(Var(\hat{\mathbf{e}})) = 5.67265\ \sigma_e^2 + 5.20319\ \sigma_a^2 \\
\hat{\mathbf{a}}' &= \mathbf{C}_a\mathbf{W}'\mathbf{y}, \\
\text{where } \mathbf{C}_a &= \text{last 10 rows of } \mathbf{C}. \\
Var(\hat{\mathbf{a}}) &= \mathbf{C}_a\mathbf{W}'\mathbf{W}\mathbf{C}_a'\ \sigma_e^2 + \mathbf{C}_a\mathbf{W}'\mathbf{A}\mathbf{W}\mathbf{C}_a'\ \sigma_a^2. \\
E(\hat{\mathbf{a}}'\mathbf{A}^{-1}\hat{\mathbf{a}}) &= tr(\mathbf{A}^{-1}\ Var(\hat{\mathbf{a}})) = .20813\ \sigma_e^2 + .24608\ \sigma_a^2. \\
\hat{\mathbf{a}}'\mathbf{A}^{-1}\hat{\mathbf{a}} &= .53929.
\end{aligned}
$$

Using these quadratics and solving for $\hat{\sigma}_e^2, \hat{\sigma}_a^2$ we obtain $\hat{\sigma}_e^2 = 2.00$, $\hat{\sigma}_a^2 = .50$.

The same estimates are obtained for any prior used for σ_e^2/σ_a^2. This is a consequence of the fact that we have a balanced design. Therefore the estimates are truly BQUE and also are REML. Further the traditional method, daughter-dam regression, gives the same estimates. These are

$$
\begin{aligned}
\hat{\sigma}_a^2 &= \text{2 times regression of daughter on dam}. \\
\hat{\sigma}_e^2 &= \text{within period mean square } - \hat{\sigma}_a^2.
\end{aligned}
$$

For unbalanced data MIVQUE is not invariant to the prior used, and daughter-dam regression is neither MIVQUE nor REML. We illustrate by assuming that y_{10} was not observed. With σ_e^2/σ_a^2 assumed equal to 2 we obtain

$$
\begin{aligned}
\hat{\mathbf{e}}'\hat{\mathbf{e}} &= 11.99524 \text{ with expectation } = 3.37891\ \sigma_e^2 + 2.90355\ \sigma_a^2. \\
\hat{\mathbf{a}}'\mathbf{A}^{-1}\mathbf{a}' &= 2.79712 \text{ with expectation } = .758125\ \sigma_e^2 + .791316\ \sigma_a^2.
\end{aligned}
$$

This gives

$$
\begin{aligned}
\hat{\sigma}_a^2 &= .75619, \\
\hat{\sigma}_e^2 &= 2.90022.
\end{aligned}
$$

When σ_e^2/σ_a^2 is assumed equal to 5, the results are

$$
\begin{aligned}
\hat{\mathbf{e}}'\hat{\mathbf{e}} &= 16.83398 \text{ with expectation } 4.9865\ \sigma_e^2 + 4.6311\ \sigma_a^2, \\
\hat{\mathbf{a}}'\mathbf{A}^{-1}\hat{\mathbf{a}}' &= .66973 \text{ with expectation } .191075\ \sigma_e^2 + .214215\ \sigma_a^2.
\end{aligned}
$$

Then

$$
\begin{aligned}
\hat{\sigma}_a^2 &= .67132, \\
\hat{\sigma}_e^2 &= 2.7524.
\end{aligned}
$$