

Chapter 28

Joint Cow and Sire Evaluation

C. R. Henderson

1984 - Guelph

At the present time, (1984), agencies evaluating dairy sires and dairy females have designed separate programs for each. Sires usually have been evaluated solely on the production records of their progeny. With the development of an easy method for computing \mathbf{A}^{-1} this matrix has been incorporated by some agencies, and that results in the evaluation being a combination of the progeny test of the individual in question as well as progeny tests of his relatives, eg., sire and paternal brothers. In addition, the method also takes into account the predictions of the merits of the sires of the mates of the bulls being evaluated. This is an approximation to the merits of the mates without using their records.

In theory one could utilize all records available in a production testing program and could compute \mathbf{A}^{-1} for all animals that have produced these records as well as additional related animals without records that are to be evaluated. Then these could be incorporated into a single set of prediction equations. This, of course, could result in a set of equations that would be much too large to solve with existing computers. Nevertheless, if we are willing to sacrifice some accuracy by ignoring the fact that animals change herds, we can set up equations that are block diagonal in form that may be feasible to solve.

1 Block Diagonality Of Mixed Model Equations

Henderson (1976) presented a method for rapid calculation of \mathbf{A}^{-1} without computing \mathbf{A} . A remarkable property of \mathbf{A}^{-1} is that the only non-zero off-diagonal elements are those pertaining to a pair of mates, and those pertaining to parent - progeny pairs. These non-zero elements can be built up by entering the data in any order, with each piece of data incorporating the individual identification number, the sire number, and the dam number. At the same time one could enter with this information the production record and elements of the incidence matrix of the individual. Now when the dam and her progeny are in different herds, we pretend that we do not know the dam of the progeny and if, when a natural service sire has progeny in more than one herd, we treat him as a different sire in each herd, there are no non-zero elements of \mathbf{A}^{-1} between herds. This strategy, along with the fact that most if not all elements of $\boldsymbol{\beta}$ are peculiar to the individual herd, results in the mixed model coefficient matrix having a block diagonal form. The elements of the model are ordered as follows

β_0 : a subvector of β common to all elements of \mathbf{y} .

\mathbf{a}_0 : a subvector of \mathbf{a} , additive genetic values, pertaining to sires used in several herds.

β_i ($i = 1, \dots$, number of herds): a subvector of β pertaining only to records in the i^{th} herd.

\mathbf{a}_i : a subvector of \mathbf{a} pertaining to animals in the i^{th} herd. \mathbf{a}_i can represent cows with records, or dams and non-AI sires of the cows with records. In computing \mathbf{A}^{-1} for the animals in the i^{th} herd the dam is assumed unknown if it is in a different herd. When Section 28.3 method is used (multiple records) no records of a cow should be used in a herd unless the first lactation record is available. This restriction prevents using records of a cow that moves to another herd subsequent to first lactation. With this ordering and with these restrictions in computing \mathbf{A}^{-1} the BLUP equations have the following form

$$\begin{pmatrix} \mathbf{C}_{00} & \mathbf{C}_{01} & \mathbf{C}_{02} & \cdots & \mathbf{C}_{0k} \\ \mathbf{C}'_{01} & \mathbf{C}_{11} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{C}'_{02} & \mathbf{0} & \mathbf{C}_{22} & \cdots & \mathbf{0} \\ \vdots & & & & \vdots \\ \mathbf{C}'_{0k} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{C}_{kk} \end{pmatrix} \begin{pmatrix} \gamma_0 \\ \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_k \end{pmatrix} = \begin{pmatrix} \mathbf{r}_0 \\ \mathbf{r}_1 \\ \mathbf{r}_2 \\ \vdots \\ \mathbf{r}_k \end{pmatrix}.$$

$$\begin{aligned} \gamma'_0 &= (\beta'_0 \ \mathbf{a}'_0), \\ \gamma'_i &= (\beta'_i \ \mathbf{a}'_i). \end{aligned}$$

Then with this form of the equations the herd unknowns can be “absorbed” into the β_0 and \mathbf{a}_0 equations provided the \mathbf{C}_{ii} blocks can be readily inverted. Otherwise one would need to solve iteratively. For example, one might first solve iteratively for β_0 and \mathbf{a}_0 sires ignoring β_i , \mathbf{a}_i . Then with these values one would solve iteratively for the herd values. Having obtained these one would re-solve for β_0 and the \mathbf{a}_0 values, adjusting the right hand sides for the previously estimated herd values.

The AI sire equations would also contain values for the “base population” sires. A base population dam with records would be included with the herd in which its records were made. Any base population dam that has no records, has only one AI son, and has no female progeny can be ignored without changing the solution.

2 Single Record On Single Trait

The simplest example of joint cow and sire evaluation with multiple herds involves a single trait and with only one record per tested animal. We illustrate this with the following example.

Base population animals

- 1 male
- 2 female with record in herd 1
- 3 female with record in herd 2

AI Sires

- 4 with parents 1 and 2
- 5 with parents 1 and 3

Other Females With Records

- 6 with unknown parents, record in herd 1
- 7 with unknown parents, record in herd 2
- 8 with parents 4 and 6, record in herd 1
- 9 with parents 4 and 3, record in herd 2
- 10 with parents 5 and 7, record in herd 2
- 11 with parents 5 and 2, record in herd 1

Ordering these animals (1,4,5,2,6,8,11,3,7,9,10) the \mathbf{A} matrix is in (28.1).

$$\begin{pmatrix} 1 & .5 & .5 & 0 & 0 & .25 & .25 & 0 & 0 & .25 & .25 \\ & 1 & .25 & .5 & 0 & .5 & .375 & 0 & 0 & .5 & .125 \\ & & 1 & 0 & 0 & .125 & .5 & .5 & 0 & .375 & .5 \\ & & & 1 & 0 & .25 & .5 & 0 & 0 & .25 & 0 \\ & & & & 1 & .5 & 0 & 0 & 0 & 0 & 0 \\ & & & & & 1 & .1875 & 0 & 0 & .25 & .0625 \\ & & & & & & 1 & .25 & 0 & .3125 & .25 \\ & & & & & & & 1 & 0 & .5 & .25 \\ & & & & & & & & 1 & 0 & .5 \\ & & & & & & & & & 1 & .1875 \\ & & & & & & & & & & 1 \end{pmatrix} \quad (1)$$

\mathbf{A}^{-1} shown in (28.2)

$$\begin{pmatrix} 2 & -1 & -1 & .5 & 0 & 0 & 0 & .5 & 0 & 0 & 0 \\ & 3 & 0 & -1 & .5 & -1 & 0 & .5 & 0 & -1 & 0 \\ & & 3 & .5 & 0 & 0 & -1 & -1 & .5 & 0 & -1 \\ & & & 2 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ & & & & 1.5 & -1 & 0 & 0 & 0 & 0 & 0 \\ & & & & & 2 & 0 & 0 & 0 & 0 & 0 \\ & & & & & & 2 & 0 & 0 & 0 & 0 \\ & & & & & & & 2 & 0 & -1 & 0 \\ & & & & & & & & 1.5 & 0 & -1 \\ & & & & & & & & & 2 & 0 \\ & & & & & & & & & & 2 \end{pmatrix} \quad (2)$$

Note that the lower 8×8 submatrix is block diagonal with two blocks of order 4×4 down the diagonal and 4×4 null off-diagonal blocks. The model assumed for our illustration is

$$y_{ij} = \mu_i + a_{ij} + e_{ij},$$

where i refers to herd and j to individual within herd. Then with ordering

$$(a_1, a_4, a_5, \mu_1, a_2, a_6, a_8, a_{11}, \mu_2, a_3, a_7, a_9, a_{10})$$

the incidence matrix is as shown in (28.3). Note that β^o does not exist.

$$\begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{pmatrix} \quad (3)$$

Suppose $\mathbf{y}' = [3, 2, 5, 6, 7, 9, 2, 3]$ corresponding to animals 2, 6, 8, 11, 3, 7, 9, 10. We assume that $h^2 = .25$ which implies $\sigma_e^2/\sigma_a^2 = 3$. Then adding $3 \mathbf{A}^{-1}$ to appropriate elements of the OLS equations we obtain mixed model equations displayed in (28.4).

$$\begin{pmatrix}
6 & -3 & -3 & 0 & 1.5 & 0 & 0 & 0 & 0 & 1.5 & 0 & 0 & 0 \\
& 9 & 0 & 0 & -3 & 1.5 & -3 & 0 & 0 & 1.5 & 0 & -3 & 0 \\
& & 9 & 0 & 1.5 & 0 & 0 & -3 & 0 & -3 & 1.5 & 0 & -3 \\
& & & 4 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
& & & & 7 & 0 & 0 & -3 & 0 & 0 & 0 & 0 & 0 \\
& & & & & 5.5 & -3 & 0 & 0 & 0 & 0 & 0 & 0 \\
& & & & & & 7 & 0 & 0 & 0 & 0 & 0 & 0 \\
& & & & & & & 7 & 0 & 0 & 0 & 0 & 0 \\
& & & & & & & & 4 & 1 & 1 & 1 & 1 \\
& & & & & & & & & 7 & 0 & -3 & 0 \\
& & & & & & & & & & 5.5 & 0 & -3 \\
& & & & & & & & & & & 7 & 0 \\
& & & & & & & & & & & & 7
\end{pmatrix}
\begin{pmatrix}
\hat{a}_1 \\
\hat{a}_4 \\
\hat{a}_5 \\
\hat{\mu}_1 \\
\hat{a}_2 \\
\hat{a}_6 \\
\hat{a}_8 \\
\hat{a}_{11} \\
\hat{\mu}_2 \\
\hat{a}_3 \\
\hat{a}_7 \\
\hat{a}_9 \\
\hat{a}_{10}
\end{pmatrix}
=
\begin{pmatrix}
0 \\
0 \\
0 \\
16 \\
3 \\
2 \\
5 \\
6 \\
21 \\
7 \\
9 \\
2 \\
3
\end{pmatrix}
\tag{4}$$

Note that the lower 10×10 block of the coefficient matrix is block diagonal with 5×5 blocks down the diagonal and 5×5 null blocks off-diagonal. The solution to (28.4) is

$[-.1738, -.3120, -.0824, 4.1568, -.1793, -.4102, -.1890, .1512, 5.2135, .0857, .6776, -.5560, -.0611]$.

Note that the solution to (a_1, a_4, a_5) could be found by absorbing the other equations as follows.

$$\left\{ \begin{pmatrix} 6 & -3 & -3 \\ & 9 & 0 \\ & & 9 \end{pmatrix} - \begin{pmatrix} 0 & 1.5 & 0 & 0 & 0 \\ 0 & -3 & 1.5 & -3 & 0 \\ 0 & 1.5 & 0 & 0 & -3 \end{pmatrix} \begin{pmatrix} 4 & 1 & 1 & 1 & 1 \\ & 7 & 0 & 0 & -3 \\ & & 5.5 & -3 & 0 \\ & & & 7 & 0 \\ & & & & 7 \end{pmatrix}^{-1} \right.$$

$$\begin{pmatrix} 0 & 0 & 0 \\ 1.5 & -3 & 1.5 \\ 0 & 1.5 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & -3 \end{pmatrix} - \begin{pmatrix} 0 & 1.5 & 0 & 0 & 0 \\ 0 & 1.5 & 0 & -3 & 0 \\ 0 & -3 & 1.5 & 0 & -3 \end{pmatrix} \begin{pmatrix} 4 & 1 & 1 & 1 & 1 \\ & 7 & 0 & -3 & 0 \\ & & 5.5 & 0 & -3 \\ & & & 7 & 0 \\ & & & & 7 \end{pmatrix}^{-1} \\
\left. \begin{pmatrix} 0 & 0 & 0 \\ 1.5 & 1.5 & -3 \\ 0 & 0 & 1.5 \\ 0 & -3 & 0 \\ 0 & 0 & -3 \end{pmatrix} \right\} \begin{pmatrix} a_1 \\ a_4 \\ a_5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} - \begin{pmatrix} 0 & 1.5 & 0 & 0 & 0 \\ 0 & -3 & 1.5 & -3 & 0 \\ 0 & 1.5 & 0 & 0 & -3 \end{pmatrix} \\
\begin{pmatrix} 4 & 1 & 1 & 1 & 1 \\ & 7 & 0 & 0 & -3 \\ & & 5.5 & -3 & 0 \\ & & & 7 & 0 \\ & & & & 7 \end{pmatrix}^{-1} \begin{pmatrix} 16 \\ 3 \\ 2 \\ 5 \\ 6 \end{pmatrix} - \begin{pmatrix} 0 & 1.5 & 0 & 0 & 0 \\ 0 & 1.5 & 0 & -3 & 0 \\ 0 & -3 & 1.5 & 0 & -3 \end{pmatrix} \\
\begin{pmatrix} 4 & 1 & 1 & 1 & 1 \\ & 7 & 0 & -3 & 0 \\ & & 5.5 & 0 & -3 \\ & & & 7 & 0 \\ & & & & 7 \end{pmatrix}^{-1} \begin{pmatrix} 21 \\ 7 \\ 9 \\ 2 \\ 3 \end{pmatrix}.$$

Iterations on these equations were carried out by two different methods. First, the herd equations were iterated 5 rounds with AI sire values fixed. Then the AI sire equations were iterated 5 rounds with the herd values fixed and so on. It required 17 cycles (85 rounds) to converge to the direct solution previously reported. Regular Gauss-Seidel iteration produced convergence in 33 rounds. The latter procedure would require more retrieval of data from external storage devices.

3 Simple Repeatability Model

As our next example we use the same animals as before but now we have records as follows.

	Herd 1				Herd 2		
	Years				Years		
Cow	1	2	3	Cow	1	2	3
2	5	6	-	3	8	-	-
6	4	5	3	7	9	8	7
8	-	7	6	9	-	8	8
11	-	-	8	10	-	-	7

We assume a model,

$$y_{ijk} = \mu_{ij} + a_{ik} + p_{ik} + e_{ijk}.$$

i refers to herd, j to year, and k to cow. It is assumed that $h^2 = .25$, $r = .45$. Then

$$\begin{aligned} \text{Var}(\mathbf{a}) &= \mathbf{A} \sigma_e^2 / 2.2. \\ \text{Var}(\mathbf{p}) &= \mathbf{I} \sigma_e^2 / 2.75. \\ \sigma_e^2 / \sigma_a^2 &= 2.2, \quad \sigma_e^2 / \sigma_p^2 = 2.75 \end{aligned}$$

The diagonal coefficients of the \mathbf{p} equations of OLS have added to them 2.75. Then $\hat{\mathbf{p}}$ can be absorbed easily. This can be done without writing the complete equations by weighting each observation by

$$\frac{2.75}{n_{ik} + \sigma_e^2 / 2.75}$$

where n_{ik} is the number of records on the ik^{th} cow. These weights are .733, .579, .478 for 1,2,3 records respectively. Once these equations are derived, we then add $2.2 \mathbf{A}^{-1}$ to appropriate coefficients to obtain the mixed model equations. The coefficient matrix is in (28.5) ... (28.7), and the right hand side vector is (0, 0, 0, 4.807, 9.917, 10.772, 6.369, 5.736, 7.527, 5.864, 10.166, 8.456, 13.109, 5.864, 11.472, 9.264, 5.131)'. The unknowns are in this order ($a_1, a_4, a_5, \mu_{11}, \mu_{12}, \mu_{13}, a_2, a_6, a_8, a_{11}, \mu_{21}, \mu_{22}, \mu_{23}, a_3, a_7, a_9, a_{10}$). Note that block diagonality has been retained. The solution is

[.1956, .3217, .2214, 4.6512, 6.127, 5.9586, .2660, -.5509, .0045, .5004, 8.3515, 7.8439, 7.1948, .0377, .0516, .2424, .0892].

Upper 8×8

$$\begin{pmatrix} 4.4 & -2.2 & -2.2 & 0 & 0 & 0 & 1.1 & 0 \\ & 6.6 & 0 & 0 & 0 & 0 & -2.2 & 1.1 \\ & & 6.6 & 0 & 0 & 0 & 1.1 & 0 \\ & & & 1.057 & 0 & 0 & .579 & .478 \\ & & & & 1.636 & 0 & .579 & .478 \\ & & & & & 1.79 & 0 & .478 \\ & & & & & & 5.558 & 0 \\ & & & & & & & 4.734 \end{pmatrix} \quad (5)$$

Upper right 8×9 and (lower left 9×8)'

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1.1 & 0 & 0 & 0 \\ -2.2 & 0 & 0 & 0 & 0 & 1.1 & 0 & -2.2 & 0 \\ 0 & -2.2 & 0 & 0 & 0 & -2.2 & 1.1 & 0 & -2.2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ .579 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ .579 & .733 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -2.2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -2.2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (6)$$

Lower right 9×9

$$\begin{pmatrix} 5.558 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ & 5.133 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & 1.211 & 0 & 0 & .733 & .478 & 0 & 0 \\ & & & 1.057 & 0 & 0 & .478 & .579 & 0 \\ & & & & 1.79 & 0 & .478 & .579 & .733 \\ & & & & & 5.133 & 0 & -2.2 & 0 \\ & & & & & & 4.734 & 0 & -2.2 \\ & & & & & & & 5.558 & 0 \\ & & & & & & & & 5.133 \end{pmatrix} \quad (7)$$

4 Multiple Traits

As a final example of joint cow and sire evaluation we evaluate on two traits. Using the same animals as before the records are as follows.

	Herd 1			Herd 2	
	Trait			Trait	
Cow	1	2	Cow	1	2
2	6	8	3	7	-
6	4	6	7	-	2
8	9	-	9	-	8
11	-	3	10	6	9

We assume a model,

$$y_{ijk} = \mu_{ij} + a_{ijk} + e_{ijk},$$

where i refers to herd, j to trait, and k to cow. We assume that the error variance-covariance matrix for a cow and the additive genetic variance-covariance matrix for a non-inbred individual are

$$\begin{pmatrix} 5 & 2 \\ 2 & 8 \end{pmatrix} \text{ and } \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix},$$

respectively. Then \mathbf{R} is

$$\begin{pmatrix} 5 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ & 8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & 5 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & & 8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & & & 5 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & & & & 8 & 0 & 0 & 0 & 0 & 0 \\ & & & & & & 5 & 0 & 0 & 0 & 0 \\ & & & & & & & 8 & 0 & 0 & 0 \\ & & & & & & & & 8 & 0 & 0 \\ & & & & & & & & & 5 & 2 \\ & & & & & & & & & & 8 \end{pmatrix}.$$

Ordering traits within animals, \mathbf{G} is composed of 2×2 blocks as follows

$$\begin{pmatrix} 2 & a_{ij} & a_{ij} \\ & a_{ij} & 3 a_{ij} \end{pmatrix}.$$

The right hand sides of the mixed model equations are (0, 0, 0, 0, 0, 0, 3.244, 1.7639, .8889, .7778, .5556, .6111, 1.8, 0, 0, .375, 2.3333, 2.1167, 1.4, 0, 0, .25, 0, 1., .8333, .9167) corresponding to ordering of equations,

$[a_{11}, a_{12}, a_{41}, a_{42}, a_{51}, a_{52}, \mu_{11}, \mu_{12}, a_{21}, a_{22}, a_{61}, a_{62}, a_{81}, a_{82}, a_{11,1}, a_{11,2}, \mu_{21}, \mu_{22}, a_{31}, a_{32}, a_{71}, a_{72}, a_{91}, a_{92}, a_{10,1}, a_{10,2}]$.

The coefficient matrix is block diagonal with two 10×10 blocks in the lower diagonal and with two 10×10 null blocks off-diagonal. The solution is

(.2087, .1766, .4469, .4665, .1661, .1912, 5.9188, 5.9184, .1351, .3356, -.1843, .1314, .6230, .5448, -.0168, -.2390, 6.0830, 6.5215, .2563, .2734, -.4158, -.9170, .4099, .5450, -.0900, .0718).

5 Summary Of Methods

The model to be used contains the following elements.

1. $\mathbf{X}_0\boldsymbol{\beta}_0$: pertaining to all records
2. $\mathbf{X}_i\boldsymbol{\beta}_i$: pertaining to records only on the i^{th} herd.
3. $\mathbf{Z}_0\mathbf{a}_0$: additive genetic values of sires used in several herds, AI sires in particular, but could include natural service sires used in several herds.

4. $\mathbf{Z}_i\mathbf{a}_i$: additive genetic values of all females that have made records in the i^{th} herd. Some of these may be dams of AI sires. Others will be daughters of AI sires, and some will be both dams and daughters of different AI sires. $\mathbf{Z}_i\mathbf{a}_i$ will also contain any sire with daughters only in the i^{th} herd or with daughters in so few other herds that this is ignored, and he is regarded as a different sire in each of the other herds. One will need to decide how to handle such sires, that is, how many to include with AI sires and how many to treat as a separate sire in each of the herds in which he has progeny.
5. \mathbf{A}^{-1} should be computed by Henderson's simple method, possibly ignoring inbreeding in large data sets, since this reduces computations markedly. In order to generate block diagonality in the mixed model equations the elements of \mathbf{A}^{-1} for animals in $\mathbf{Z}_i\mathbf{a}_i$ should be derived only from sires in \mathbf{a}_0 and from dams and sires in \mathbf{a}_i (same herd). This insures that there will be no non-zero elements of \mathbf{A}^{-1} between any pair of herds, provided ordering is done according to the following

- (1) $\mathbf{X}_0\boldsymbol{\beta}_0$
- (2) $\mathbf{Z}_0\mathbf{a}_0$
- (3) $\mathbf{X}_1\boldsymbol{\beta}_1$
- (4) $\mathbf{Z}_1\mathbf{a}_1$
- (5) $\mathbf{X}_2\boldsymbol{\beta}_2$
- (6) $\mathbf{Z}_2\mathbf{a}_2$
- ⋮
- etc.

6 Gametic Model To Reduce The Number Of Equations

Quaas and Pollak (1980) described a gametic additive genetic model that reduces the number of equations needed for computing BLUP. The only breeding values appearing in the equations are those of animals having tested progeny. Then individuals with no progeny can be evaluated by taking appropriate linear functions of the solution vector. The paper cited above dealt with multiple traits. We shall consider two situations, (1) single traits with one or no record per trait and (2) single traits with multiple records and the usual repeatability model assumed. If one does not choose to assume the repeatability model, the different records in a trait can be regarded as multiple traits and the Quaas and Pollak method used.

6.1 Single record model

Let the model be

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_a\mathbf{a} + \text{other possible random factors} + \mathbf{e}.$$

There are b animals with tested progeny, and $c \leq b$ of these parents are tested. There are d tested animals with no progeny. Thus \mathbf{y} has $c + d$ elements. In the regular mixed model \mathbf{a} has $b + d$ elements. \mathbf{Z}_a is formed from an identity matrix of order $b + d$ and then deleting $b - c$ rows corresponding to parents with no record.

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

Now in the gametic model, which is linearly equivalent to the model above, \mathbf{a} has only b elements corresponding to the animals with tested progeny. As before \mathbf{y} has $c + d$ elements, and is ordered such that records of animals with progeny appear first.

$$\mathbf{Z}_a = \begin{pmatrix} \mathbf{P} \\ \mathbf{Q} \end{pmatrix}.$$

\mathbf{P} is a $c \times b$ matrix formed from an identity matrix of order, b , by deleting $b - c$ rows corresponding to parents without a record. \mathbf{Q} is a $d \times b$ matrix with all null elements except the following. For the i^{th} individual .5 is inserted in the i^{th} row of \mathbf{Q} in columns corresponding to its parents in the \mathbf{a} vector. Thus if both parents are present, the row contains two “.5’s”. If only one parent is present, the row contains one “.5”. If neither parent is present, the row is null. Now, of course, \mathbf{A} has order, b , referring to those animals with tested progeny. $\text{Var}(\mathbf{e})$ is no longer $\mathbf{I}\sigma_e^2$. It is diagonal with diagonal elements as follows for noninbred animals.

- (1) σ_e^2 for parents.
- (2) $\sigma_e^2 + .5 \sigma_a^2$ for progeny with both parents in \mathbf{a} .
- (3) $\sigma_e^2 + .75 \sigma_a^2$ for progeny with one parents in \mathbf{a} .
- (4) $\sigma_e^2 + \sigma_a^2$ for progeny with no parent in \mathbf{a} .

This model results in d less equations than in the usual model and a possible large reduction in time required for a solution to the mixed model equations.

Computation of \hat{a}_i , BLUP of a tested individual not in the solution for \mathbf{a} but providing data in \mathbf{y} , is simple.

$$\hat{e}_i = y_i - \mathbf{x}'_i\boldsymbol{\beta}^o - \mathbf{z}'_i\hat{\mathbf{u}} - .5 (\text{Sum of parental } \hat{\mathbf{a}}).$$

\mathbf{x}'_i is the incidence matrix for the i^{th} animal with respect to $\boldsymbol{\beta}$.

\mathbf{z}'_i is the incidence matrix for the i^{th} animal with respect to $\hat{\mathbf{u}}$, other random factors in the model. Then

$$\begin{aligned}\hat{a}_i &= .5 (\text{sum of parental } \hat{\mathbf{a}}) + k_i \hat{e}_i, \\ \text{where } k_i &= .5 \sigma_a^2 / (.5 \sigma_a^2 + \sigma_e^2) \text{ if both parents known,} \\ &= .75 \sigma_a^2 / (.75 \sigma_a^2 + \sigma_e^2) \text{ if one parent known,} \\ &= \sigma_a^2 / (\sigma_a^2 + \sigma_e^2) \text{ if neither parent known.}\end{aligned}$$

The solution for an animal with no record and no progeny is .5 (sum of parental $\hat{\mathbf{a}}$), provided these parents, if known, are included in the b elements of $\hat{\mathbf{a}}$ in the solution.

A simple sire model for single traits can be considered a special case of this model. The incidence matrix for sires is the same as in Chapter 23 except that it is multiplied by .5. The “error” variance is $\mathbf{I}(\sigma_e^2 + .75 \sigma_a^2)$. The \mathbf{G} submatrix for sires is $\mathbf{A}\sigma_a^2$ rather than $.25 \sigma_a^2 \mathbf{A}$. Then the evaluations from this model for sires are exactly twice those of Chapter 23.

A sire model containing sires of the mates but not the mates’ records can be formulated by the gametic model. Then \mathbf{a} would include both sires and grandsires. The incidence matrix for a progeny would contain elements .5 associated with sire and .25 associated with grandsire. Then the “error” variance would contain $\sigma_e^2 + .6875 \sigma_a^2$, $\sigma_e^2 + .75 \sigma_a^2$, or $\sigma_e^2 + .9375 \sigma_a^2$ for progeny with both sire and grandsire, sire only, or grandsire only respectively.

We illustrate the methods of this section with a very simple example. Animals 1, . . . , 4 have records (5,3,2,8). $\mathbf{X}' = (1 \ 2 \ 1 \ 3)$. Animals 1 and 2 are the parents of 3, and animal 1 is the parent of 4. The error variance is $\sigma_e^2 = 10$ and $\sigma_a^2 = 4$. We first treat this as an individual animal model where

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & .5 & .5 \\ & 1 & .5 & 0 \\ & & 1 & .25 \\ & & & 1 \end{pmatrix}.$$

The mixed model equations are

$$\begin{pmatrix} 1.5 & .1 & .2 & .1 & .3 \\ & .558333 & .125 & -.25 & -.166667 \\ & & .475 & -.25 & 0 \\ & & & .6 & 0 \\ & & & & .433333 \end{pmatrix} \begin{pmatrix} \hat{\beta} \\ \hat{u}_1 \\ \hat{u}_2 \\ \hat{u}_3 \\ \hat{u}_4 \end{pmatrix} = \begin{pmatrix} 3.7 \\ .5 \\ .3 \\ .2 \\ .8 \end{pmatrix}. \quad (8)$$

The solution is

$$(2.40096, .73264, -.57212, .00006, .46574).$$

Now in the gametic model the incidence matrix is

$$\begin{pmatrix} 1 & 1 & 0 \\ 2 & 0 & 1 \\ 1 & .5 & .5 \\ 3 & .5 & 0 \end{pmatrix}, \mathbf{G} = \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix}, \mathbf{R} = \text{dg}(10, 10, 12, 13).$$

$$12 = 10 + .5(4), 13 = 10 + .75(4).$$

Then the mixed model equations are

$$\begin{pmatrix} 1.275641 & .257051 & .241667 \\ & .390064 & .020833 \\ & & .370833 \end{pmatrix} \begin{pmatrix} \hat{\beta} \\ \hat{u}_1 \\ \hat{u}_2 \end{pmatrix} = \begin{pmatrix} 3.112821 \\ .891026 \\ .383333 \end{pmatrix}. \quad (9)$$

The solution is

$$(2.40096, .73264, -.57212). \quad (10)$$

This is the same as the first 3 elements of (28.8).

$$\begin{aligned} \hat{e}_3 &= 2 - 2.40096 - .5(.73264 - .57212) = -.48122. \\ \hat{u}_3 &= .5(.73264 - .57212) + 2(-.48122)/12 = .00006. \\ \hat{e}_4 &= 8 - 3(2.40096) - .5(.73264) = .43080. \\ \hat{u}_4 &= .5(.73264) + 3(.43080)/13 = .46574. \end{aligned}$$

\hat{u}_3, \hat{u}_4 are the same as in (28.8).

6.2 Repeated records model

This section is concerned with multiple records in a single trait and under the assumption that

$$\text{Var} \begin{pmatrix} y_{i1} \\ y_{i2} \\ y_{i3} \\ \vdots \end{pmatrix} = \begin{pmatrix} 1 & r & r & \cdots \\ r & 1 & r & \cdots \\ r & r & 1 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \sigma_y^2,$$

where \mathbf{y} has been adjusted for random factors other than producing ability and random error. The subscript i refers to a particular animal. The model is

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_a\mathbf{a} + \mathbf{Z}_p\mathbf{p} + \text{possibly other random factors} + \mathbf{e}.$$

$$Var \begin{pmatrix} \mathbf{a} \\ \mathbf{p} \\ \mathbf{e} \end{pmatrix} = \begin{pmatrix} \mathbf{A}\sigma_a^2 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}\sigma_p^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}\sigma_e^2 \end{pmatrix}.$$

In an unselected population $\sigma_a^2 = h^2\sigma_y^2$, $\sigma_p^2 = (r - h^2)\sigma_y^2$, $\sigma_e^2 = (1 - r)\sigma_y^2$, after adjusting \mathbf{y} for other random factors. As before b animals have progeny; $c \leq b$ of these have records. These records number n_1 . Also as before d animals with records have no progeny. The number of records made by these animals is n_2 .

First we state the model as described in Chapter 24. \mathbf{X} , \mathbf{Z}_a , \mathbf{Z}_b all have $n_1 + n_2$ rows. The number of elements in \mathbf{a} is $b + d$. The \mathbf{Z}_a matrix is the same as in the conventional model of Section 28.6.1 except that the row pertaining to an individual with records is repeated as many times as there are records on that animal. The number of elements in \mathbf{p} is $c + d$ corresponding to these animals with records. \mathbf{Z}_p would be an identity matrix with order $c + d$ if the $c + d$ animals with records had made only one record each. Then the row of this matrix corresponding to an animal is repeated as many times as the number of records in that animal. Since $\mathbf{Z}_p'\mathbf{Z}_p + (\mathbf{I}\sigma_p^2)^{-1}$ is diagonal, $\hat{\mathbf{p}}$ can be “absorbed” easily to reduce the number of equations to $b + d$ plus the number of elements in $\boldsymbol{\beta}$. The predicted real producing ability of the i^{th} animal is $\hat{a}_i + \hat{p}_i$, with $\hat{p}_i = 0$ for animals with no records.

Now we state the gametic model for repeated records. As for single records, \mathbf{a} now has b elements corresponding to the b animals with progeny. \mathbf{Z}_a is exactly the same as in the gametic model for single records except that the row pertaining to an animal is repeated as many times as the number of records for that animal. As in the conventional method for repeated records, \mathbf{p} has $c + d$ elements and \mathbf{Z}_p is the same as in that model.

Now Mendelian sampling is taken care of in this model by altering $Var(\mathbf{p})$ rather than $Var(\mathbf{e})$ as was done in the single record gametic model. For the parents $Var(\mathbf{p})$ remains diagonal with the first c diagonals being σ_p^2 . The remaining d have the following possible values.

- (1) $\sigma_p^2 + .5\sigma_a^2$ if both parents are in \mathbf{a} ,
- (2) $\sigma_p^2 + .75\sigma_a^2$ if one parent is in \mathbf{a} ,
- (3) $\sigma_p^2 + \sigma_a^2$ if no parent is in \mathbf{a} .

Again we can absorb “p” to obtain a set of equations numbering b plus the number of elements in $\boldsymbol{\beta}$, a reduction of c from the conventional equations. The computation of \hat{a} for the d animals with no progeny is simple.

$$\begin{aligned} \hat{a}_i &= .5(\text{sum of parental } \hat{\mathbf{a}}) + k_i\hat{p}_i. \\ \text{where } k_i &= .5\sigma_a^2/(\sigma_p^2 + .5\sigma_a^2) \text{ for animals with 2 parents in } \mathbf{a}. \\ &= .75\sigma_a^2/(\sigma_p^2 + .75\sigma_a^2) \text{ for those with one parent in } \mathbf{a}. \\ &= \sigma_a^2/(\sigma_p^2 + \sigma_a^2) \text{ for those with no parent in } \mathbf{a}. \end{aligned}$$

These two methods for repeated records are illustrated with the same animals as in Section 28.8 except now there are repeated records. The 4 animals have 2,3,1,2 records respectively. These are (5,3,4,2,3,6,7,8). $\mathbf{X}' = (1\ 2\ 3\ 1\ 2\ 2\ 3\ 2)$. Let

$$\begin{aligned}\sigma_a^2 &= .25, \\ \sigma_p^2 &= .20, \\ \sigma_e^2 &= .55.\end{aligned}$$

Then the regular mixed model equations are in (28.11).

$$\begin{pmatrix} 65.455 & 5.455 & 10.909 & 3.636 & 9.091 & 5.455 & 10.909 & 3.636 & 9.091 \\ & 10.970 & 2.0 & -4.0 & -2.667 & 3.636 & 0 & 0 & 0 \\ & & 11.455 & -4.0 & 0 & 0 & 5.455 & 0 & 0 \\ & & & 9.818 & 0 & 0 & 0 & 1.818 & 0 \\ & & & & 8.970 & 0 & 0 & 0 & 3.636 \\ & & & & & 8.636 & 0 & 0 & 0 \\ & & & & & & 10.455 & 0 & 0 \\ & & & & & & & 6.818 & 0 \\ & & & & & & & & 8.636 \end{pmatrix}$$

$$\begin{pmatrix} \hat{\beta} \\ \hat{\mathbf{a}} \\ \hat{\mathbf{p}} \end{pmatrix} = \begin{pmatrix} 145.455 \\ 14.546 \\ 16.364 \\ 10.909 \\ 27.273 \\ 14.546 \\ 16.364 \\ 10.909 \\ 27.273 \end{pmatrix}. \quad (11)$$

Note that the right hand sides for $\hat{\mathbf{a}}$ and $\hat{\mathbf{p}}$ are identical. The solution is

$$(1.9467, .8158, -.1972, .5660, 1.0377, .1113, -.3632, .4108, .6718). \quad (12)$$

Next the solution for the gametic model is illustrated with the same data. The incidence matrix is

$$\begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 1 & 0 & 0 & 0 \\ 3 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 2 & 0 & 1 & 0 & 1 & 0 & 0 \\ 2 & .5 & .5 & 0 & 0 & 1 & 0 \\ 3 & .5 & 0 & 0 & 0 & 0 & 1 \\ 2 & .5 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

corresponding to $\beta, a_1, a_2, p_1, p_2, p_3, p_4$.

$$\begin{aligned} \text{Var}(\mathbf{e}) &= .55 \mathbf{I}, \\ \text{Var}(\mathbf{a}) &= \begin{pmatrix} .25 & 0 \\ 0 & .25 \end{pmatrix}, \\ \text{Var}(\mathbf{p}) &= \text{diag} (.2, .2, .325, .3875). \end{aligned}$$

Then the mixed model equations are

$$\begin{pmatrix} 65.454 & 11.818 & 12.727 & 5.454 & 10.909 & 3.636 & 9.091 \\ & 9.0 & .454 & 3.636 & 0 & .909 & 1.818 \\ & & 9.909 & 0 & 5.454 & .909 & 0 \\ & & & 8.636 & 0 & 0 & 0 \\ & & & & 10.454 & 0 & 0 \\ & & & & & 4.895 & 0 \\ & & & & & & 6.217 \end{pmatrix} \begin{pmatrix} \hat{\beta} \\ a_1 \\ a_2 \\ p_1 \\ p_2 \\ p_3 \\ p_4 \end{pmatrix} = \begin{pmatrix} 145.454 \\ 33.636 \\ 21.818 \\ 14.546 \\ 16.364 \\ 10.909 \\ 27.273 \end{pmatrix} \quad (13)$$

Note that the coefficient submatrix for $\hat{\mathbf{p}}$ is diagonal. The solution is

$$(1.9467, .8158, -.1972, .1113, -.3632, .6676, 1.3017). \quad (14)$$

Note that $\hat{\beta}, \hat{a}_1, \hat{a}_2, \hat{p}_1, \hat{p}_2$ are the same as in (28.12). Now

$$\begin{aligned} \hat{a}_3 &= .5(.8158 - .1972) + .125 (.6676)/.325 = .5660. \\ \hat{a}_4 &= .5(.8158) + .1875(1.3017)/.3875 = 1.0377. \end{aligned}$$

These are the same results for \hat{a}_3 and \hat{a}_4 as (28.12).