

Pedigree options in ASReml

Arthur Gilmour^{a*} Greg Dutkowski^b

^a*Mathematics and Statistics, University of Wollongong, NSW, Australia*

^b*University of Tasmania, Hobart, Tasmania, Australia*

Abstract

ASReml has several options when forming the pedigree. This paper discusses the new developments for Release 2. Briefly, these are constraints on the Genetic groups, changes to the way the Maternal Grandsire model is fitted and provision for a degree of selfing and inbreeding.

Key words: BLUP, linear mixed models, REML, variance components, genetic groups, average information, ASReml, Maternal Grandsire Model, Selfing

1 Introduction

An important development in the use of mixed models was the incorporation of the Numerator Relationship Matrix to account for genetic correlations among animals. The wonder of this approach is that the inverse relationship matrix is in general sparser than the relationship matrix itself and not difficult to form in a recursive manner when parents are listed before their offspring.

However, a few special situations arise in practise which we consider here. First, in some breeding situations, sires and sire groups are tracked but individual dam pedigrees are not tracked. This leads to the Maternal Grandsire model. ASReml Release 1 accomodated this model by simply inserting a dummy dam into the pedigree. We now form the inverse omitting this dummy dam. A further modification is for sex-linked traits when it is just inheritance through the X chromosome that is considered.

* Corresponding author.

Email address: `arthur.gilmour@cargovale.com.au` (Arthur Gilmour).

In open pollinated forest trees, a proportion of the open-pollinated flowers are self pollinated. We derive the relationship matrix for a given proportion of selfing when the second parent is unspecified.

In other plant species, for example cereals, the plants are mainly self pollinated and lines are selfed for several generations to produce varieties for release. Alternatively, the pedigree can have varying levels of selfing. The pedigree relationship needs to take this into account.

ASReml now has 3 Ainverse routines. These extensions are implemented as for the method of Meuwissen and Lou (1992) which should be slightly faster than the other methods for large pedigrees.

2 Standard relationship matrix

The standard procedures can be summarised as follows. Let $\mathbf{A} = \{a_{ij}\}$ be the relationship matrix. Let $\mathbf{a}_{i,-j}$ be the i th row of \mathbf{A} except for the j th element.

- (1) Assume the relationship matrix for the base animals is known for example, are unrelated and not inbred so that their relationship matrix is an Identity matrix.
- (2) The row of the relationship matrix for the progeny (i) of two parents (s and d) is generated as the average of the relationship matrix rows for the parents.

$$\mathbf{a}_{i,-i} = \frac{1}{2}(\mathbf{a}_{s,-i} + \mathbf{a}_{d,-i})$$

- (3) The diagonal element of this new animal is $a_{i,i} = 1 + \frac{1}{2}a_{s,d} = 1 + f_i$ where f_i is the inbreeding coefficient.

Applying this rule for two fullsibs and a half-sib generates the following matrix.

$$\begin{bmatrix} \text{Sire} & 1. & 0. & 0. & 0.5 & 0.5 & 0.5 \\ \text{Dam1} & 0. & 1. & 0. & 0.5 & 0.5 & 0. \\ \text{Dam2} & 0. & 0. & 1. & 0. & 0. & 0.5 \\ \text{SibS1} & 0.5 & 0.5 & 0. & 1. & 0.5 & 0.25 \\ \text{SibS1} & 0.5 & 0.5 & 0. & 0.5 & 1.0 & 0.25 \\ \text{SibS2} & 0.5 & 0. & 0.5 & 0.25 & 0.25 & 1.0 \end{bmatrix}$$

Progressively inverting this matrix gives the sequence:

$$\begin{bmatrix} \text{Sire} & 1 & 0 & 0 \\ \text{Dam1} & 0 & 1 & 0 \\ \text{Dam2} & 0 & 0 & 1 \end{bmatrix} \Rightarrow \begin{bmatrix} \text{Sire} & 1.5 & 0.5 & 0 & -1 \\ \text{Dam1} & 0.5 & 1.5 & 0 & -1 \\ \text{Dam2} & 0 & 0 & 1 & 0 \\ \text{SibS1} & -1 & -1 & 0 & 2 \end{bmatrix} \Rightarrow \begin{bmatrix} \text{Sire} & 2 & 1 & 0 & -1 & -1 \\ \text{Dam1} & 1 & 2 & 0 & -1 & -1 \\ \text{Dam2} & 0 & 0 & 1 & 0 & 0 \\ \text{SibS1} & -1 & -1 & 0 & 2 & 0 \\ \text{SibS1} & -1 & -1 & 0 & 0 & 2 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} \text{Sire} & 2.5 & 1 & 0.5 & -1 & -1 & -1 \\ \text{Dam1} & 1 & 2 & 0 & -1 & -1 & 0 \\ \text{Dam2} & 0.5 & 0 & 1.5 & 0 & 0 & -1 \\ \text{SibS1} & -1 & -1 & 0 & 2 & 0 & 0 \\ \text{SibS1} & -1 & -1 & 0 & 0 & 2 & 0 \\ \text{SibS2} & -1 & 0 & -1 & 0 & 0 & 2 \end{bmatrix}$$

The rule for progressively generating these inverses is derived as follows.

Assume we have \mathbf{A}_1 and \mathbf{A}_1^{-1} for a set of animals including the parents of a new animal. Let \mathbf{p} be a vector which performs the operation of averaging the parental rows of \mathbf{A}_1 . It will be all zero except that the positions corresponding to the parents will be 0.5.

Then $\mathbf{A}_2 = \begin{pmatrix} \mathbf{A}_1 & \mathbf{A}_1\mathbf{p} \\ \mathbf{p}'\mathbf{A}_1 & 1 + \frac{1}{2}a_{s,d} \end{pmatrix}$ is the expanded relationship matrix.

The standard expression for the inverse of a partitioned matrix is

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}' & \mathbf{C} \end{pmatrix}^{-1} = \begin{pmatrix} \mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{B}\mathbf{Q}\mathbf{B}'\mathbf{A}^{-1} & \mathbf{A}^{-1}\mathbf{B}\mathbf{Q} \\ \mathbf{Q}\mathbf{B}'\mathbf{A}^{-1} & \mathbf{Q} \end{pmatrix} \text{ where}$$

$\mathbf{Q} = (\mathbf{C} - \mathbf{B}'\mathbf{A}^{-1}\mathbf{B})^{-1}$. In this case, $\mathbf{B} = \mathbf{A}_1\mathbf{p}$ so that $\mathbf{A}^{-1}\mathbf{B} = \mathbf{p}$.

Thus $\mathbf{A}_2^{-1} = \begin{pmatrix} \mathbf{A}_1^{-1} + q\mathbf{p}\mathbf{p}' & -q\mathbf{p} \\ -q\mathbf{p}' & q \end{pmatrix}$ where $q = (1 + \frac{1}{2}a_{s,d} - \mathbf{p}'\mathbf{A}_1\mathbf{p})^{-1}$.

Since \mathbf{p} has such a simple structure, $q = 1/(1 - \frac{1}{4}(a_{s,s} + a_{d,d})) = 4/(2 - (f_s + f_d))$ which simply requires the inbreeding coefficients for each parent.

To test **ASReml** we obtain \mathbf{A}^{-1} , invert it, and check the result. We will use the following pedigree which includes inbreeding and selfing.

```

1 0 0
2 0 0
3 0 0
4 1 1
5 1 1
6 2 2
7 4 6
8 5 6
9 7 8
10 9 9

```

The following was produced with **ASReml** qualifiers **!METHOD 0 !GIV !DIAG**. First is the relationship matrix obtained by inverting the Inverse relationship matrix. Rounding errors occur because of the moderate precision of the Ainverse .giv file. Next is the Ainverse as returned in the .giv file and finally the inbreeding coefficients as returned in the ainverse.dia file.

```

1 1.0
2 0.0 1.0
3 0.0 0.0 1.0
4 1.0 0.0 0.0 1.500
5 1.0 0.0 0.0 1.000 1.500
6 0.0 1.0 0.0 0.000 0.000 1.500
7 0.5 0.5 0.0 0.750 0.500 0.750 1.0000
8 0.5 0.5 0.0 0.500 0.750 0.750 0.6250 1.0000
9 0.5 0.5 0.0 0.625 0.625 0.750 0.8125 0.8125 1.3125
10 0.5 0.5 0.0 0.625 0.625 0.750 0.8125 0.8125 1.3125 1.65625

```

```

1 5.0
2 .0 3.0
3 .0 .0 1.0
4 -2.0 .0 .0 3.0
5 -2.0 .0 .0 .0 3.0
6 .0 -2.0 .0 1.0 1.0 4.0
7 .0 .0 .0 -2.0 .0 -2.0 4.5
8 .0 .0 .0 .0 -2.0 -2.0 0.5 4.5
9 .0 .0 .0 .0 .0 .0 -1.0 -1.0 4.90909
10 .0 .0 .0 .0 .0 .0 .0 .0 -2.90909 2.90909

```

Identity	Inbreeding	DiagofAinverse
1	0.0000	5.0000
2	0.0000	3.0000
3	0.0000	1.0000
4	0.50000	3.0000
5	0.50000	3.0000
6	0.50000	4.0000
7	0.0000	4.5000

8	0.0000	4.5000
9	0.31250	4.9091
10	0.65625	2.9091

Wed 08/09/2004

3 Maternal Grandsire Model

In this situation, dam pedigree is incomplete. The dam of the dam is regarded as a base animal but the sire of the dam is known. Consider the relationship matrix for a Maternal Grandsire, Sire, Dam and Progeny given by

$$\begin{bmatrix} \text{Sire of Dam} & 1 + f_g & r & 0 & (1 + f_g)/2 & (1 + f_g + 2r)/4 \\ \text{Sire} & r & 1 + f_s & 0 & r/2 & (2 + 2f_s + r)/4 \\ \text{Dam of Dam} & 0 & 0 & 1 & 0.5 & 0.25 \\ \text{Dam} & (1 + f_g)/2 & r/2 & 0.5 & 1 & (1 + r/2)/2 \\ \text{Progeny} & \frac{(1+f_g+2r)}{4} & \frac{(2+2f_s+r)}{4} & \frac{1}{4} & \frac{(1+r/2)}{2} & 1 + \frac{r}{4} \end{bmatrix}$$

We see that the Progeny line is half the Sire line plus a quarter of the Maternal Grandsire line except that the diagonal element is 1 plus a quarter of the relationship between the sire and the maternal grandsire.

The algebra is as before except for a different definition of \mathbf{p} and q .

For the non-zero elements of \mathbf{p}_s being $\frac{1}{2}$ and $\frac{1}{4}$ for sire and maternal grandsire

respectively, $\mathbf{A}_m = \begin{pmatrix} \mathbf{A}_1 & \mathbf{A}_1 \mathbf{p}_s \\ \mathbf{p}'_s \mathbf{A}_1 & 1 + \frac{r}{4} \end{pmatrix}$ and $\mathbf{A}_m^{-1} = \begin{pmatrix} \mathbf{A}_1^{-1} + \mathbf{p}_s q_s \mathbf{p}'_s & -\mathbf{p}_s q_s \\ -\mathbf{p}'_s q_s & q_s \end{pmatrix}$

where

$$q_s = (1 + r/4 - \mathbf{p}'_s \mathbf{A}_s \mathbf{p}_s)^{-1} = (1 - (1 + f_g + 4(1 + f_s))/16)^{-1} = \frac{16}{11 - (f_g + 4f_s)}.$$

The following results are from ASReml using the same pedigree file as before but with qualifiers !METHOD 0 !MGS !GIV !DIAG where the !MGS qualifier causes the third field to be interpreted as the maternal grandsire.

1	1.0000					
2	0.0000	1.0000				
3	0.0000	0.0000	1.			
4	0.7500	0.0000	0.	1.2500		
5	0.7500	0.0000	0.	0.5625	1.2500	
6	0.0000	0.7500	0.	0.0000	0.0000	1.2500
7	0.3750	0.1875	0.	0.6250	0.2812	0.3125
						1.0000

```

8 0.3750 0.1875 0. 0.2813 0.6250 0.3125 0.2188 1.0000
9 0.2812 0.1406 0. 0.3828 0.2969 0.2344 0.5547 0.3594 1.0547
10 0.2109 0.1055 0. 0.2871 0.2227 0.1758 0.4160 0.2695 0.7910 1.2637

```

```

1 1 2.63636
2 2 1.81818
3 3 1.00000
4 1 -1.09091 4 1.86480
5 1 -1.09091 5 1.86480
6 2 -1.09091 4 0.205128 5 0.205128 6 1.65967
7 4 -0.820513 6 -0.410256 7 2.00466
8 5 -0.820513 6 -0.410256 7 0.181818 8 1.73193
9 7 -0.727273 8 -0.363636 9 2.29358
10 9 -1.11872 10 1.49162

```

Identity	Inbreeding	DiagofAinverse
1	0.0000	2.6364
2	0.0000	1.8182
3	0.0000	1.0000
4	0.25000	1.8648
5	0.25000	1.8648
6	0.25000	1.6597
7	0.0000	2.0047
8	0.0000	1.7319
9	0.54688E-01	2.2936
10	0.26367	1.4916

4 Honey Bee breeding Model

Bienefeld et al (2007) describes a breeding program for Honey Bees which leads to another variation on the A matrix rules. The male bee is haploid. There is a mother queen mated to produce 'drone bearing' queens that produce the drones for mating. Each mother queen has several daughters and these have several sons. The scheme is quantified by two parameters q , the number of drone bearing queens, and d the average number of drones a queen mates with. So a queen is produced by combining the genetic material in a drone with that of the queen. But the drone is derived from the mother queen.

This scenario is similar to the Maternal Grand sire Model except the vector \mathbf{p} will have coefficients $\frac{1}{2}$ for the mating queen and p for the mother queen (which produced the mating drone).

$$?? \mathbf{A}_m = \begin{pmatrix} \mathbf{A}_1 & \mathbf{A}_1 \mathbf{p}_s \\ \mathbf{p}'_s \mathbf{A}_1 & 1 + \frac{r}{4} \end{pmatrix} \text{ and } \mathbf{A}_m^{-1} = \begin{pmatrix} \mathbf{A}_1^{-1} + \mathbf{p}_s q_s \mathbf{p}'_s & -\mathbf{p}_s q_s \\ -\mathbf{p}_s q_s & q_s \end{pmatrix} \text{ where}$$

$$q_s = (1 + r/4 - \mathbf{p}'_s \mathbf{A}_s \mathbf{p}_s)^{-1} = (1 - (1 + f_g + 4(1 + f_s))/16)^{-1} = \frac{16}{11 - (f_g + 4f_s)}.$$

The following results are from ASReml using the same pedigree file as before but with qualifiers !METHOD 0 !BEE 0.367 !GIV !DIAG where the !BEE 0.367 qualifier causes the third field to be interpreted as the mother queen with path coefficient $p = 0.367$.

```

1.0000                A matrix
0.0000 1.0000
0.0000 0.0000    1
0.8670 0.0000    0 1.3670
0.8670 0.0000    0 0.7517 1.3670
0.0000 0.8670    0 0.0000 0.0000 1.3670
0.4335 0.3182    0 0.6835 0.3758 0.5017 0.9488
0.4335 0.3182    0 0.3758 0.6835 0.5017 0.3720 0.9488
0.3758 0.2759    0 0.4797 0.4388 0.4350 0.6110 0.5342 1.0854
0.3259 0.2392    0 0.4159 0.3804 0.3771 0.5297 0.4632 0.9410 1.3472
0.1629 0.1196    0 0.2079 0.1902 0.1886 0.2649 0.2316 0.4705 0.6736 1.3368
0.1196 0.0878    0 0.1526 0.1396 0.1384 0.1944 0.1700 0.3454 0.4944 0.2472 1.1814

Inbreeding
0.0 0.0 0.0 0.3670 0.3670 0.3670 -0.0512 -0.0512 0.0854 0.3472 0.3368 0.1814

A-inverse
3.4433
0.0000 2.2216
0.0000 0.0000    1
-1.4090 0.0000    0 2.2163
-1.4090 0.0000    0 0.0000 2.2163
0.0000 -1.4090    0 0.4338 0.4338 2.2621
0.0000 0.0000    0 -1.1821 0.0000 -0.8677 2.7925
0.0000 0.0000    0 0.0000 -1.1821 -0.8677 0.3143 2.5950
0.0000 0.0000    0 0.0000 0.0000 0.0000 -0.8564 -0.6286 3.1276
0.0000 0.0000    0 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 -1.6318 2.2668
0.0000 0.0000    0 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 -0.5000 1.0
0.0000 0.0000    0 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 -0.3670 0.0 1.000

Identity Inbreeding  Diag_of_Ainv Offspring-Parent  Pedigree
1         0.0000      3.4433      Parent                1 0 0
2         0.0000      2.2216      Parent                2 0 0
3         0.0000      1.0000      Offspring             3 0 0
4         0.50000     2.2163      Parent                4 1 1
5         0.50000     2.2163      Parent                5 1 1
6         0.50000     2.2621      Parent                6 2 2
7         0.81836E-01 2.7925      Parent                7 4 6
8         0.81836E-01 2.5950      Parent                8 5 6
9         0.21838     3.1276      Parent                9 7 8
10        0.48017     2.2668      Parent               10 9 9
11        0.0000      1.0000      Offspring             11 10 0
12        0.0000      1.0000      Offspring             12 0 10

```

The above numbers are not correct because Inbreeding cannot be negative!

5 Genetic groups

The ASReml qualifier !GROUPS g indicates that the first g identifiers in the pedigree file relate to genetic groups rather than to individuals in the popula-

tion. When genetic groups are present, the SIRE and DAM fields should both be zero for the group lines. All other lines must specify one of the genetic groups as SIRE or DAM if the actual parent is unknown.

In release 1.62, the option was added to apply 'sum to zero' constraints on group effects. If the pedigree file includes genetic groups without constraint, the 'constant term' will be included in the genetic group effects. The constant term is unlikely to be the population average and may generate singularities in unpredictable locations because of the ordering of the 'sparse' equations. This may be undesirable. The !LAST qualifier has been added to force the 'group' equations to be absorbed last so that any singularities in them will appear where expected. However, singularities in the pedigree factor also mess up the REML likelihood evaluation because they mess up the degrees of freedom count.

'Sum to Zero' constraints are applied to sets of genetic group effects by following the set with a dummy genetic group (one with no animals in it). ASReml then modifies the A inverse to apply a constraint in place of this dummy genetic group. That is, when there are two or more groups containing animals followed by a group with no animals, the A inverse line for the empty group is replaced with a line containing 1's for the preceding set of groups, and zero otherwise. This Lagrangian constrains the genetic group effects to sum to zero.

The following results from ASReml !METHOD 0 !GRP 3 !GIV !DIAG. We use the same pedigree file as before but now the first three lines are genetic groups. Notice that there are no individuals assigned to the third group. When ASReml has formed the A-inverse, it notes that this 'group' is empty and inserts Lagrangian off diagonal elements for this equation.

For the purpose of checking the inverse, the A matrix was calculated after zeroing the intersection between groups and individuals.

```
%ex/ex11/amg00.txt
\input /data/ex/ex11/amg00.txt
 1 0.33333
 2-0.33333 0.33333
 3 0.33333 0.66667-0.66667
 4 0.00000 0.00000 0.00000 1.00
 5 0.00000 0.00000 0.00000 0.00 1.00
 6 0.00000 0.00000 0.00000 0.00 0.00 1.0
 7 0.00000 0.00000 0.00000 0.50 0.00 0.5 1.000
 8 0.00000 0.00000 0.00000 0.00 0.50 0.5 0.250 1.000
 9 0.00000 0.00000 0.00000 0.25 0.25 0.5 0.625 0.625 1.125
10 0.00000 0.00000 0.00000 0.25 0.25 0.5 0.625 0.625 1.125 1.5625

      1      1      2.00000
```


2	2	1.00000						
3	1	1.00000	2	1.00000	3	0.00000		
4	1	-1.00000	4	1.50000				
5	1	-1.00000	5	1.50000				
6	2	-1.00000	4	0.500000	5	0.500000	6	2.00000
7	4	-1.00000	6	-1.00000	7	2.50000		
8	5	-1.00000	6	-1.00000	7	0.500000	8	2.50000
9	7	-1.00000	8	-1.00000	9	4.28571		
10	9	-2.28571	10	2.28571				
Identity		Inbreeding		DiagofAinverse				
	1	-1.0000		2.0000				
	2	-1.0000		1.0000				
	3	-1.0000		0.0000				
	4	0.0000		1.5000				
	5	0.0000		1.5000				
	6	0.0000		2.0000				
	7	0.0000		2.5000				
	8	0.0000		2.5000				
	9	0.12500		4.2857				
	10	0.56250		2.2857				

Wed 08/09/2004

6 Partial selfing

$$\left[\begin{array}{l} \text{Female Parent} \\ \text{Male Parent} \\ \text{Selfed} \\ \text{Crossed} \end{array} \begin{array}{cccc} 1 + f_f & r & 1 + f_f & (1 + f_f + r)/2 \\ r & 1 + f_m & r & (1 + f_m + r)/2 \\ 1 + f_f & r & 1 + \frac{1+f_f}{2} & (1 + f_f + r)/2 \\ \frac{1+f_f+r}{2} & \frac{1+f_m+r}{2} & \frac{1+f_f+r}{2} & 1 + r/2 \end{array} \right]$$

Let s be the proportion selfed. The contribution of the female parent is then $(s + \frac{1}{2}(1-s))\mathbf{a}_f = \frac{1+s}{2}\mathbf{a}_f$. The contribution from the male parent is $\frac{1-s}{2}\mathbf{a}_m$. The diagonal element will be $s(1 + \frac{1+f_f}{2}) + (1-s)(1 + r/2) = 1 + \frac{s(1+f_f)+(1-s)r}{2}$.

Thus, the non zero elements of \mathbf{p}_w are $\left(\frac{1+s}{2} \frac{1-s}{2} \right)$ and

$$\begin{aligned} q_w^{-1} &= 1 + \frac{s(1+f_f)+(1-s)r}{2} - \left(\frac{1+s}{2} \frac{1-s}{2} \right) \begin{pmatrix} 1 + f_f & r \\ r & 1 + f_m \end{pmatrix} \begin{pmatrix} \frac{1+s}{2} \\ \frac{1-s}{2} \end{pmatrix} \\ &= 1 - \frac{1}{4}((1 + s^2)(1 + f_f) + 2sr(1 - s) + (1 - s)^2(1 + f_m)) \end{aligned}$$

Now this expression for q_w involves the covariance between the parents which previously cancelled out and which is not easily available from the existing algorithm. However, we expect that if the plant is outcrossed, the male parent

is not known. We could therefore nominate a degree of relationship for seeds collected in the wild, or just assume the male parent is unrelated.

As a starting point, ASReml assumes that if the male parent is unknown, there is proportional selfing but if the male parent is identified, that a controlled mating has occurred. This complicates the algorithm because in the former case, the nonzero element of \mathbf{p} is $\frac{1}{2}(1 + s)$ while in the other case, there are two values of $\frac{1}{2}$

The following results from ASReml !METHOD 0 !SELF 0.3 !GIV !DIAG for the pedigree displayed at the bottom.

```

1 1.0000
2 0.0000 1.0000
3 0.0000 0.0000 1.
4 1.0000 0.0000 0. 1.5000
5 0.6500 0.0000 0. 0.9750 1.2250
6 0.0000 0.6500 0. 0.0000 0.0000 1.1500
7 0.5000 0.3250 0. 0.7500 0.4875 0.5750 1.0000
8 0.3250 0.3250 0. 0.4875 0.6125 0.5750 0.5312 1.0000
9 0.4125 0.3250 0. 0.6188 0.5500 0.5750 0.7656 0.7656 1.2656
10 0.2681 0.2112 0. 0.4022 0.3575 0.3737 0.4977 0.4977 0.8227 1.1898

1 1 3.00000
2 2 1.58076
3 3 1.00000 1 -2.00000
4 4 3.45533
5 4 -1.09937 5 2.30672
6 2 -0.893471 4 0.740741 5 0.615385 6 2.73070
7 4 -1.48148 6 -1.48148 7 3.46296
8 5 -1.23077 6 -1.23077 7 0.500000 8 2.96154
9 7 -1.00000 8 -1.00000 9 2.64492
10 9 -0.992189 10 1.52644
Identity Female Male Inbreeding DiagofAinverse
1 0 0 0.0000 3.0000
2 0 0 0.0000 1.5808
3 0 0 0.0000 1.0000
4 1 1 0.50000 3.4553
5 4 0 0.22500 2.3067
6 2 0 0.15000 2.7307
7 4 6 0.0000 3.4630
8 5 6 0.0000 2.9615
9 7 8 0.26562 2.6449
10 9 0 0.18984 1.5264
Fri 17/09/2004

```

7 Inbred lines

In this situation, after a cross is made, it is selfed a large number of times to produce an inbred line. The procedure is basically as before except that the inbreeding is always 1.

For the non-zero elements of \mathbf{p}_i being $\frac{1}{2}$ and $\frac{1}{2}$, $\mathbf{A}_i = \begin{pmatrix} \mathbf{A}_1 & \mathbf{A}_1\mathbf{p}_i \\ \mathbf{p}'_i\mathbf{A}_1 & 2 \end{pmatrix}$ and

$\mathbf{A}_i^{-1} = \begin{pmatrix} \mathbf{A}_1^{-1} + \mathbf{p}_i q_i \mathbf{p}'_i & -\mathbf{p}_i q_i \\ -\mathbf{p}_i q_i & q_i \end{pmatrix}$ where $q_i^{-1} = 2 - \mathbf{p}'_i \mathbf{A}_1 \mathbf{p}_i = 1 - \frac{1}{2} a_{mf}$. This

last term comes about from $2 - \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 2 & a_{mf} \\ a_{mf} & 2 \end{pmatrix} \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$

The special cases are that where no parents are known, $\mathbf{p}_i = (0)$ so that $q_i^{-1} = 2$ and if only one parent is known, \mathbf{p}_i contains a single $\frac{1}{2}$ so that $q_i^{-1} = 2 - \frac{1}{2} 2 \frac{1}{2} = 1.5$

The following results are from ASReml using the pedigree file displayed below with qualifiers !METHOD 0 !INBRED 1.0 !GIV !DIAG where the !INBRED qualifier implies inbred lines. Note that 'selfing' is not permitted in the pedigree with inbred lines as it will generate a singularity.

```

1 2.0000
2 0.0000 2.0000
3 0.0000 0.0000 2.00000
4 1.0000 0.0000 1.0000 2.0000
5 0.0000 1.0000 1.0000 0.5000 2.0000
6 0.5000 0.5000 1.0000 1.2500 1.2500 2.0000
7 0.7500 0.2500 1.0000 1.6250 0.8750 1.6250 2.0000
8 1.3750 0.1250 0.5000 1.3125 0.4375 1.0625 1.3750 2.0000
9 1.0625 0.1875 0.7500 1.4687 0.6562 1.3437 1.6875 1.6875 2.0000
10 0.7812 0.3437 0.8750 1.3594 0.9531 1.6719 1.6562 1.3750 1.6719 2.0

```

```

1      1      1.15000
2      2      0.750000
3      1 0.250000      2 0.250000      3      1.00000
4      1 -0.500000      3 -0.500000      4      2.00000
5      2 -0.500000      3 -0.500000      4 0.333333      5      1.33333
6      4      0.00000      5 -0.666667      6      2.76190
7      1 0.400000      4 -1.33333      6 -1.33333      7      3.86667
8      1 -0.800000      7      0.00000      8      2.40000
9      6 0.761905      7 -1.60000      8 -1.60000      9      3.96190
10     6 -1.52381      9 -1.52381      0      3.04762

```

Identity	Female	Male	Inbreeding	DiagofAinverse
1	0	0	1.0000	1.1500
2	0	0	1.0000	0.75000
3	0	0	1.0000	1.0000
4	1	3	1.0000	2.0000
5	2	3	1.0000	1.3333
6	4	5	1.0000	2.7619
7	6	4	1.0000	3.8667
8	7	1	1.0000	2.4000
9	8	7	1.0000	3.9619
10	9	6	1.0000	3.0476

Thu 16/09/2004

For example, the coefficient $3.0476 = 1/(1 - 43/64)$ where $43/64 = \frac{1}{2}a_{96}$. The calculation of this relationship required development of a recursive routine which works as follows. Define three vectors \mathbf{H} to hold the index of the higher numbered parent, \mathbf{L} to hold the index of the other parent, and \mathbf{c} to hold the contributions to relationship. Initialise $a_{69} = 0$, $k = 1$, $H_1 = 9$, and $L_1 = 6$.

While $k > 0$,

if H_k equals L_k , add c_k to a and decrement k .

else identify sire and dam of H_k (s and d) and replace (H_k, L_k, c_k) with $(\max(s, L_k), \min(s, L_k, \frac{1}{2}c_k)$ and $(H_{k+1}, L_{k+1}, c_{k+1})$ with $(\max(d, L_k), \min(d, L_k, \frac{1}{2}c_k)$ (but omit operations if s and/or d are 0, adjusting k accordingly).

H,L,c	k=1	k=2	k=3	k=4	k=5	k=6	a
	9,6,1						
	8,6,1/2	7,6,1/2					
	8,6,1/2	6,4,1/4	6,6,1/4				
	8,6,1/2	5,4,1/8	4,4,1/8				1/4
	8,6,1/2	4,3,1/16	4,2,1/16				3/8
	8,6,1/2	4,3,1/16	3,2,1/32	2,1,1/32			3/8
	8,6,1/2	3,3,1/32	3,1,1/32				3/8
	7,6,1/4	6,1,1/4					13/32
	7,6,1/4	4,1,1/8	5,1,1/8				13/32
	7,6,1/4	4,1,1/8	2,1,1/16	3,1,1/16			13/32
	7,6,1/4	1,1,1/16	3,1,1/16				13/32
	6,6,1/8	6,4,1/8					15/32
	6,6,1/8	4,4,1/16	5,4,1/16				15/32
	6,6,1/8	4,4,1/16	4,2,1/32	4,3,1/32			15/32
	6,6,1/8	4,4,1/16	4,2,1/32	3,3,1/64	3,1,1/64		15/32
	6,6,1/8	4,4,1/16	1,2,1/64	3,2,1/64			31/64
	6,6,1/8						35/64
							43/64

While the ASReml syntax allows for an inbreeding coefficient less than one to be specified, the algebra has not been worked through for that case.

8 Mixed Model equations

If \mathbf{y} denotes the $n \times 1$ vector of observations, the linear mixed model can be written as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\tau} + \mathbf{Z}\mathbf{u} + \mathbf{e} \quad (1)$$

where $\boldsymbol{\tau}$ is the $p \times 1$ vector of fixed effects, \mathbf{X} is an $n \times p$ design matrix of full column rank which associates observations with the appropriate combination of fixed effects, \mathbf{u} is the $q \times 1$ vector of random effects, \mathbf{Z} is the $n \times q$ design matrix which associates observations with the appropriate combination of random effects, and \mathbf{e} is the $n \times 1$ vector of residual errors.

The model (1) is called a linear mixed model or linear mixed effects model. It is assumed

$$\begin{bmatrix} \mathbf{u} \\ \mathbf{e} \end{bmatrix} \sim N \left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \mathbf{G}(\boldsymbol{\gamma}) & \mathbf{0} \\ \mathbf{0} & \mathbf{R}(\boldsymbol{\phi}) \end{bmatrix} \right) \quad (2)$$

where the matrices \mathbf{G} and \mathbf{R} are functions of parameters $\boldsymbol{\gamma}$ and $\boldsymbol{\phi}$, respectively.

Details of the AI algorithm for REML estimation can be found in several places including Gilmour et al. (1995). It revolves around the mixed model equations derived from the objective function

$$\log f_{\mathbf{Y}}(\mathbf{y} | \mathbf{u}; \boldsymbol{\tau}, \mathbf{R}) + \log f_{\mathbf{U}}(\mathbf{u}; \mathbf{G}) .$$

which is the log-joint distribution of (\mathbf{Y}, \mathbf{u}) . Differentiating with respect to $\boldsymbol{\tau}$ and \mathbf{u} leads to the mixed model equations (Robinson, 1991) which are given by

$$\begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\tau}} \\ \hat{\mathbf{u}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \end{bmatrix} . \quad (3)$$

These can be written as

$$\mathbf{C}\tilde{\boldsymbol{\beta}} = \mathbf{W}'\mathbf{R}^{-1}\mathbf{y}$$

where $\mathbf{C} = \mathbf{W}'\mathbf{R}^{-1}\mathbf{W} + \mathbf{G}^*$, $\boldsymbol{\beta} = [\boldsymbol{\tau}' \ \mathbf{u}']'$ and $\mathbf{G}^* = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}^{-1} \end{bmatrix}$.

The solution of (3) requires values for the variance parameters $\boldsymbol{\gamma}$ and $\boldsymbol{\phi}$. In practice we replace $\boldsymbol{\gamma}$ and $\boldsymbol{\phi}$ by their REML estimates $\hat{\boldsymbol{\gamma}}$ and $\hat{\boldsymbol{\phi}}$.

The AI algorithm uses the trace($(\partial\mathbf{C}) \mathbf{C}^{-1}$) when calculating the score and uses $\mathbf{Y}'\mathbf{P}\mathbf{Y}$ as the the Average Information matrix where $\mathbf{y}'\mathbf{P}\mathbf{y}$ is the usual residual sum of squares and \mathbf{Y} is a matrix of working variables, one for each variance parameter, given by either $(\partial\mathbf{R}) \mathbf{R}^{-1}\mathbf{e}$ for parameters of \mathbf{R} or $\mathbf{Z}(\partial\mathbf{G}) \mathbf{G}^{-1}\mathbf{u}$ for parameters of \mathbf{G} .

$\mathbf{Y}'\mathbf{P}\mathbf{Y}$ is formed by an absorption process (Gilmour et al 1995) which is quite efficient when \mathbf{C} is sparse provided a judicious ordering is used. The other part of the process requires calculation of trace($(\partial\mathbf{C}) \mathbf{C}^{-1}$). We only need the elements of \mathbf{C}^{-1} which correspond to non-zero elements in \mathbf{C} when calculating this trace since the derivative $\partial\mathbf{C}$ is zero for all zero elements of \mathbf{C} . There is a therefore a huge advantage in forming as few extra elements as possible when forming the required elements of \mathbf{C}^{-1} . This is controlled by the order in which the equations are processed.

The AI algorithm requires space for \mathbf{C}^{-1} and for working variables and their cross-products. ASReml uses the same memory to form the mixed model equations, absorb them and then hold the (partial) inverse.

For convenience, we have described the situation where both \mathbf{R} and \mathbf{G} are positive definite but we in fact also allow special cases where \mathbf{G} is singular or negative definite (for example Thompson et al 2003).

9 Discussion

10 Acknowledgements

11 References

Gilmour, A.R., Thompson, R. and Cullis, B.R. (1995). Average Information, an efficient algorithm for REML estimation in linear mixed models. *Biometrics* **51**, 1440-50.

Gilmour, A.R., Cullis, B.R., Gogel, B.J., Welham, S.J. and Thompson, R. (2002). ASREML User Guide. VSN International, Hemel Hempstead, United

Kingdom. <http://www.asreml.co.uk>

Derivation of the Reduced Animal Model

Consider an animal model analysis which can be represented by the model

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

$$\text{Var} \begin{pmatrix} \mathbf{u} \\ \mathbf{e} \end{pmatrix} = \begin{pmatrix} \sigma_A^2 \mathbf{A} & 0 \\ 0 & \sigma_E^2 \mathbf{I} \end{pmatrix}$$

Letting $\gamma = \sigma_A^2/\sigma_E^2$; $\lambda = 1./\gamma$, the mixed model equations can be represented by

$$\begin{pmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z} \\ \mathbf{Z}'\mathbf{X} & \mathbf{Z}'\mathbf{Z} + \lambda\mathbf{A}^{-1} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta} \\ \mathbf{u} \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{Z}'\mathbf{y} \end{pmatrix}$$

Now order the data file and pedigree file so that the last p rows relate to p progeny with own data but no descendants in the data or pedigree file.

Represent the data/design as

$$\begin{pmatrix} \mathbf{y}_a & \mathbf{X}_a & \mathbf{Z}_a & \mathbf{0} \\ \mathbf{y}_p & \mathbf{X}_p & \mathbf{0} & \mathbf{I}_p \end{pmatrix}$$

and $\mathbf{A}^{-1} = \begin{pmatrix} \mathbf{A}_a^{-1} + \mathbf{A}_a^{-1}\mathbf{B}\mathbf{Q}\mathbf{B}'\mathbf{A}_a^{-1} & -\mathbf{A}_a^{-1}\mathbf{B}\mathbf{Q} \\ -\mathbf{Q}\mathbf{B}'\mathbf{A}_a^{-1} & \mathbf{Q} \end{pmatrix}$ noting that in this case of a relationship matrix, \mathbf{Q} is diagonal.

The elements of \mathbf{Q} are calculated from the inbreeding coefficients of the parents as $1/(1 - (a_{jj} + a_{kk})/4)$ where $a_{jj} = 1 + i_j$ and $a_{kk} = 1 + i_k$ and the relationship matrix coefficients for the parents, with inbreeding coefficients i_j and i_k respectively. The elements of $\mathbf{B}'\mathbf{A}^{-1}$ are all zero except cells ij and ik are 0.5.

The mixed model equations are then represented by

$$\begin{pmatrix} \mathbf{X}'_a\mathbf{X}_a + \mathbf{X}'_p\mathbf{X}_p & \mathbf{X}'_a\mathbf{Z}_a & \mathbf{X}'_p \\ \mathbf{Z}'_a\mathbf{X}_a & \mathbf{Z}'_a\mathbf{Z}_a + \lambda\mathbf{A}_a^{-1} + \lambda\mathbf{A}_a^{-1}\mathbf{B}\mathbf{Q}\mathbf{B}'\mathbf{A}_a^{-1} - \lambda\mathbf{A}_a^{-1}\mathbf{B}\mathbf{Q} \\ \mathbf{X}_p & -\lambda\mathbf{Q}\mathbf{B}'\mathbf{A}_a^{-1} & \mathbf{I} + \lambda\mathbf{Q} \end{pmatrix} \times$$

$$\begin{pmatrix} \beta \\ \mathbf{u}_a \\ \mathbf{u}_p \end{pmatrix} = \begin{pmatrix} \mathbf{X}'_a \mathbf{vec} \mathbf{y}_a + \mathbf{X}'_p \mathbf{y}_p \\ \mathbf{Z}'_a \mathbf{y}_a \\ \mathbf{y}_p \end{pmatrix}$$

Absorbing the progeny equations gives

$$\begin{pmatrix} \mathbf{X}'_a \mathbf{X}_a + \mathbf{X}'_p [\mathbf{I} - (\mathbf{I} + \lambda \mathbf{Q})^{-1}] \mathbf{X}_p & \mathbf{X}'_a \mathbf{Z}_a + \mathbf{X}'_p (\mathbf{I} + \lambda \mathbf{Q})^{-1} \lambda \mathbf{Q} \mathbf{B}' \mathbf{A}_a^{-1} \\ \mathbf{Z}'_a \mathbf{X}_a + \lambda \mathbf{A}_a^{-1} \mathbf{B} \mathbf{Q} (\mathbf{I} + \lambda \mathbf{Q})^{-1} \mathbf{X}_p & \mathbf{Z}'_a \mathbf{Z}_a + \lambda \mathbf{A}_a^{-1} + \lambda \mathbf{A}_a^{-1} \mathbf{B} \mathbf{Q} [\mathbf{I} - (\mathbf{I} + \lambda \mathbf{Q})^{-1} \lambda \mathbf{Q}] \mathbf{B}' \mathbf{A}_a^{-1} \end{pmatrix} \times$$

$$\begin{pmatrix} \beta \\ \mathbf{u}_a \end{pmatrix} = \begin{pmatrix} \mathbf{X}'_a \mathbf{y}_a + \mathbf{X}'_p [\mathbf{I} - (\mathbf{I} + \lambda \mathbf{Q})^{-1}] \mathbf{y}_p \\ \mathbf{Z}'_a \mathbf{y}_a + \lambda \mathbf{A}_a^{-1} \mathbf{B} \mathbf{Q} (\mathbf{I} + \lambda \mathbf{Q})^{-1} \mathbf{y}_p \end{pmatrix}$$

$$\text{Let } \mathbf{W}_p = \mathbf{I} - (\mathbf{I} + \lambda \mathbf{Q})^{-1} = (\mathbf{I} + \lambda \mathbf{Q} - \mathbf{I})(\mathbf{I} + \lambda \mathbf{Q})^{-1} = \lambda \mathbf{Q} (\mathbf{I} + \lambda \mathbf{Q})^{-1}$$

$$\lambda \mathbf{Q} [\mathbf{I} - (\mathbf{I} + \lambda \mathbf{Q})^{-1} \lambda \mathbf{Q}] = \lambda \mathbf{Q} (\mathbf{I} + \lambda \mathbf{Q})^{-1} [\mathbf{I} + \lambda \mathbf{Q} - \lambda \mathbf{Q}] = \mathbf{W}_p$$

giving

$$\begin{pmatrix} \mathbf{X}'_a \mathbf{X}_a + \mathbf{X}'_p \mathbf{W}_p \mathbf{X}_p & \mathbf{X}'_a \mathbf{Z}_a + \mathbf{X}'_p \mathbf{W}_p \mathbf{B}' \mathbf{A}_a^{-1} \\ \mathbf{Z}'_a \mathbf{X}_a + \mathbf{A}_a^{-1} \mathbf{B} \mathbf{W}_p \mathbf{X}_p & \mathbf{Z}'_a \mathbf{Z}_a + \lambda \mathbf{A}_a^{-1} + \mathbf{A}_a^{-1} \mathbf{B} \mathbf{W}_p \mathbf{B}' \mathbf{A}_a^{-1} \end{pmatrix} \times$$

$$\begin{pmatrix} \beta \\ \mathbf{u}_a \end{pmatrix} = \begin{pmatrix} \mathbf{X}'_a \mathbf{y}_a + \mathbf{X}'_p \mathbf{W}_p \mathbf{y}_p \\ \mathbf{Z}'_a \mathbf{y}_a + \mathbf{A}_a^{-1} \mathbf{B} \mathbf{W}_p \mathbf{y}_p \end{pmatrix}$$

At this point, let $\mathbf{Z}_p = \mathbf{B}' \mathbf{A}_a^{-1}$ and the equations become

$$\begin{pmatrix} \mathbf{X}'_a \mathbf{X}_a + \mathbf{X}'_p \mathbf{W}_p \mathbf{X}_p & \mathbf{X}'_a \mathbf{Z}_a + \mathbf{X}'_p \mathbf{W}_p \mathbf{Z}_p \\ \mathbf{Z}'_a \mathbf{X}_a + \mathbf{Z}'_p \mathbf{W}_p \mathbf{X}_p & \mathbf{Z}'_a \mathbf{Z}_a + \lambda \mathbf{A}_a^{-1} + \mathbf{Z}'_p \mathbf{W}_p \mathbf{Z}_p \end{pmatrix} \times$$

$$\begin{pmatrix} \beta \\ \mathbf{u}_a \end{pmatrix} = \begin{pmatrix} \mathbf{X}'_a \mathbf{y}_a + \mathbf{X}'_p \mathbf{W}_p \mathbf{y}_p \\ \mathbf{Z}'_a \mathbf{y}_a + \mathbf{Z}'_p \mathbf{W}_p \mathbf{y}_p \end{pmatrix}$$

So, we have a reduced set of equations formed in the normal way with respect to the parental data, and with weights \mathbf{W}_p and a special design matrix (\mathbf{Z}_p) for the offspring data.

The weights are derived from the diagonal of the inverse of the A matrix and the special design matrix is from the parent/offspring block of the A-inverse.

Consider a six animal pedigree

```
201 101 102
202 101 102
301 201 202
302 201 202
```

The A-inverse (obtained from ASReml using the !GIV qualifier) is

```
      101  102  201  202  301  302
101    2
102    1    2
201   -1   -1    3
202   -1   -1    1    3
301    0    0   -1   -1    2
302    0    0   -1   -1    0    2
```

$$\text{So } \mathbf{Q} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}; \mathbf{W}_p = \lambda \mathbf{Q}(\mathbf{I} + \lambda \mathbf{Q})^{-1}$$

\mathbf{Z}_a is defined in **ASReml** as `id.Parent`, \mathbf{Z}_p is defined as `and(sire.Proj.Half) and(Half.dam.Proj)` where `Half` is a variate with values all 0.5 and the `and()` function overlays the design matrix.

So, to use the reduced animal model, we need to augment the data file by the weights. The `!DIAG` qualifier used on the whole pedigree will write the \mathbf{Q} values to `ainverse.dia`. Alternatively, the \mathbf{Q} values can be worked out from the inbreeding coefficients of the parents.

Extension to Maternal Grandsire model

Consider we have an existing relationship matrix and its inverse involving **MGS** (row 1) and **SIRE** (row 2) represented by

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}^{-1} = \begin{pmatrix} a^{11} & a^{12} \\ a^{21} & a^{22} \end{pmatrix}$$

. Add rows for an unrelated **MGD** (row 3) and their progeny **DAM** (row 4).

$$\begin{pmatrix} a^{11} + q_d/4 & a^{12} & 0 + q_d/4 & 0 - q_d/2 \\ a^{21} & a^{22} & 0 & 0 \\ 0 + q_d/4 & 0 & 1 + q_d/4 & 0 - q_d/2 \\ 0 - q_d/2 & 0 & 0 - q_d/2 & 0 + q_d \end{pmatrix}$$

where $q_d = 1/(1 - (a_{11} + a_{33})/4)$ and $a_{33} = 1$ so that $q_d = 4/(3 - a_{11})$.

We no longer require **MGD** so absorb row 3 to give

$$\begin{pmatrix} a^{11} + 1/(4 - a_{11}) & a^{12} & -2/(4 - a_{11}) \\ a^{21} & a^{22} & 0 \\ -2/(4 - a_{11}) & 0 & 4/(4 - a_{11}) \end{pmatrix}$$

since

$$\begin{aligned} q_d[1 - q_d/4/(1 + q_d/4)] &= q_d[1 - 1/(3 - a_{11})/[(3 - a_{11} + 1)/(3 - a_{11})]] \\ &= q_d[1 - 1/(4 - a_{11})] = 4/(4 - a_{11}) \end{aligned}$$

Adding the progeny of **SIRE** and **DAM** as new row 4 gives

$$\begin{pmatrix} a^{11} + 1/(4 - a_{11}) & a^{12} & -2/(4 - a_{11}) & 0 \\ a^{21} & a^{22} + q_p/4 & 0 + q_p/4 & -q_p/2 \\ -2/(4 - a_{11}) & 0 + q_p/4 & 4/(4 - a_{11}) + q_p/4 & -q_p/2 \\ 0 & -q_p/2 & -q_p/2 & q_p \end{pmatrix}$$

where $q_p = 1/(1 - (a_{22} + a_{44})/4)$

and $a_{44} = 1$ since its dam was unknown so that $q_p = 4/(3 - a_{22})$.

Finally, we absorb **DAM** (row 3).

Let

$$\begin{aligned} D &= 1/[4/(4 - a_{11}) + q_p/4] \\ &= 1/[4/(4 - a_{11}) + 1/(3 - a_{22})] \\ &= (4 - a_{11})(3 - a_{22})/(4(3 - a_{22}) + 4 - a_{11}) \\ &= (4 - a_{11})(3 - a_{22})/(16 - 4a_{22} - a_{11}) \end{aligned}$$

Cell 1, 1 becomes

$$a^{11} + 1/(4 - a_{11}) - 4D/(4 - a_{11})^2 = a^{11} + (1 - 4(3 - a_{22})/(16 - 4a_{22} - a_{11}))/ (4 - a_{11})$$

$$= a^{11} + 1/(16 - 4a_{22} - a_{11})$$

Cell 2, 1=Cell 1, 2 becomes

$$a^{21} + 2D/(4 - a_{11})/(3 - a_{22}) = a^{21} + 2/(16 - 4a_{22} - a_{11}).$$

Cell 2, 2 becomes

$$\begin{aligned} a^{22} + q_p/4 - Dq_p^2/16 &= a^{22} + (1 - (4 - a_{11})/(16 - 4a_{22} - a_{11}))/ (3 - a_{22}) \\ &= a^{22} + 4/(16 - 4a_{22} - a_{11}) \end{aligned}$$

Cell 4, 1 (1, 4) becomes

$$-2Dq_p/2/(4 - a_{11}) = -4/(16 - 4a_{22} - a_{11}).$$

Cell 4, 2 (2, 4) becomes

$$\begin{aligned} -q_p/2 + Dq_p^2/8 &= -2(1 - (4 - a_{11})/(16 - 4a_{22} - a_{11}))/ (3 - a_{22}) \\ &= -8/(16 - 4a_{22} - a_{11}). \end{aligned}$$

Cell 4, 4 becomes

$$q_p - Dq_p^2/4 = q_p(1 - Dq_p/4) = 16/(16 - 4a_{22} - a_{11}).$$

Let $Q_p = 16/(16 - 4a_{22} - a_{11})$ and the matrix becomes

$$\begin{pmatrix} a^{11} + Q_p/16 & a^{12} + Q_p/8 & -Q_p/4 \\ a^{21} + Q_p/8 & a^{22} + Q_p/4 & -Q_p/2 \\ -Q_p/4 & -Q_p/2 & Q_p \end{pmatrix}$$

For non inbred parents, $a_{11} = a_{22} = 1$ and $Q_p = 16/11$

Consequently, we can also use the RAM method with a maternal grandsire pedigree by using the model terms

\mathbf{Z}_a is defined in **ASReml** as `id.Parent`, \mathbf{Z}_p is defined as `and(Half.sire*Proj)`
`and(mgs.Proj, 0.25)`

and using weights for the progeny records calculated using $Q_p = 16/(16 - 4a_{22} - a_{11}) = 16/(11 - 4i_s - i_{mgs})$.

Estimating the variance parameters under the RAM model

As formulated, the RAM model cannot be used to estimate the variance ratio because **ASReml** cannot handle the differential of the weight with respect to

the variance ratio. It can however be estimated by putting the extra progeny variance into an extra variance component. It is not obvious that this would provide any computational advantage over using the IAM since the original order (number of) of equations is retained.

The expanded equations become

$$\begin{pmatrix} \mathbf{X}'_a \mathbf{X}_a + \mathbf{X}'_p \mathbf{X}_p & \mathbf{X}'_a \mathbf{Z}_a + \mathbf{X}'_p \mathbf{Z}_p & \mathbf{X}'_p \\ \mathbf{Z}'_a \mathbf{X}_a + \mathbf{Z}'_p \mathbf{X}_p & \mathbf{Z}'_a \mathbf{Z}_a + \lambda \mathbf{A}_a^{-1} + \mathbf{Z}'_p \mathbf{Z}_p & \mathbf{Z}'_p \\ \mathbf{X}_p & \mathbf{Z}_p & \mathbf{I} + \lambda \mathbf{Q} \end{pmatrix} \times$$

$$\begin{pmatrix} \beta \\ \mathbf{u}_a \\ \epsilon_p \end{pmatrix} = \begin{pmatrix} \mathbf{X}'_a \mathbf{y}_a + \mathbf{X}'_p \mathbf{y}_p \\ \mathbf{Z}'_a \mathbf{y}_a + \mathbf{Z}'_p \mathbf{y}_p \\ \mathbf{y}_p \end{pmatrix}$$

noting that λ appears in two places. The extra random effect is then the deviation of the animal's BLUP from its midparent value.

An example

Following is a data file (`anim.ped`) generated in S-Plus containing 4 generations. We then show an **ASReml** job which fits the full animal model and estimates the variance components, and an **ASReml** job which fits the reduced animal model to estimated the BLUP values.

The first 70 lines of `anim.ped` relate to Parents, the remaining 70 lines relate to progeny without descendents. The first three columns contain the full pedigree. The column `ramid` is like `animal` except that it is zero for the progeny lines. The `progeny` variable is 0 for parents, 1 for progeny (it could be created from the `ramid` column via transformation). `ibv` is the inbreeding associated with the animal (times 32 to make it an integer) and is there just for interest. The progeny rows of `Q` contain the diagonal of the inverse of the full `A` inverse for these animals.

```

animal sire dam ramid y progeny ibv Q
  11 1 4  11 10.3908857477546  0 0 0
  12 1 5  12 11.1104657392240  0 0 0
  13 2 6  13 11.7779552726571  0 0 0
  14 2 4  14 8.6741538043638  0 0 0
  15 2 7  15 7.91890378030102  0 0 0
  16 3 5  16 9.71655429296733  0 0 0
  17 3 8  17 10.8904473767759  0 0 0
  18 3 9  18 9.61207680583325  0 0 0
  19 3 6  19 10.5227820706866  0 0 0

```

20	1	10	20	9.34303055021789	0	0	0
21	1	9	21	9.36021403905914	0	0	0
22	1	10	22	11.0300042996657	0	0	0
23	2	8	23	10.2060540668193	0	0	0
24	2	9	24	10.3261486305873	0	0	0
25	2	10	25	10.1435062911543	0	0	0
26	2	4	26	9.38946394377953	0	0	0
27	3	5	27	10.3082467403304	0	0	0
28	3	6	28	10.7775996888753	0	0	0
29	3	7	29	9.45255896227042	0	0	0
30	3	8	30	10.0005912543929	0	0	0
31	19	26	31	10.6833370443366	0	0	0
32	12	24	32	10.5686070634204	0	0	0
33	20	24	33	11.1517440567010	0	0	0
34	16	30	34	10.0286689932973	0	4	0
35	19	22	35	11.1329794620133	0	0	0
36	14	23	36	9.80581432197048	0	4	0
37	15	25	37	10.0165440616599	0	4	0
38	16	27	38	9.63089952169726	0	8	0
39	15	27	39	9.07073787302496	0	0	0
40	20	24	40	9.46343346201765	0	0	0
41	13	25	41	10.9546878217872	0	4	0
42	11	21	42	9.69861973870184	0	4	0
43	16	23	43	9.1415847358994	0	0	0
44	12	21	44	8.95129961418688	0	4	0
45	18	25	45	11.2014778719284	0	0	0
46	16	21	46	10.0834481927450	0	0	0
47	19	21	47	9.36277506264529	0	0	0
48	11	25	48	11.3906966649587	0	0	0
49	16	26	49	9.3760267049459	0	0	0
50	14	30	50	8.9478418940027	0	0	0
51	14	22	51	10.5752778989598	0	0	0
52	12	22	52	8.636365924336	0	4	0
53	14	28	53	10.5596974215369	0	0	0
54	11	24	54	10.6235872497733	0	0	0
55	17	30	55	10.7208445263010	0	8	0
56	11	30	56	11.1845127701036	0	0	0
57	13	21	57	10.6176220363483	0	0	0
58	18	29	58	10.2376357012145	0	4	0
59	15	29	59	9.50752581920866	0	4	0
60	14	25	60	11.5727423713477	0	4	0
61	11	29	61	10.3706427153618	0	0	0
62	12	25	62	9.84133070107534	0	0	0
63	15	28	63	9.56209348619	0	0	0
64	15	28	64	9.48594143828686	0	0	0
65	19	26	65	10.5423939080058	0	0	0
66	15	22	66	9.51586858551116	0	0	0

67	14	30	67	9.43485295607768	0	0	0
68	16	23	68	9.75749335918498	0	0	0
69	14	23	69	9.49082318556744	0	4	0
70	11	22	70	10.180415181553	0	4	0
71	14	28	71	10.5015621557079	0	0	0
72	15	28	72	9.38552780572869	0	0	0
73	12	25	73	10.0811885517726	0	0	0
74	20	24	74	9.98029685584483	0	0	0
75	18	27	75	10.3921923488348	0	4	0
76	12	21	76	11.2095368426887	0	4	0
77	14	30	77	11.4028159057297	0	0	0
78	16	27	78	9.03775130221661	0	8	0
79	18	22	79	11.3811226071947	0	0	0
80	18	25	80	7.8124711265838	0	0	0
81	37	54	0	9.41301907943905	1	2	2.1333
82	32	56	0	10.9478210963465	1	1	2.0000
83	43	51	0	10.4108227824721	1	1	2.0000
84	38	63	0	9.50198191412658	1	2	2.2857
85	49	56	0	10.1836742661691	1	2	2.0000
86	36	67	0	9.43561878646791	1	6	2.1333
87	39	54	0	8.96051177807472	1	1	2.0000
88	37	57	0	10.6006629223889	1	2	2.1333
89	40	61	0	10.8951841890337	1	1	2.0000
90	31	59	0	11.6171466747894	1	2	2.1333
91	34	71	0	8.82615618148376	1	2	2.1333
92	48	75	0	10.0784745397543	1	0	2.1333
93	39	62	0	10.1091215894326	1	2	2.0000
94	44	73	0	10.0999858381977	1	5	2.1333
95	32	74	0	8.91637153793524	1	5	2.0000
96	49	74	0	9.98534429588077	1	1	2.0000
97	42	80	0	9.72385207966892	1	1	2.1333
98	50	74	0	10.4435420931982	1	1	2.0000
99	40	51	0	10.1924373987657	1	3	2.0000
100	46	57	0	9.61540429177981	1	4	2.0000
101	49	69	0	10.2832108644091	1	3	2.1333
102	44	51	0	10.7724842239921	1	2	2.1333
103	40	53	0	9.84204430655675	1	1	2.0000
104	39	57	0	7.91867388041925	1	1	2.0000
105	36	62	0	10.1864922259464	1	2	2.1333
106	44	80	0	10.3704117751821	1	1	2.1333
107	43	64	0	10.4920572539863	1	2	2.0000
108	42	72	0	8.36290471333196	1	0	2.1333
109	45	79	0	10.2650472908691	1	5	2.0000
110	42	56	0	10.2986886395126	1	5	2.1333
111	44	80	0	10.6009452530656	1	1	2.1333
112	38	74	0	8.67042728685292	1	0	2.2857
113	40	72	0	11.0634565743273	1	1	2.0000

114	31	61	0	8.89534464048755	1	2	2.0000
115	49	71	0	10.8962430535315	1	3	2.0000
116	39	62	0	9.18754090244638	1	2	2.0000
117	46	66	0	10.8691783663719	1	1	2.0000
118	36	77	0	10.2807432360335	1	6	2.1333
119	39	56	0	9.10687842385982	1	1	2.0000
120	46	52	0	11.8720666576208	1	3	2.1333
121	47	72	0	10.1850603552501	1	2	2.0000
122	46	54	0	9.6558797999729	1	2	2.0000
123	32	57	0	9.22850429171406	1	3	2.0000
124	45	66	0	8.15136825542446	1	2	2.0000
125	46	73	0	7.88941183740888	1	2	2.0000
126	39	53	0	9.01615309035975	1	2	2.0000
127	31	55	0	10.3523192574918	1	2	2.2857
128	43	79	0	10.2148135862004	1	1	2.0000
129	44	76	0	8.50957680401474	1	10	2.2857
130	32	54	0	9.53000303131414	1	5	2.0000
131	44	52	0	9.33721586426789	1	7	2.2857
132	36	78	0	8.48635421527905	1	0	2.4615
133	31	77	0	11.3329385837141	1	3	2.0000
134	37	56	0	10.4134913778961	1	0	2.1333
135	46	70	0	9.56323081820872	1	2	2.1333
136	47	61	0	9.730332323743	1	2	2.0000
137	50	79	0	8.65562598882075	1	1	2.0000
138	35	54	0	9.97523579491704	1	1	2.0000
139	34	79	0	8.2347003348346	1	2	2.1333
140	49	60	0	10.2526254965016	1	3	2.1333
141	46	69	0	9.72524883824503	1	0	2.1333
142	34	70	0	10.2568000361074	1	0	2.2857
143	32	69	0	10.2691991448205	1	2	2.1333
144	44	79	0	7.99985920710376	1	3	2.1333
145	45	51	0	10.1699822966765	1	2	2.0000
146	31	61	0	9.35249803893117	1	2	2.0000
147	50	80	0	9.2115314170863	1	2	2.0000
148	42	63	0	10.7952258148718	1	0	2.1333
149	39	57	0	9.8647096375117	1	1	2.0000
150	45	80	0	10.4279842491734	1	8	2.0000

The first job fits the animal model (estimating the variance component).

```
Analysis by animal model
```

```
animal !P
```

```
sire !P
```

```
dam !P
```

```
ramid Y
```

```
anim.ped !skip 1 !make !diag
```

```
anim.ped !skip 1 !extra 3
Y ~ mu !r anim
```

An extract from the .asr file follows.

```
ASReml 1.57 [04 Dec 2003] Analysis by animal model
16 Jan 2004 14:46:52.155 64.00 Mbyte Windows anim
```

Folder: C:\data\proj\agbu

animal !P

sire !P

dam !P

Reading pedigree file anim.ped : skipping 1 lines

PEDIGREE [anim.ped] has 150 identities, 547 Non zero elements

QUALIFIERS: !SKIP 1 !Evec{X}TRA 3

Reading anim.ped FREE FORMAT skipping 1 lines

Univariate analysis of Y

Using 140 records of 140 read

Model term		Size	Minimum	Mean	Maximum	#zero	#miss
1 animal	!P	150	3.000	80.24	150.0	0	0
2 sire	!P	150	2.000	25.43	50.00	0	0
3 dam	!P	150	1.000	43.07	80.00	0	0
4 ramid			11.00	22.75	80.00	70	0
5 Y	Variate		7.812	9.944	11.87	0	0
6 mu		1					

Forming 151 equations: 1 dense.

Initial updates will be shrunk by factor 0.316

1 LogL=-53.3627	S2= 0.70623	139 df	0.1000	1.000
2 LogL=-53.3585	S2= 0.70315	139 df	0.1060	1.000
3 LogL=-53.3550	S2= 0.69912	139 df	0.1140	1.000
4 LogL=-53.3538	S2= 0.69537	139 df	0.1216	1.000
5 LogL=-53.3538	S2= 0.69485	139 df	0.1226	1.000
6 LogL=-53.3538	S2= 0.69479	139 df	0.1228	1.000
7 LogL=-53.3538	S2= 0.69478	139 df	0.1228	1.000
8 LogL=-53.3538	S2= 0.69478	139 df	0.1228	1.000

Final parameter values 0.12278 1.0000

Degrees of Freedom and Stratum Variances

67.43	0.844141	1.8	1.0
71.57	0.694780	0.0	1.0

Source	Model	terms	Gamma	Component	Comp/SE	% C
animal	150	150	0.122781	0.853059E-01	0.80	0 P
Variance	140	139	1.00000	0.694780	5.98	0 P

Analysis of Variance	NumDF	DenDF	F-incr	F-adj	Prob
6 mu	1	2.6	5366.78	NA	NA

	Estimate	Standard Error	T-value	T-prev
6 mu				
1	9.97036	0.136099	73.26	
1 animal		150 effects fitted		
Finished: 16 Jan 2004 14:46:53.737		LogL Converged		

The !DIAG qualifier caused the elements of Q to be written to `ainverse.dia` from whence they were copied into the data file. The Q values can also be worked out directly (not within **ASReml**) from the inbreeding coefficients of the parents. For example, **81** has parents **37** and **54** with inbreeding coefficients of $4/32$ and $0/32$ respectively. $q_{81} = 1/(1 - (1 + 4/32 + 1 + 0/32)/4) = 32/(16 - 1) = 2.13333$. For non inbred parents, the coefficient is 2. **84** has parents **38** and **63** with inbreeding coefficients of $8/32$ and $0/32$ respectively. $q_{84} = 1/(1 - (1 + 8/32 + 1 + 0/32)/4) = 32/(16 - 2) = 2.28571$.

The job to estimate the effects using the reduced animal model is

```

Analysis by Reduced animal model
! From animal model, gamma = 0.122781
animal
sire !P !*V6
dam !P !*V6
ramid !P
Y
Progeny Parent !=1 !-Prog
Weight !/0.122781 !+1 !^-1 !-1 !*-1 !+Parent
ram.ped !skip 1
anim.ped !skip 1 !MAvec{X}IT 1 !BLUP 2

Y !wt We ~ mu !r ramid .122781 , and(sire,0.5) and(dam,0.5)

```

This job uses the same data file but the pedigree file just contains the parents (the first 70 lines of `anim.ped`)

The `sire` and `dam` fields are multiplied by `Progeny (V6)` to annihilate the information in them in the parent records. `ibv` is overwritten by a created variable `Parent` which is the complement of `Progeny` and is subsequently used to give a weight of 1. to the parent records.

Q is converted to the weight using $-((Q/\gamma + 1)^{-1} - 1)$. Since Q was supplied as 0 for the parents, this generates 0 for the parental records hence we then added the `Parent` indicator variable to set the weight to 1 for the parental records.

Running this job gave

ASReml 1.57 [04 Dec 2003] Analysis by Reduced animal model
16 Jan 2004 20:10:44.768 64.00 Mbyte Windows ram

! From animal model, gamma = 0.122781
Folder: C:\data\proj\agbu
sire !P !*V6
dam !P !*V6
ramid !P
Progeny Parent !=1 !-Prog Weight !/0.122781 !+1 !^-1 !-1 !*-1 !+Parent
A-inverse retrieved from ainverse.bin
PEDIGREE [ram.ped] has 80 identities, 271 Non zero elements
QUALIFIERS: !SKIP 1 !MAvec{X}IT 1 !BLUP 2
Reading anim.ped FREE FORMAT skipping 1 lines

Univariate analysis of Y

Using 140 records of 140 read

Model term	Size	Minimum	Mean	Maximum	#zero	#miss
1 animal		11.00	80.50	150.0	0	0
2 sire	!P 80	31.00	20.43	50.00	70	0
3 dam	!P 80	51.00	32.62	80.00	70	0
4 ramid	!P 80	3.000	22.49	80.00	70	0
5 Y	Variate	7.812	9.944	11.87	0	0
6 Progeny		1.000	0.5000	1.000	70	0
7 Parent		1.000	0.5000	1.000	70	0
8 Weight	Weight	0.9422	0.9721	1.000	0	0
9 mu	1					
10 and(sire,0.5)	80					
11 and(mgs,0.5)	80					

Forming 81 equations: 1 dense.

1 LogL=-53.3538 S2= 0.69478 139 df

Warning: The estimation was ABORTED because the !BLUP qualifier was specified.
The Standard Errors of variance components (and solutions?)
and derived quantities are wrong.
Use !MAvec{X}IT 1 instead of !BLUP to get correct standard errors.

...

	Estimate	Standard Error	T-value	T-prev
9 mu				
	1 9.97036	0.136099	73.26	
4 ramid		80 effects fitted		

Finished: 16 Jan 2004 20:10:45.899 BLUP run done

Notice that the LogL, S2 and mu values agree with the full animal model results.

Furthermore, the BLUPS for the parents agree. The following shows some

values.

		----- anim.sln ----		--- ram.sln -----	
mu	1	9.970	0.1361	9.970	0.1361
animal	4	0.6236E-01	0.2705	0.6236E-01	0.2705
animal	1	0.2862E-01	0.2606	0.2862E-01	0.2606
animal	11	0.1169	0.2535	0.1169	0.2535
animal	5	-0.1774	0.2681	-0.1774	0.2681
animal	12	-0.6694E-01	0.2497	-0.6694E-01	0.2497
animal	6	0.2221	0.2737	0.2221	0.2737
animal	2	-0.1436E-01	0.2607	-0.1436E-01	0.2607
animal	13	0.2136	0.2643	0.2136	0.2643
animal	14	-0.2214E-02	0.2482	-0.2214E-02	0.2482
animal	7	-0.2042	0.2755	-0.2042	0.2755
animal	37	-0.7297E-01	0.2714	-0.7297E-01	0.2714
animal	54	0.2736E-01	0.2599	0.2736E-01	0.2599
animal	70	0.1271	0.2727	0.1271	0.2727
animal	71	0.8816E-01	0.2644	0.8816E-01	0.2644
animal	72	-0.1188	0.2614	-0.1188	0.2614
animal	73	-0.1705E-01	0.2631	-0.1705E-01	0.2631
animal	74	-0.2541E-01	0.2601	-0.2541E-01	0.2601
animal	75	-0.1041	0.2768	-0.1041	0.2768
animal	76	-0.5544E-01	0.2693	-0.5544E-01	0.2693
animal	77	0.1469	0.2615	0.1469	0.2615
animal	78	-0.2678	0.2840	-0.2678	0.2840
animal	79	-0.1235E-01	0.2592	-0.1235E-01	0.2592
animal	80	-0.8484E-01	0.2598	-0.8484E-01	0.2598
animal	81	-0.5190E-01	0.2748		
animal	82	0.1257	0.2744		
animal	83	0.4467E-01	0.2753		
animal	84	-0.1789	0.2732		
animal	85	0.6887E-01	0.2767		

ASReml does not work out the Progeny breeding values but they can be obtained as follows. Animal 81 has parents 37 and 54 whose BLUPS are

animal	37	-0.7297E-01	0.2714
animal	54	0.2736E-01	0.2599

and has residual of

71	9.9476	-0.5345	1.000
----	--------	---------	-------

So its BLUP = $0.5(-.07297+.02736) + 1/(1+Q)(-.5345)$ where $Q = 2.1333/.122781$

= 17.37484

BLUP = $0.5(-.04561) - .02909 = -.02280 - .02909 = -0.05189$ which compares well with the animal model solution.

Further checking

Modifying pedigree of the last three animals, the last three lines of the data file now read

```
147 50 80 0 9.2115314170863 1 0 2 2.0000
148 42 0 0 10.7952258148718 1 2 0 1.3913
149 0 57 0 9.8647096375117 1 0 1 1.3333
150 0 0 0 10.4279842491734 1 0 8 1.0000
```

where the Q values have also been updated to reflected the changed pedigree (diagonal elements of \mathbf{A}^{-1}).

Running the IAM on the revised data gave LogL=-53.2775, S2=0.68607 and gamma=0.138732. Selected BLUPS are

animal	31	0.1804	0.2715
animal	42	-0.5272E-01	0.2795
animal	50	-0.9671E-01	0.2716
animal	57	0.1613E-02	0.2722
animal	61	-0.2036E-01	0.2743
animal	80	-0.1335	0.2739
animal	147	-0.1579	0.2899
animal	148	0.4939E-01	0.2824
animal	149	-0.1071E-01	0.2850
animal	150	0.5383E-01	0.2896

Re running the RAM job (referring to the modified data file and with gamma=0.138732) gave identical values for LogL, S2, gamma and the BLUPS. The BLUPS for the last 4 animals are given by

147: $(-0.09671 - .1335)/2 + (-.6595)/(1 + 2 \cdot .138732) = -.15788$

148: $(-0.05272 + 0.000)/2 + (0.8355)/(1 + 1.3913 \cdot .138732) = 0.049397$

149: $(0.00 + .001613)/2 + (-.1222)/(1 + 1.3333 \cdot .138732) = -.01071077$

150: $(0.00 + 0.00)/2 + (0.4419)/(1 + 1 \cdot .138732) = 0.053837$

Extension to Maternal Grandsire model

The machinery also works for the maternal grandsire model. Using the same

data file except for the Q column, we compare the results from a direct MGS analysis performed using

```

Analysis by animal model - Maternal grandsires pedigree
animal !P
sire !P
mgs !P
ramid Y
anim.mgs !skip 1 !mgs
anim.mgs !skip 1 !extra 3
Y ~ mu !r anim

```

and a RAM model analysis using

```

Analysis by Reduced animal model : Maternal grandsire pedigree
! From animal model, gamma = 0.245788
animal
sire !P !*V6
mgs !P !*V6
ramid !P
Y
Progeny Skip Parent !=1 !-Prog
Weight !/16 !*0.245788 !^-1 !V10=1 !+V9 !V9 !/V10 !+Parent
ram.ped !skip 1 !MGS !DIAG
anim.mgs !skip 1 !MAvec{X}IT 1 !BLUP 2
Y !wt We ~ mu !r ramid .245788 , and(sire,0.5) and(mgs,0.25)

```

As part of this exercise, I made **ASReml** report inbreeding coefficients as an extension of the !DIAG qualifier output. Previous versions printed the diagonal elements of \mathbf{A}^{-1} to AINVERSE.DIA. The new version also prints the inbreeding coefficients. After an initial run of the second job to obtain the inbreeding coefficients, I calculated $11 - 4i_s - i_{mgs}$ for the progeny records and placed it in the 9th data field. Transformations then converted this to the weight.

Portion of the output from the IAM run follows:

```

ASReml 1.58 [20 Jan 2004]   Analysis by animal model

      8 LogL=-52.9871      S2= 0.62605      139 df      0.2458      1.000

Source          Model terms      Gamma      Component      Comp/SE      % C
animal          290      290      0.245788      0.153877      1.06      0 P
Variance        140      139      1.00000      0.626054      4.46      0 P

      Estimate      Standard Error      T-value      T-prev
6 mu          1      9.96253      0.125665      79.28

```

1 animal 290 effects fitted
 Finished: 28 Jan 2004 13:52:22.651 LogL Converged

The RAM model gave equivalent results:

ASReml 1.58 [20 Jan 2004] Analysis by Reduced animal model

1 LogL=-52.9871 S2= 0.62605 139 df

	Estimate	Standard Error	T-value	T-prev
10 mu	1 9.96253	0.125665	79.28	
4 ramid				150 effects fitted

Finished: 28 Jan 2004 13:53:21.165 BLUP run done

Following is a comparison of breeding values. **ASReml** fits the maternal grand-sire model at present by inserting a dummy DAM for each offspring. Given the algebra in chapter 1, it would be possible to modify the Ainverse algorithm to directly form the MGS Ainverse without inserting the dummy dams.

Identity	IAM-solutions		RAM_solutions	
4	0.3351E-01	0.3818	0.3351E-01	0.3818
999999998	0.1122	0.3732	0.1122	0.3732
1	0.1314	0.3435	0.1314	0.3435
11	0.2491	0.3183	0.2491	0.3183
5	-0.8852E-01	0.3812	-0.8852E-01	0.3812
999999994	-0.3013E-01	0.3726	-0.3013E-01	0.3726
12	0.6950E-01	0.3145	0.6950E-01	0.3145
6	0.2155	0.3831	0.2155	0.3831
999999991	0.2758	0.3780	0.2758	0.3780
2	-0.2119E-01	0.3390	-0.2119E-01	0.3390
13	0.3514	0.3366	0.3514	0.3366
999999988	-0.1295E-01	0.3709	-0.1295E-01	0.3709
14	-0.5668E-01	0.3090	-0.5668E-01	0.3090
7	-0.1837	0.3847	-0.1837	0.3847
..
77	0.2457	0.3399	0.2457	0.3399
999999855	-0.1147	0.3707	-0.1147	0.3707
78	-0.2681	0.3473	-0.2681	0.3473
999999853	0.1617	0.3706	0.1617	0.3706
79	0.1154	0.3389	0.1154	0.3389
999999851	-0.9385E-01	0.3705	-0.9385E-01	0.3705
80	-0.2626	0.3390	-0.2626	0.3390
999999849	0.3336E-01	0.3717		
81	-0.1087	0.3445	-0.10864	
999999847	0.2351	0.3716		
82	0.2383	0.3451	0.23831	
999999845	0.1202	0.3718		
83	0.5213E-01	0.3463	0.05212	

999999843	-0.1291	0.3721	
84	-0.2099	0.3448	-0.20990
..	
147	-0.2627	0.3463	-0.26271
999999715	-0.3639E-01	0.3722	
148	0.7755E-01	0.3433	0.07754
999999713	0.9237E-01	0.3720	
149	-0.2316	0.3438	-0.23162
999999711	-0.9452E-01	0.3708	
150	0.6621E-01	0.3499	0.06621

Again, ASReml does not work out the progeny values directly but they are easily obtained from the parental blups and the residual as
 $BLUP = Sire/2 + MGS/4 + Res / (1 + Q/0.245788)$.

Pedigree	Sire	MGS	Residual	1/(16Q)	BLUP
81 37 54	-0.1457	0.1480	-0.5136	10.75	-0.10864
82 32 56	0.05798	0.3324	0.8732	11.	0.23831
83 43 51	-0.1133	0.1673	0.4631	11.	0.05212
84 38 63	-0.2315	-0.2120	-0.2918	10.6875	-0.20990
147 50 80	-0.2291	-0.2626	-0.5708	11.	-0.26271
148 42 63	0.01169	-0.2120	0.8798	10.75	0.07754
149 39 57	-0.5885	0.1601	0.1564	11.	-0.23162
150 45 80	0.1288	-0.2626	0.4667	11.	0.06621

RAM estimation

The **ASReml** job to estimate the variance ratio using the RAM modelling is

```

Analysis by Reduced animal model
! From animal model, gamma = 0.122781
animal
sire !P !*V6
dam !P !*V6
ramid !P
Y
Progeny #Skip Parent !=1 !-Prog Weight !/0.122781 !+1 !^-1 !-1 !*-1 !+Parent
ram.ped !skip 1
ram.giv # Diagonal matrix containing Q
anim.ped !skip 1 !VCC

```

Y ~ mu !r ramid .122781 , and(sire,0.5) and(dam,0.5) uni(Prog,1,70)

0 0 1
uni 1
uni 0 GIV1 .122781
+ 2 7

which produced the following results:

ASReml 1.58 [20 Jan 2004] Analysis by Reduced animal model
02 Feb 2004 14:52:33.287 64.00 Mbyte Windows rame

! From animal model, gamma = 0.122781
Folder: C:\data\proj\agbu
sire !P !*V6
dam !P !*V6
ramid !P
Reading pedigree file ram.ped : skipping 1 lines
PEDIGREE [ram.ped] has 80 identities, 271 Non zero elements
Reading ram.giv skipping 0 header lines
Inverse G structure of 70 rows having 70 non zero cells read from ram.giv
QUALIFIERS: !SKIP 1 !VCC
Reading anim.ped FREE FORMAT skipping 1 lines
Univariate analysis of Y
Using 140 records of 140 read

Model term	Size	Minimum	Mean	Maximum	#zero	#miss
1 animal		11.00	80.50	150.0	0	0
2 sire	!P 80	31.00	20.43	50.00	70	0
3 dam	!P 80	51.00	32.62	80.00	70	0
4 ramid	!P 80	3.000	22.49	80.00	70	0
5 Y	Variate	7.812	9.944	11.87	0	0
6 Progeny		1.000	0.5000	1.000	70	0
7 mu	1					
8 and(sire,0.5)	80					
9 and(dam,0.5)	80					
10 uni(Prog,1,70)	70	1.000	0.5000	1.000	70	0
70 ram.giv	0.1228					

Structure for uni(Prog,1,70) has 70 levels defined
Forming 151 equations: 1 dense.
Initial updates will be shrunk by factor 0.316
For setting constraints, the variance parameters are numbered 2 to 7

1 LogL=-53.3538	S2= 0.69478	139 df
2 LogL=-53.3538	S2= 0.69478	139 df

Source	Model terms	Gamma	Component	Comp/SE	% C
ramid	80 80	0.122782	0.853063E-01	0.80	0 P 2


```

Variance          140    139    1.00000    0.694780    5.98    0 P
uni(Prog,1,70)   ram.giv    70    0.122782    0.853063E-01    0.00    0 C  2

              Estimate      Standard Error      T-value      T-prev
7 mu
              1    9.97036      0.136099      73.26
4 ramid
              80 effects fitted
10 uni(Prog,1,70)
              70 effects fitted
Finished: 02 Feb 2004 14:52:34.419    LogL Converged

```

Simulated Tree example

This exercise was stimulated by a comparison of Treeplan with ASReml in which ASReml took 30 times longer than Treeplan and this was attributed to the use of the reduced animal model.

Following are results from ASReml using a simulated data set with size and structure somewhat like that used in the comparison with Treeplan. In particular, we have 71000 progeny of 500 parents (randomly allocated to progeny). For the exercise, parents are unrelated. The progeny have been allocated to 71 groups of 1000 each. For the exercise, there is no data on parents.

The following table compares 6 runs. The greatest amount of time is taken forming the A-inverse matrix (about 126 seconds) for the IAM-1 run. Runs IAM-2 and IAM-3 used the A-inverse formed when IAM-1 was run. It took 13 seconds to read the data and set up the design matrix (the data was held as a .csv file). IAM-4 uses a different subroutine to form the A-inverse.

The default in ASReml is to seek to find an optimum equation order to use. This process took 8 seconds but was omitted in IAM-3 as the natural order of equations [groups, parents, progeny] is ideal.

With RAM, we trade a smaller system of equations with a more complex design matrix. The reduced size of the Ainverse matrix means that little time is taken forming A-inverse (0.04 seconds?) but it takes a little (2s) longer to set up the design matrix. The more complex design matrix means it takes a little longer to form the SSP which is now much more dense. Processing it is now a second longer although finding the order is much quicker.

Writing the solutions is slightly faster because BLUPS for the PROGENY are not reported.

Process	IAM-1	IAM-2	IAM-3	IAM-4	RAM-1	RAM-2
Getting Started	129.406	13.179	13.079	21.451	15.152	15.112
R&Gformed	0.150	0.150	0.160	0.180	0.030	0.030
SSP formed	0.070	0.070	0.070	0.060	0.591	0.631

Add Ginverse	0.090	0.090	0.090	0.090	0.010	0.000
Order found	8.372	8.412	-	-	0.881	0.901
SSP absorbed	5.308	5.298	4.336	4.306	6.219	6.219
Iteration complete	0.120	0.090	0.090	0.080	0.080	0.080
Report	5.5	5.5	5.5	5.5	3.0	3.0
Total elapsed time	156.104	36.172	26.929	40.919	28.501	29.853

The bottom line is that for this size problem, there appears little advantage to RAM except the cost of forming the large A-inverse matrix for IAM. Even allowing for this, the difference is 5-fold, not 30-fold.

The A-inverse is formed using, by default, a subroutine obtained from Robin. I had not optimised it but to calculate inbreeding it has a loop across all remaining animals. This searching for animals who are descendants of the current animal is the expensive process. In version 1.58 I have added a test which checks whether there are any offspring first. Using it, the 'Getting started' time reduced to 21.75 seconds (i.e. 8 seconds to form the A-inverse). There is another A-inverse routine in ASReml invoked by the !OLD qualifier which does not handle genetic groups but did have the pre-test for calculating contributions to inbreeding. Using it, 'Getting started' took 21.541 seconds (i.e. it took about 8 seconds to form the A-inverse). Unfortunately, the old code did not handle 'selfing' properly which my test pedigree had some 150 cases of. This is also now fixed in 1.58.

The preceding results were obtained with version 1.57 (1.58). Below is a comparison of various versions of ASReml running the model as in IAM-1.

Version	Elapsed time	
	AIM-1	RAM-1
Jul 1999	173.74	25.85
Nov 2001	157.92	19.92, 15.20
Sep 2003 [110]	156.94	32.25, 28.24
Dec 2003 [157]	151.60	30.92
Feb 2004 [158]	37.24	29.35

These timings were obtained on an ACER Travelmate with 504 MB RAM and a 1000 MHz processor. It is evident PC timings are somewhat variable (10I have not explored at the moment). The Nov 2001 version seems to be reading the data file much faster, probably reflecting less options for decoding the data.

The IAM runs require more memory than the RAM models. If the test machine had less RAM (e.g. 64MB rather than 504 MB), then paging would slow the analysis. The RAM model needed S3 (64MB) but the IAM model needed S4 (128MB) to run. If the IAM model omitted the !BLUP 2 qualifier and had !MAXIT 1 instead, the elapsed time would be increased 5s because of the

extra processing involved.

Discussion and Conclusion

In chapter 1, I have outlined some matrix results supporting the use of **ASReml** for fitting a Reduced Animal model. Chapter 2 provides an example with inbreeding.

RAM is fitted in **ASReml** using two particular features of the program, weights and the `and()` model function. Both of these features are discussed in the 1999 ASReml Reference manual. The only thing that is a little tricky is the working out of the appropriate weights for the RAM analysis and the backsolving for the BLUPs of the absorbed progeny.

The former is based on the diagonal elements of the A^{-1} , and without inbreeding, these values are 1 if no parent is known, $4/3$ if one parent is known, $16/11$ if sire and maternal grandsire is known, 2 if both parents known. With inbreeding, they can be calculated from the parental inbreeding values or taken from the A^{-1} that includes the progeny. The `!DIAG` qualifier reports the latter.

I understand some European colleagues have used the `and()` function in models similar to RAM.

It is not possible as things stand to directly estimate the genetic variances using the RAM model because the variance ratio is used in the weights and the weights are not differentiated with respect to the ratio. It can of course be done if the residual is split into the genetic and residual components and this is shown.

ASReml has benefitted from a review of the A^{-1} algorithm to speed it up. However, in a sparse matrix implementation, there is little advantage in using the RAM model per se.

I am interested in understanding what additional time costs can be identified which made the Treeplan comparison so much faster than the ASReml run, given I have identified only 20 percent of the difference within ASReml.