

## Maternal Genetic Models

In mammalian species of livestock, such as beef cattle, sheep or swine, the female provides an environment for its offspring to survive and grow in terms of protection and nourishment. Females vary in their ability to provide a suitable environment for their offspring, and this variability has a genetic basis. Offspring directly inherit an ability to grow (or survive) from both parents, and environmentally do better or poorer depending on their dam's genetic maternal ability. Maternal ability is a genetic trait expressed by the dam in the offsprings' performance, and is transmitted, like all genetic traits, from both parents. Maternal ability is only expressed by females when they have offspring (i.e. much like milk yield in dairy cows).

A model to account for maternal ability is

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{Z}_1\mathbf{a} + \mathbf{Z}_2\mathbf{m} + \mathbf{Z}_3\mathbf{p} + \mathbf{e},$$

where  $\mathbf{y}$  is the growth trait of a young animal,  $\mathbf{b}$  is a vector of fixed factors influencing growth, such as contemporary group, sex of the offspring, or age of dam,  $\mathbf{a}$  is a vector of random additive genetic effects (i.e. direct genetic effects) of the animals,  $\mathbf{m}$  is a vector of random maternal genetic (dam) effects, and  $\mathbf{p}$ , in this model, is a vector of maternal permanent environmental effects (because dams may have more than one offspring in the data).

The expectations of the random vectors,  $\mathbf{a}$ ,  $\mathbf{m}$ ,  $\mathbf{p}$ , and  $\mathbf{e}$  are all null vectors in a model without selection, and the variance-covariance structure is

$$Var \begin{pmatrix} \mathbf{a} \\ \mathbf{m} \\ \mathbf{p} \\ \mathbf{e} \end{pmatrix} = \begin{pmatrix} \mathbf{A}\sigma_a^2 & \mathbf{A}\sigma_{am} & \mathbf{0} & \mathbf{0} \\ \mathbf{A}\sigma_{am} & \mathbf{A}\sigma_m^2 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}\sigma_p^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I}\sigma_e^2 \end{pmatrix},$$

where  $\sigma_a^2$  is the additive genetic variance,  $\sigma_m^2$  is the maternal genetic variance,  $\sigma_{am}$  is the additive genetic by maternal genetic covariance, and  $\sigma_p^2$  is the maternal permanent environmental variance. Also,

$$\begin{pmatrix} \mathbf{a} \\ \mathbf{m} \end{pmatrix} \Big| \mathbf{A}, \mathbf{G} \sim N \left( \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \mathbf{G} \otimes \mathbf{A} \right),$$

where

$$\mathbf{G} = \begin{pmatrix} \sigma_a^2 & \sigma_{am} \\ \sigma_{am} & \sigma_m^2 \end{pmatrix},$$

and

$$\mathbf{p} \mid \mathbf{I}, \sigma_p^2 \sim N(\mathbf{0}, \mathbf{I}\sigma_p^2),$$

and

$$\mathbf{e} \sim N(\mathbf{0}, \mathbf{I}\sigma_e^2).$$

In this model, a female animal,  $i$ , could have its own growth record for estimating  $\hat{a}_i$ . The same female could later have offspring for estimating  $\hat{m}_i$  and  $\hat{p}_i$ , and the offspring would also

contribute towards  $\hat{a}_i$ . The maternal effects model can be more complicated if, for example, embryo transfer is practiced. Recipient dams would have maternal effects, but would not have direct genetic effects on that calf, see Schaeffer and Kennedy (1989).

To better understand this model, simulation of records is again useful. Let

$$\mathbf{G} = \begin{pmatrix} \sigma_a^2 & \sigma_{am} \\ \sigma_{am} & \sigma_m^2 \end{pmatrix} = \begin{pmatrix} 49 & -7 \\ -7 & 26 \end{pmatrix}.$$

Any positive definite matrix can be partitioned into the product of a matrix times its transpose (i.e. Cholesky decomposition), or

$$\begin{aligned} \mathbf{G} &= \mathbf{L}\mathbf{L}' \\ \mathbf{L} &= \begin{pmatrix} 7 & 0 \\ -1 & 5 \end{pmatrix}. \end{aligned}$$

Let  $\sigma_p^2 = 9$  and  $\sigma_e^2 = 81$ . This model differs from previous models in that both the additive genetic and maternal genetic effects need to be generated simultaneously because these effects are genetically correlated. The same is true for multiple trait models. Consider three animals,  $A$ ,  $B$ , and  $C$ , where  $C$  is an offspring of sire  $A$  and dam  $B$ .

## 0.1 Genetic Values

1. For  $A$ , generate a vector of two random normal deviates which will be pre-multiplied by  $\mathbf{L}$ . Animals  $A$  and  $B$  are base population animals that are unrelated to each other. Let the vector of random normal deviates be  $\mathbf{w}' = (2.533 \quad -.299)$ , then for  $A$

$$\begin{aligned} \begin{pmatrix} a_A \\ m_A \end{pmatrix} &= \mathbf{L}\mathbf{w} \\ &= \begin{pmatrix} 7 & 0 \\ -1 & 5 \end{pmatrix} \begin{pmatrix} 2.533 \\ -.299 \end{pmatrix} \\ &= \begin{pmatrix} 17.731 \\ -4.028 \end{pmatrix}. \end{aligned}$$

2. Similarly for animal  $B$ ,

$$\begin{aligned} \begin{pmatrix} a_B \\ m_B \end{pmatrix} &= \begin{pmatrix} 7 & 0 \\ -1 & 5 \end{pmatrix} \begin{pmatrix} -1.141 \\ .235 \end{pmatrix} \\ &= \begin{pmatrix} -7.987 \\ 2.316 \end{pmatrix}. \end{aligned}$$

3. Creating a progeny's true breeding value is similar to the scalar version. Take the average of the parents' true breeding values and add a random Mendelian sampling term.

$$\begin{pmatrix} a_C \\ m_C \end{pmatrix} = \frac{1}{2} \begin{pmatrix} a_A + a_B \\ m_A + m_B \end{pmatrix} + (b_{ii})^{.5} \mathbf{L}\mathbf{w}$$

$$\begin{aligned}
&= \frac{1}{2} \begin{pmatrix} 17.731 - 7.987 \\ -4.028 + 2.316 \end{pmatrix} + \left(\frac{1}{2}\right)^5 \mathbf{L} \begin{pmatrix} .275 \\ .402 \end{pmatrix} \\
&= \begin{pmatrix} 6.233 \\ .371 \end{pmatrix}.
\end{aligned}$$

All animals have both direct and maternal genetic breeding values.

## 0.2 Maternal Permanent Environmental Values

For all dams, a maternal permanent environmental effect should be generated. In this case only for animal  $B$ , multiply a random normal deviate by  $\sigma_p = 3$ , suppose it is  $-4.491$ .

## 0.3 Phenotypic Record

An observation for animal  $C$  is created by following the model equation,

$$\begin{aligned}
y &= \text{Fixed Effects} + a_C + m_B + p_B + \sigma_e * RND \\
&= 140 + 6.233 + 2.316 + (-4.491) + (9)(1.074) \\
&= 153.724.
\end{aligned}$$

The Fixed Effects contribution of 140 was arbitrarily chosen for this example. The main point is that the observation on animal  $C$  consists of the direct genetic effect of animal  $C$  plus the maternal genetic effect of the dam ( $B$ ) plus the maternal permanent environmental effect of the dam ( $B$ ) plus a residual.

# 1 HMME

To illustrate the calculations, assume the data as given in the table below.

Animal	Sire	Dam	CG	Weight
5	1	3	1	156
6	2	3	1	124
7	1	4	1	135
8	2	4	2	163
9	1	3	2	149
10	2	4	2	138

CG stands for contemporary group, the only fixed effect in this example. Assume that the appropriate variance parameters are those which were used in the simulation in the previous

section. Based on the matrix formulation of the model, the MME are

$$\begin{pmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z}_1 & \mathbf{X}'\mathbf{Z}_2 & \mathbf{X}'\mathbf{Z}_3 \\ \mathbf{Z}'_1\mathbf{X} & \mathbf{Z}'_1\mathbf{Z}_1 + \mathbf{A}^{-1}k_{11} & \mathbf{Z}'_1\mathbf{Z}_2 + \mathbf{A}^{-1}k_{12} & \mathbf{Z}'_1\mathbf{Z}_3 \\ \mathbf{Z}'_2\mathbf{X} & \mathbf{Z}'_2\mathbf{Z}_1 + \mathbf{A}^{-1}k_{12} & \mathbf{Z}'_2\mathbf{Z}_2 + \mathbf{A}^{-1}k_{22} & \mathbf{Z}'_2\mathbf{Z}_3 \\ \mathbf{Z}'_3\mathbf{X} & \mathbf{Z}'_3\mathbf{Z}_1 & \mathbf{Z}'_3\mathbf{Z}_2 & \mathbf{Z}'_3\mathbf{Z}_3 + \mathbf{I}k_{33} \end{pmatrix} \begin{pmatrix} \hat{\mathbf{b}} \\ \hat{\mathbf{a}} \\ \hat{\mathbf{m}} \\ \hat{\mathbf{p}} \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{Z}'_1\mathbf{y} \\ \mathbf{Z}'_2\mathbf{y} \\ \mathbf{Z}'_3\mathbf{y} \end{pmatrix},$$

where

$$\begin{aligned} \begin{pmatrix} k_{11} & k_{12} \\ k_{12} & k_{22} \end{pmatrix} &= \begin{pmatrix} \sigma_a^2 & \sigma_{am} \\ \sigma_{am} & \sigma_m^2 \end{pmatrix}^{-1} \sigma_e^2, \\ &= \begin{pmatrix} 49 & -7 \\ -7 & 26 \end{pmatrix}^{-1} (81), \\ &= \begin{pmatrix} 1.7192 & .4628 \\ .4628 & 3.2400 \end{pmatrix}. \end{aligned}$$

Note that these numbers are not equal to

$$\begin{pmatrix} 81/49 & 81/(-7) \\ 81/(-7) & 81/26 \end{pmatrix}.$$

Finally,  $k_{33} = \sigma_e^2/\sigma_p^2 = 81/9 = 9$ .

The matrices are

$$\mathbf{X} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{pmatrix}, \quad \mathbf{X}'\mathbf{y} = \begin{pmatrix} 415 \\ 450 \end{pmatrix},$$

$$\mathbf{Z}_1 = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

$$\mathbf{Z}_2 = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

$$\mathbf{Z}_3 = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathbf{Z}'_3\mathbf{y} = \begin{pmatrix} 429 \\ 436 \end{pmatrix}.$$

The other two right hand side matrices can be easily obtained from  $\mathbf{y}$  and  $\mathbf{Z}'_3\mathbf{y}$ . Thus, the order of the MME will be 24. The inverse of the relationship matrix is

$$\mathbf{A}^{-1} = \frac{1}{2} \begin{pmatrix} 5 & 0 & 2 & 1 & -2 & 0 & -2 & 0 & -2 & 0 \\ 0 & 5 & 1 & 2 & 0 & -2 & 0 & -2 & 0 & -2 \\ 2 & 1 & 5 & 0 & -2 & -2 & 0 & 0 & -2 & 0 \\ 1 & 2 & 0 & 5 & 0 & 0 & -2 & -2 & 0 & -2 \\ -2 & 0 & -2 & 0 & 4 & 0 & 0 & 0 & 0 & 0 \\ 0 & -2 & -2 & 0 & 0 & 4 & 0 & 0 & 0 & 0 \\ -2 & 0 & 0 & -2 & 0 & 0 & 4 & 0 & 0 & 0 \\ 0 & -2 & 0 & -2 & 0 & 0 & 0 & 4 & 0 & 0 \\ -2 & 0 & -2 & 0 & 0 & 0 & 0 & 0 & 4 & 0 \\ 0 & -2 & 0 & -2 & 0 & 0 & 0 & 0 & 0 & 4 \end{pmatrix}.$$

The solutions to the MME are

$$\hat{\mathbf{b}} = \begin{pmatrix} 137.8469 \\ 150.4864 \end{pmatrix}, \quad \hat{\mathbf{p}} = \begin{pmatrix} .0658 \\ -.0658 \end{pmatrix},$$

$$\hat{\mathbf{a}} = \begin{pmatrix} 2.3295 \\ -2.3295 \\ .1280 \\ -.1280 \\ 5.1055 \\ -4.1143 \\ .2375 \\ 2.0161 \\ .5447 \\ -3.7896 \end{pmatrix}, \quad \text{and} \quad \hat{\mathbf{m}} = \begin{pmatrix} -.3328 \\ .3328 \\ .1646 \\ -.1646 \\ -.6379 \\ .6792 \\ -.1254 \\ -.3795 \\ .0136 \\ .4499 \end{pmatrix}.$$

No correlations with true values were calculated for this small example.

The influence of the correlation between direct and maternal effects can be a matter of concern. If the correlation between direct and maternal true breeding values is negative (-0.1), for example, and if an animal has a high direct EBV based on its own growth record, then the maternal EBV could be very negative due to the correlation alone. Thus, if few of the animals with growth records have progeny too, then the relationship between direct and maternal EBVs will be strongly negative (like -0.8)(reflecting the assumed negative correlation amongst true breeding values). However, if the data are complete and animals have both their own records and those of several progeny, then the correlation between direct and maternal EBVs should more closely follow the assumed genetic correlation. This situation can also affect the correct

estimation of this genetic correlation. Estimates of this correlation in beef cattle has ranged from -0.5 to +0.5, and this mostly reflects the differences in quality (completeness) of data used.

In experimental station herds with several generations and fairly complete data, the estimates have tended to be zero or slightly positive between direct and maternal effects. On the other hand, in field data with almost no ties between growth of calves with performance of offspring as a dam, the estimates of the correlation have tended to be negative. To determine if your data are complete, create a file that has an animal's own record plus the average growth record of its progeny, to do a sort of dam-offspring phenotypic correlation. If you have 3 million records, but only 100 dam-offspring pairs, then the reliability of the estimated correlation between direct and maternal effects will be low.

## 2 Cytoplasmic Effects

Cytoplasmic effects are created by mitochondrial DNA that is passed through the oocyte to each offspring of a female. This DNA does not undergo meiosis, but is transmitted directly and wholly to each oocyte. The size of this maternal effect is not known, but several attempts were made to estimate the effect. Brian Kennedy was noted for debunking the incorrect models that were used to estimate cytoplasmic effects (J. Dairy Sci. 1986, 69:3100-3105). Kennedy suggested the use of an animal model (with additive genetic relationships), which included a female line of origin effect. That is, each animal would need to be traced back to a female in the base population with unknown parents. All such females would represent a different line of cytoplasmic effects. The variance of the line effects would need to be estimated. Most studies using this model showed very small levels of cytoplasmic effects in dairy cattle.

## 3 Embryo Transfer

Embryo transfer is used widely in dairy and beef cattle breeding. A genetically superior female is identified, and the owner desires to have many offspring from this individual. The cow is superovulated and the multiple embryos are harvested and inseminated, then at a certain stage of development, the embryos are implanted into recipient cows (of a different breed possibly). The recipients serve as a surrogate mother. When the calf is born, the maternal environment that it lives with is that of the recipient cow, and not that of its biological female parent. This can be handled in the maternal genetics effects model. The biological parent is included in the inverse of the relationship matrix, but the design matrix ( $\mathbf{Z}_2$ ) indicates the surrogate dam of that calf. The maternal genetic effects in the calf are inherited from the biological parents, but the maternal environment that influences its own growth is from the surrogate dam.

## 4 Data Structure

Data structure is very important for estimating the covariance between direct and maternal genetic effects. Mahyar Heydarpour (2006) recently completed a study of this problem with multiple trait models, and showed that estimates could be biased by poor data structure. At least two features must be present to enable proper estimation of the covariance matrices.

1. Females must appear in the data as a calf with their own early growth records, and they must appear later as the dam of other calves where they express their maternal genetic influence. There needs to be a high percentage of such ties (through a common ID number) in the data structure in order for the covariance between direct and maternal genetic effects to be estimable.
2. Sires must have daughters that appear as dams of calves.

Often the identification of a calf (registration number) is unknown in dairy and therefore, the link between the female calf growth record with that animal's progeny in later years is lost. There is a good link for a dam with all of her progeny, but not with her own growth data when she was a calf. Without this link, the covariance between direct and maternal genetic effects is often very highly negative. This can be shown to be a mathematical consequence of the data structure, if a negative correlation is used as a prior value.

If the data structure is poor, then use of a zero covariance may be better than trying to estimate the covariance

## 5 EXERCISES

Analyze the following data sets using an appropriate maternal effects model. Also, perform one iteration or one round of Gibbs sampling or EM REML.

### 5.1 Data Set 1

Weaning weights (pounds) of Hereford, female, beef calves from three contemporary groups (CG) and at different ages at weaning.

Calf	Sire	Dam	CG	Age	Weight(lb)
5	17	2	1	205	500
6	18	1	1	216	580
7	18	3	1	190	533
8	17	4	1	196	535
9	18	1	2	210	507
10	17	2	2	221	555
11	19	3	2	175	461
12	19	4	2	184	467
13	18	5	3	212	548
14	17	7	3	214	605
15	20	3	3	202	480
16	20	4	3	236	576

Assume that

$$\mathbf{G} = \begin{pmatrix} 2122 & 338 \\ 338 & 1211 \end{pmatrix},$$

$\sigma_p^2 = 476$ , and  $\sigma_e^2 = 5962$ .

What are the direct and maternal heritabilities, and direct-maternal correlation?

## 5.2 Data Set 2

This is an example of some animals being produced from embryo transfer. An animal from ET has a recipient female as the “dam” providing the maternal environment. Assume recipients are unrelated to any other animals genetically.

Length of rat pups after one week (cm).

Pup	Sire	Dam	Sex	ET	Length
5	1	3	No	M	7.3
6	1	3	Yes	F	6.5
7	2	4	No	F	7.5
8	2	4	No	M	8.7
9	2	4	Yes	M	9.4

Assume that

$$\mathbf{G} = \begin{pmatrix} 0.05 & -0.01 \\ -0.01 & 0.02 \end{pmatrix},$$

$\sigma_p^2 = 0.014$ , and  $\sigma_e^2 = 0.12$ .

What are the direct and maternal heritabilities, and direct-maternal correlation?