#### ANIMAL BREEDING NOTES

#### **CHAPTER 20**

#### ADDITIVE GENETIC GROUP MODELS

The models studied until now have all assumed that the expected value of the BLUP of the animals' genetic values was zero, i.e.,  $\mathbf{E}[\mathbf{u}] = \mathbf{0}$ . This was an assumption made when the BLUP was derived. It means that a priori all animals are considered to be equal. Also, the information used to evaluate an animal is regressed towards zero. However, in some cases  $\mathbf{E}[\mathbf{u}] \neq \mathbf{0}$ . This can happen when:

- (i) Animals come from various genetically distinct populations.
- (ii) Animals were born over a number of years within a population that underwent genetic changes due to selection over time.

These cases can be handled by adding a genetic group factor to the model. The grouping strategy, however, will differ in cases (i) and (ii). Additive group models (AGM) will be explained based on the sire-maternal grandsire model (SMM).

### **Standard Additive Genetic Group Models (SAGM)**

The SAGM can be used, for example, when animals belong to different breeds or different countries, and there is no migration across the various breeding populations. A linear model for such cases is:

$$y = Xb + Zu + e$$

$$E[y] = Xb + ZQg$$

$$var \begin{bmatrix} u \\ e \end{bmatrix} = \begin{bmatrix} A\gamma & 0 \\ 0 & R \end{bmatrix} \sigma_e^2$$

where

g = vector of genetic groups,

Q = incidence matrix relating elements of u to elements of g, and

$$\gamma = (1/4) \frac{\sigma_A^2}{\sigma_e^2}$$

The remaining terms are as defined for the **SMM**.

#### **Remarks:**

- (i) The above linear model assumes **equal** additive genetic and residual variances and covariances in all genetic populations, which may **not** be true.
- (ii) Sires are nested within genetic groups.

To include genetic groups in the above model, add and subtract ZQg, i.e.,

$$y = Xb + Z(Qg - Qg + u) + e$$

Let

$$s = u - Qg$$

Thus, the **SAGM** is:

$$y = Xb + ZQg + Zs + e$$

$$E[y] = Xb + ZQg$$

$$var \begin{bmatrix} s \\ e \end{bmatrix} = \begin{bmatrix} A\gamma & 0 \\ 0 & R \end{bmatrix} \sigma_e^2$$

The usual MME for the SAGM are:

$$\begin{bmatrix} X'R^{-1}X & X'R^{-1}ZQ & X'R^{-1}Z \\ Q'Z'R^{-1}X & Q'Z'R^{-1}ZQ & Q'Z'R^{-1}Z \\ Z'R^{-1}X & Z'R^{-1}ZQ & Z'R^{-1}Z+A^{-1}\gamma^{-1} \end{bmatrix} \begin{bmatrix} b \\ g \\ s \end{bmatrix} = \begin{bmatrix} X'R^{-1}y \\ Q'Z'R^{-1}y \\ Z'R^{-1}y \end{bmatrix}$$

The equations for the fixed effects b are usually absorbed. Thus, the MME for the **SAGM** become:

$$\begin{bmatrix} Q'Z'MZQ & Q'Z'MZ \\ Z'MZQ & Z'MZ+A^{-1}\gamma^{-1} \end{bmatrix} \begin{bmatrix} g \\ s \end{bmatrix} = \begin{bmatrix} Q'Z'My \\ Z'My \end{bmatrix}$$

where

$$M = R^{-1} - R^{-1}X(X'R^{-1}X)X'R^{-1}$$

Because sires are nested within genetic groups, the programming strategy to form the MME involves:

- (i) computing the Z'MZ matrix,
- (ii) summing the appropriate rows of Z'MZ to build the Q'Z'MZ matrix,
- (iii) summing the appropriate columns of Q'Z'MZ to form the Q'Z'MZQ matrix,
- (iv) computing Z'My,
- (v) adding appropriate rows of Z'My to obtain Q'Z'My, and
- (vi) computing the  $\{d^{ii}\}$  using Quaas' algorithm and  $A^{-1}\gamma$  using Henderson's rules.

The objectives of the **SAGM** could be:

(i) to compare populations, i.e., we want to compute:

$$\{g_i^{\circ} - g_{i'}^{\circ}\}$$

where

$$E[g_i^{\circ} - g_{i'}^{\circ}] = g_i - g_{i'}$$

(ii) to rank sires within a population, i.e., we need to compute:

$$\{\hat{\boldsymbol{s}}_{ij} - \hat{\boldsymbol{s}}_{i'j}\}$$

(iii) to compare sires across populations, i.e., we must obtain:

$$\{{u_{ij}}^{\circ} - {u_{i'j'}}^{\circ}\}$$

where

$$u_{ij}^{\circ} = g_i^{\circ} + \hat{s}_{ij}$$

$$u_{i'j'}^{\circ} = g_{i'}^{\circ} + \hat{g}_{i'j'}$$

Usually the  $g_i^{\circ}$  are computed as deviations from a base population defined as a single population or a weighted average of all of the populations involved. In such cases,

$$E[g_i^{\circ}] = g_i - g_{base}$$

The vector of predictions of additive genetic values of animals in u, i.e.,

$$\hat{\mathbf{u}} = \mathbf{Q} \ \hat{\mathbf{g}} + \hat{\mathbf{s}}$$

can be obtained directly by modifying the absorbed MME for the **SAGM** as follows:

(i) Let

$$P = \begin{bmatrix} I & 0 \\ Q & I \end{bmatrix} \quad \Rightarrow \quad P^{-1} = \begin{bmatrix} I & 0 \\ -Q & I \end{bmatrix}$$

(ii) Premultiply the LHS and the RHS of the absorbed MME by  $(P^{-1})'$ , where

$$(P^{-1})' = \begin{bmatrix} I & -Q' \\ 0 & I \end{bmatrix}$$

to obtain:

$$\begin{bmatrix} 0 & -Q'A^{-1}\gamma^{-1} \\ Z'MZQ & Z'MZ+A^{-1}\gamma^{-1} \end{bmatrix} \begin{bmatrix} g \\ s \end{bmatrix} = \begin{bmatrix} 0 \\ Z'My \end{bmatrix}$$

(iii) Insert  $P^{-1}P$  between the LHS and the vector of unknowns of the MME in (ii), i.e., postmultiply the LHS by  $P^{-1}$ , and premultiply the vector of unknowns by P. The resulting set of equations is called the **modified MME** for the **SAGM**. These equations are:

$$\begin{bmatrix} Q'A^{-1}Q\gamma^{-1} & -Q'A^{-1}\gamma^{-1} \\ -A^{-1}Q\gamma^{-1} & Z'MZ+A^{-1}\gamma^{-1} \end{bmatrix} \begin{bmatrix} g \\ u \end{bmatrix} = \begin{bmatrix} 0 \\ Z'My \end{bmatrix}$$

These equations converged much faster than the usual MME for the SAGM (Van Vleck and Dwyer, 1985). Multiple trait versions of the modified MME were first implemented at Cornell University in the mid-1980's (North East Dairy Sire Evaluation and the American Simmental Sire Evaluation).

#### **Remarks:**

(i) From the first equation of the modified MME:

$$Q'A^{-1}Q\gamma^{-1}g = Q'A^{-1}\gamma^{-1}u$$

$$\Rightarrow \qquad g = (Q'A^{-1}Q)^{-1} Q'A^{-1} u$$

i.e., group effects are linear combinations of the ui

 $\rightarrow$  the BLUP of g (as a deviation from  $g_{base}$ ) is:

$$\hat{g} = (Q'A^{-1}Q)^{-1}Q'A^{-1} \hat{u}$$

(ii) From the second equation of the modified MME:

$$-A^{-1}O\gamma^{-1}g + (Z'MZ + A^{-1}\gamma^{-1})u = Z'Mv$$

$$\hat{\mathbf{u}} = (Z'MZ + A^{-1}\gamma^{-1})^{-1}Z'My + (Z'MZ + A^{-1}\gamma^{-1})^{-1}A^{-1}Q\gamma^{-1} \hat{\mathbf{g}}$$

→ As the amount of information per animal increases, the second term will go to zero.

### Also, note that:

$$\begin{split} (Z'MZ + A^{-1}\gamma^{-1})^{-1}A^{-1}Q\gamma^{-1} \; g &= (Z'MZ + A^{-1}\gamma^{-1})^{-1}[(Z'MZ + A^{-1}\gamma^{-1}) - Z'MZ]Q \; g \\ &= Q \; g - (Z'MZ + A^{-1}\gamma^{-1})^{-1}Z'MZQ \; g \end{split}$$

$$\Rightarrow \hat{u} = (Z'MZ + A^{-1}\gamma^{-1})^{-1}Z'M(y - ZQ\hat{g}) + Q\hat{g}$$

→ As the amount of information per animal increases, the importance of the group component

decreases.

(iii) Premultiplying both sides of the MME of the **SAGM** (i.e., the unmodified MME) with b absorbed by [-I Q'] yields:

$$Q'A^{-1}\gamma^{-1} \hat{s} = 0$$

$$\Rightarrow \qquad Q'A^{-1} \hat{s} \qquad = 0$$

i.e., the weighted sum of the  $\hat{s}_{ij}$  sum to zero within genetic groups, where the weights are the elements of the inverse of the numerator relationship matrix.

If there is only one population, Q = 1, thus,

$$1'A^{-1} \hat{s} = 1'A^{-1} \hat{u} = 0$$

and when A = I,

Q'  $\hat{s} = 0$  if there are several genetic groups,

 $1'\hat{s} = 0$  if there is only one genetic group.

A proof that  $1'A^{-1} \hat{u} = 0$  in a single population (Quaas, 1986) is the following.

The MME for a single population are:

$$(Z'MZ + A^{-1}\gamma^{-1}) \hat{\mathbf{u}} = Z'My$$

where

$$Z'MZ = Z'Z - Z'X(X'X)^{-1}X'Z$$

so,

$$M = I - X(X'X)^{-1}X'$$

and

$$X'M = X' - X'X(X'X)^{-1}X' = 0$$

Notice that  $1_n$  (n = number of records in vector y) is in the column space of X. Thus,  $1_n$  can be

written as Xk,

$$\rightarrow$$
  $1_n' = k'X'$ 

$$\rightarrow$$
  $1_n'M = k'X'M = 0$ 

Also,

$$1_{n'} = Z1_{ns}$$
, where ns = number of sires

$$\rightarrow$$
 1<sub>ns</sub>'Z'MZ = 1'MZ = 0

Thus,

$$1_{ns}'(Z'MZ + A^{-1}\gamma^{-1}) \hat{\mathbf{u}} = Z'My$$
  
 $1_{ns}'A^{-1} \hat{\mathbf{u}} = 0$ 

$$\Rightarrow 1_{ns'} \hat{\mathbf{u}} = 0 \qquad \text{if } \mathbf{A} = \mathbf{I}$$

 $\rightarrow$  the  $\hat{\mathbf{u}}$  sum to zero "within" the overall mean of a single population.

The **SAGM** can also be used to account for differences in the mean genetic values of animals born in different years, generations, etc. However, a more realistic model is the accumulated group model.

# **Accumulated Additive Genetic Group Models (AAGM)**

The **AAGM** was introduced by Thompson (1979). The **AAGM** for a sire model, using scalar notation, is:

$$y_{ij} \qquad = \quad x_{ij}{'}b + \sum_k a_{ik}g_k + s_i + e_{ij}$$

$$\mathrm{E}[y_{ij}] = x_{ij}'b + \sum_{k} a_{ik}g_k$$

$$\operatorname{var} \begin{bmatrix} s_{i} \\ e_{ij} \end{bmatrix} = \begin{bmatrix} a_{s_{i}s_{i}} \gamma & 0 \\ 0 & I \end{bmatrix} \sigma_{e}^{2}$$

where

b = vector of fixed effects,

 $a_{ik}$  = additive genetic relationship between animals i and k,

 $g_k$  = additive genetic group which animal k belongs to, and

 $\sum$  = summation over all ancestors of animal i.

Thus, instead of having one group effect for the i<sup>th</sup> animal as in the **SAGM**, the **AAGM** contains a weighted sum of the genetic group effects of all the ancestors of an animal in addition to that animal's group, e.g., for the sire model above:

$$\sum_{k} a_{ik} g_{k} = \text{sire genetic group}$$

+

1/2 grandsire genetic group

+

1/4 great grandsire genetic group

+

:

The additive genetic value of sire i is:

$$u_{s_i} = \sum_{k} a_{ik} g_k + s_i$$

# **Derivation of the AAGM using matrix notation**

Let a linear model for vector y be:

$$y = Xb + Zu + e$$

$$E[y] = Xb + Z E[u]$$

$$\operatorname{var} \left[ \begin{array}{c} u \\ e \end{array} \right] = \left[ \begin{array}{cc} A\gamma & 0 \\ 0 & R \end{array} \right] \sigma_e^2$$

But

$$A = TDT'$$

where

$$D = var (\phi)$$

$$T = \begin{cases} (I - \frac{1}{2}P)^{-1} & \text{for models including sires and dams} \\ (I - \frac{1}{2}P_s - \frac{1}{4}P_{mgs})^{-1} & \text{for models including sires and mgs} \end{cases}$$

Thus, the following equivalent **AAGM** model can now be considered:

$$y = Xb + Z(TT^{-1})u + e$$

$$y = Xb + ZT(T^{-1}u) + e$$

$$y = Xb + ZT\phi + e$$

$$E[y] = Xb + ZT E[\phi]$$

$$\operatorname{var} \begin{bmatrix} \phi \\ e \end{bmatrix} = \begin{bmatrix} D\gamma & 0 \\ 0 & R \end{bmatrix} \sigma_{e}^{2}$$

The MME for this equivalent **AAGM** model are:

$$\begin{bmatrix} X'R^{-1}X & X'R^{-1}ZT \\ T'Z'R^{-1}X & T'Z'R^{-1}ZT + D^{-1}\gamma^{-1} \end{bmatrix} \begin{bmatrix} b \\ \phi \end{bmatrix} = \begin{bmatrix} X'R^{-1}y \\ T'Z'R^{-1}y \end{bmatrix}$$

# Composition of the accumulated groups

Recall that for the **SMM**,

$$\phi = (I - \frac{1}{2}P_{s} - \frac{1}{4}P_{mgs}) u$$
$$= T^{-1} u$$

Consequently, the accumulated groups will include all the genetic effects contained in  $\varphi$  whose expectation is different from zero. Thus,

(i) if the sire and the mgs of animal i are identified,

$$\begin{split} \varphi_i &= \ ^{1\!\!/_2} u_{s_i} + ^{1\!\!/_2} u_{d_i} + \dot{\varphi}_i - ^{1\!\!/_2} u_{s_i} - ^{1\!\!/_4} u_{mgs_i} \\ &= \ ^{1\!\!/_2} \left( u_{d_i} - ^{1\!\!/_2} u_{mgs_i} \right) + \dot{\varphi}_i \\ &= \ ^{1\!\!/_4} u_{mgs_i} + ^{1\!\!/_4} u_{mgd_i} + ^{1\!\!/_2} \varphi_{d_i} - ^{1\!\!/_4} u_{mgs_i} + \dot{\varphi}_i \\ &= \ ^{1\!\!/_4} u_{mgd_i} + ^{1\!\!/_2} \hat{\varphi}_{d_i} + \dot{\varphi}_i \\ &= \ ^{1\!\!/_4} u_{mgd_i} \right] \quad \text{Additive genetic values of bull mgds} \\ &+ \\ &E \left[ \dot{\varphi}_i \right] = \left\{ \begin{array}{ll} E \left[ ^{1\!\!/_4} u_{mgd_i} \right] & \text{Selection of bull dams} \\ &+ \\ &E \left[ \dot{\varphi}_i \right] & \text{Selection of the bulls themselves} \end{array} \right. \end{split}$$

(ii) if the sire of animal i is identified only,

$$\begin{array}{rcl} \varphi_{i} & = & \sqrt[1]{2} \, u_{s_{i}} + \sqrt[1]{2} \, u_{d_{i}} + \dot{\varphi}_{i} - \sqrt[1]{2} \, u_{s_{i}} \\ \\ & = & \sqrt[1]{2} \, u_{d_{i}} + \dot{\varphi}_{i} \\ \\ & E \left[ \dot{\varphi}_{i} \right] = & E \left[ \sqrt[1]{2} \, u_{d_{i}} \right] + E \left[ \dot{\varphi}_{i} \right] \end{array}$$

$$E\left[ \begin{matrix} | 4 u_{mgs_i} | \\ + \\ | E[ \rlap/ 4 u_{mgd_i} | \\ + \\ | E[ \rlap/ 4 u_{mgd_i} | \\ + \\ | E[ \rlap/ 2 \dot{\phi}_{d_i} | \\ + \\ | E[ \rlap/ 2 \dot{\phi}_{d_i} | \\ + \\ | E[ \dot{\phi}_i ] \qquad \text{Selection of bull dams} \\ + \\ | E[ \dot{\phi}_i ] \qquad \text{Selection of the bulls themselves} \\ \label{eq:energy}$$

(iii) if the maternal grandsire of animal i is identified only,

$$\begin{split} \varphi_i &= \ \ \frac{1}{2} \, u_{s_i} + \frac{1}{2} \, u_{d_i} + \dot{\varphi}_i - \frac{1}{4} \, u_{mgs_i} \\ &= \ \ \frac{1}{2} \, u_{s_i} + \frac{1}{2} \, \left( u_{d_i} - \frac{1}{2} \, u_{mgs_i} \right) + \dot{\varphi}_i \\ &= \ \ \frac{1}{2} \, u_{s_i} + \frac{1}{4} \, u_{mgd_i} + \frac{1}{2} \, \varphi_{d_i} + \dot{\varphi}_i \\ &= \ \ \left[ E \left[ \frac{1}{2} \, u_{s_i} \right] \quad \text{Additive genetic values of bull sires} \right. \\ &+ \\ &E \left[ \frac{1}{4} \, u_{mgd_i} \right] \quad \text{Additive genetic values of bull mgds} \\ &+ \\ &E \left[ \frac{1}{2} \, \dot{\phi}_{d_i} \right] \quad \text{Selection of bull dams} \\ &+ \\ &E \left[ \dot{\phi}_i \right] \quad \text{Selection of the bulls themselves} \end{split}$$

(iv) if neither the sire nor the mgs of animal i are identified,

$$\phi_i = \frac{1}{2} u_{s_i} + \frac{1}{2} u_{d_i} + \dot{\phi}_i$$

$$E\left[\phi_{i}\right] = \begin{cases} E\left[\frac{1}{2}u_{s_{i}}\right] & \text{Additive genetic values of bull sires} \\ + \\ E\left[\frac{1}{2}u_{d_{i}}\right] & \text{Additive genetic value of bull dams} \\ + \\ E\left[\dot{\phi}_{i}\right] & \text{Selection of the bulls themselves} \end{cases}$$

## Criteria used to construct groups

(a) Ancestors identified, e.g., in the sire-mgs case, there are four groups:

group 1 = sire and mgs identified, group 2 = sire identified only, group 3 = mgs identified only, and

group 4 = neither sire nor mgs identified.

(b) In addition to the grouping criterion (a) we can add time, stud, region, breed, etc. For instance, if there were only one more criterion: time and only 2 generations, there would be eight groups in all.

In terms of the **AAGM** model we have that:

$$E [\phi] = Q_0 g$$

where

 $Q_o = \text{matrix that relates the } \phi_i$  to genetic groups g according to criteria (a) and (b) above.

Thus, the complete specification of the equivalent **AAGM** model is:

$$y = Xb + ZT\phi + e$$
 
$$E[y] = Xb + ZTQ_{o}g$$
 
$$var \begin{bmatrix} \phi \\ e \end{bmatrix} = \begin{bmatrix} D\gamma & 0 \\ 0 & R \end{bmatrix} \sigma_{e}^{2}$$

But,  $u = T \varphi$ . Thus,

$$y = Xb + Zu + e$$

$$E[y] = Xb + ZTQ_{0}g$$

$$var\begin{bmatrix} u \\ 0 \end{bmatrix} = \begin{bmatrix} A\gamma & 0 \\ 0 & R \end{bmatrix} \sigma_{e}^{2}$$

This equivalent **AAGM** model can be rewritten as follows:

$$\begin{array}{rcl} y & = & Xb + Z(TQ_og - TQ_og + u) + e \\ \\ y & = & Xb + ZTQ_og + (u - TQ_og) + e \\ \\ y & = & Xb + ZTQ_og + s + e, \text{ where } s = u - TQ_og \\ \\ E[y] & = & Xb + ZTQ_og \\ \\ var \begin{bmatrix} s \\ e \end{bmatrix} & = & \begin{bmatrix} A\gamma & 0 \\ 0 & R \end{bmatrix} \sigma_e^2 \end{array}$$

with MME:

$$\begin{bmatrix} X'R^{-1}X & X'R^{-1}ZTQ_0 & X'R^{-1}Z \\ Q_0'T'Z'R^{-1}X & Q_0'T'Z'R^{-1}ZTQ_0 & Q_0'T'Z'R^{-1}Z \\ Z'R^{-1}X & Z'R^{-1}ZTQ_0 & Z'R^{-1}Z + A^{-1}\gamma^{-1} \end{bmatrix} \begin{bmatrix} b \\ g \\ s \end{bmatrix} = \begin{bmatrix} X'R^{-1}y \\ Q_0'T'Z'R^{-1}y \\ Z'R^{-1}y \end{bmatrix}$$

These equations for the **AAGM** are identical to those of the **SAGM**, with  $TQ_0$  substituted for Q, and using the fact that A = TDT' and  $A^{-1} = T^{-1}D^{-1}T^{-1}$ . The transformed MME to compute vectors u and g instead of vectors s and g can be obtained using a similar procedure to the one utilized with the **SAGM**.

Thus,

(i) Let

$$P = \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & TQ_0 & I \end{bmatrix}, P^{-1} = \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & -TQ_0 & I \end{bmatrix}, and (P^{-1})' = \begin{bmatrix} I & 0 & 0 \\ 0 & I & -Q_0'T' \\ 0 & 0 & I \end{bmatrix}$$

(ii) Premultiply the LHS and the RHS of the **AAGM** MME by  $(P^{-1})'$ , and then postmultiply the LHS of the resulting MME by  $P^{-1}$  and premultiply the vector of unknowns by P.

The resulting MME are the modified MME for the **AAGM**:

$$\begin{bmatrix} X \, {}^{\prime} R^{-1} X & 0 & X \, {}^{\prime} R^{-1} Z \\ 0 & Q_{o} \, {}^{\prime} D^{-1} Q_{o} \gamma^{-1} & -Q_{o} \, {}^{\prime} D^{-1} T^{-1} \gamma^{-1} \\ Z \, {}^{\prime} R^{-1} X & -T^{-T} D^{-1} Q_{o} \gamma^{-1} & Z \, {}^{\prime} R^{-1} Z + A^{-1} \gamma^{-1} \end{bmatrix} \begin{bmatrix} b \\ g \\ u \end{bmatrix} = \begin{bmatrix} X \, {}^{\prime} R^{-1} y \\ 0 \\ Z \, {}^{\prime} R^{-1} y \end{bmatrix}$$

The MME for the **AAGM** are simple to program (Quaas, 1986). Usually the fixed effects are absorbed, so the modified MME for the **AAGM** become:

$$\begin{bmatrix} Q_o ' D^{-1} Q_o \gamma^{-1} & -Q_o ' D^{-1} T^{-1} \gamma^{-1} \\ -T^{-T} D^{-1} Q_o \gamma^{-1} & Z' M Z + A^{-1} \gamma^{-1} \end{bmatrix} \begin{bmatrix} g \\ u \end{bmatrix} = \begin{bmatrix} 0 \\ Z' M y \end{bmatrix}$$

where

$$M = R^{-1} - R^{-1}X(X'R^{-1}X)^{-1}X'R^{-1}$$
.

### These MME can be built as follows:

- (i) Compute and half-store the matrix Z'MZ,
- (ii) Compute and store the vector Z'My, and
- (iii) Compute and half-store the matrix.

$$\begin{bmatrix} A^{-1}\gamma^{-1} & \vdots & -T^{-T}D^{-1}Q_{o}\gamma^{-1} \\ ----- & \vdots & ----- \\ -Q_{o}, D^{-1}T^{-1}\gamma^{-1} & \vdots & Q_{o}, D^{-1}Q_{o}\gamma^{-1} \end{bmatrix}$$

The contributions of a bull to the submatrices in (iii) are:

mgs of bull	1/16 d <sub>ii</sub> <sup>-1</sup>			
sire of bull	1/8 d <sub>ii</sub> <sup>-1</sup>	$1/4 d_{ii}^{-1}$		
bull	-1/4 d <sub>ii</sub> <sup>-1</sup>	$-1/2 d_{ii}^{-1}$	$\mathbf{d}_{\mathrm{ii}}^{-1}$	
bull group	$1/4 d_{ii}^{-1}$	$1/2 d_{ii}^{-1}$	$-d_{ii}^{-1}$	$d_{ii}^{-1}$
	mgs of bull	sire of bull	bull	bull group

# Relationship between the SAGM and the AAGM

The **SAGM** is;

$$y = Xb + Zu + e$$

$$E[y] = Xb + ZQg, \qquad Q = \text{incidence matrix relating sires to groups}$$

$$var \begin{bmatrix} u \\ e \end{bmatrix} = \begin{bmatrix} A\gamma & 0 \\ 0 & R \end{bmatrix} \sigma_e^2$$

The **AAGM** is:

$$y = Xb + Zu + e$$

 $E[y] = Xb + ZTQ_0g$ ,  $Q_0 =$  incidence matrix relating residual genetic terms  $(\phi_{ij})$  to

groups

$$var \begin{bmatrix} u \\ e \end{bmatrix} = \begin{bmatrix} A\gamma & 0 \\ 0 & R \end{bmatrix} \sigma_e^2$$

In the **SAGM**, bulls are placed into groups. Thus,

$$E[u_{ij}] = g_i$$

On the other hand, residual genetic terms ( $\phi_{ij}$ ) are grouped in the AAGM. Thus,

$$E[\phi_{ij}] \ = \ g_i$$

But

$$\varphi = T^{-1} u$$

$$\rightarrow$$
  $u = T \varphi$ 

$$\ \, \Rightarrow \ \, E[u_{ij}] \quad = \quad \sum_i t_{ij} \, g_i$$

where

 $t_{ij}$  = elements of the  $i^{th}$  row of t, and

 $T = (I - \frac{1}{2}P)^{-1}$  if sires and dams are included in the model

 $T = (I - \frac{1}{2}P_s - \frac{1}{4}P_{mgs})^{-1}$  if sires and mgs' are included in the model

Let

$$g_i^* = \sum_i t_{ij} g_i$$

Then, for the **SMM** we have that the elements of T are:

$$t_{ij} = \delta_{s_i} \left( \frac{1}{2} t_{s_i, j} \right) + \delta_{mgs_i} \left( \frac{1}{4} t_{mgs_i, j} \right)$$

$$t_{ii} = 1$$

$$\Rightarrow g_{i}^{*} = (1)g_{i} + \delta_{s_{ij}} \left| \frac{1}{2} \sum_{k} t_{s_{ij},k} g_{k} \right| + \delta_{mgs_{ij}} \left| \frac{1}{4} \sum_{k} t_{mgs_{ij},k} g_{k} \right|$$

$$\Rightarrow \qquad g_{i}^{\,*} \; = \quad g_{i}^{\,} + \delta_{s_{ij}} \left( \! \frac{1}{2} \, g_{s_{ij}}^{\,*} \right) \! + \delta_{mgs_{ij}} \left( \! \frac{1}{4} \, g_{mgs_{ij}}^{\,*} \right)$$

$$\rightarrow E[u_{ij}] = g_i + \delta_{s_{ii}} \left( \frac{1}{2} E[u_{s_{ii}}] \right) + \delta_{mgs_{ii}} \left( \frac{1}{4} E[u_{mgs_{ii}}] \right)$$

$$\Rightarrow \qquad g_i = E[u_{ij}] - \delta_{s_{ij}} \left( \frac{1}{2} E[u_{s_{ij}}] \right) - \delta_{mgs_{ij}} \left( \frac{1}{4} E[u_{mgs_{ij}}] \right)$$

or

$$E\left[\phi_{ij}\right] = E\left[u_{ij}\right] - \delta_{s_{ij}} \left(\frac{1}{2} E\left[u_{s_{ij}}\right]\right) - \delta_{mgs_{ij}} \left(\frac{1}{4} E\left[u_{mgs_{ij}}\right]\right)$$

→ the SAGM and the AAGM will be equivalent only if the pedigrees of all bulls are the same and all bulls belong to the same group. If so, the ranking of bulls will be the same in the SAGM and in the **AAGM**.

#### **Remarks:**

- (i) The composition of the residual genetic terms  $\phi_{ij}$  will depend on the ancestors identified in an individual.
- (ii) The expected values of the  $\phi_{ij}$  will be determined by the expected values of the additive genetic values of the ancestors of individual ij included in  $\phi_{ij}$  and by the expected values of the Mendelian sampling terms (i.e., the  $\dot{\phi}$ ) belonging to these ancestors and to animal ij.
- (iii) The expected values of the Mendelian sampling terms (i.e., the  $\dot{\phi}$ ) reflect the occurrence of selection in a group of animals.

Let

 $g_i = genetic group value of the animals in the i<sup>th</sup> genetic group, where the i<sup>th</sup> genetic group may be defined to be, for instance, all animals belonging to the i<sup>th</sup> (time × breed) subclass. Because <math>g_i$  is not estimable,  $(g_i - g_{base})$  is used to obtain a unique  $\hat{g}$  in bull evaluations. Notice that if there is only one (time × breed) subclass  $(g_i - g_{base}) = (g_{base} - g_{base}) = 0$ .

Then,

(a) if the i<sup>th</sup> genetic group is **unselected**,

$$E[\dot{\phi}_{ii}] = g_i$$

where

ij = j<sup>th</sup> animal from the i<sup>th</sup> genetic group

(b) if there was selection of animals in the i<sup>th</sup> genetic group,

$$E[\dot{\phi}_{ij}] = g_i + (g_{selected_i} - g_i)$$

$$E[\dot{\phi}_{ij}] = g_i + \Delta_i$$

where

 $g_{\text{selected}_i}$  = additive genetic value of the selected group of animals from the  $i^{\text{th}}$  (time  $\times$  breed) subclass, and

 $\Delta_i$  = genetic selection differential for the i<sup>th</sup> genetic group.

## **Example:**

If bulls of breed i are mated to dams of breed i', the expected value of the  $\dot{\phi}$  of their progeny is equal to:

$$(1) \qquad E[\dot{\phi}_{ij}] = g_i$$

$$= \frac{1}{2} E[\dot{\phi}_{ij}] + \frac{1}{2} E[\dot{\phi}_{i'j'}]$$

$$\downarrow \qquad \qquad \downarrow$$

$$sires \qquad dams$$

$$= E[\epsilon_{ij}] + E[\epsilon_{i'j'}]$$

$$= \frac{1}{2} g_i + \frac{1}{2} g_{i'}$$

if both parents of animal j are identified and the sire and the dam genetic groups are **unselected**,

(2) 
$$\begin{split} E[\dot{\phi}_{ij}] &= E\left[\epsilon_{ij}\right] + E\left[\epsilon_{i'j'}\right] \\ &= \frac{1}{2}\left(g_{i} + \Delta_{i}\right) + \frac{1}{2}\left(g_{i'} + \Delta_{i'}\right) \end{split}$$

if both parents of the animal j are known and both parental genetic groups are selected.

#### Understanding the solutions obtained using the AAGM

Consider the MME without absorbing the b's, i.e.,

$$\begin{bmatrix} X'R^{-1}X & X'R^{-1}Z & 0 \\ Z'R^{-1}X & Z'R^{-1}Z + A^{-1}\gamma^{-1} & -T^{-T}D^{-1}Q_{o}\gamma^{-1} \\ 0 & -Q_{o}'D^{-1}T^{-1}\gamma^{-1} & Q_{o}'D^{-1}Q_{o}\gamma^{-1} \end{bmatrix} \begin{bmatrix} b \\ u \\ g \end{bmatrix} = \begin{bmatrix} X'R^{-1}y \\ Z'R^{-1}y \\ 0 \end{bmatrix}$$

### (1) Group solutions

From the 3<sup>rd</sup> equation of the MME:

$$\begin{array}{rclcrcl} Q_{o}{'}D^{-1}T^{-1}\gamma^{-1} & \hat{u} & = & Q_{o}{'}D^{-1}Q_{o}\gamma^{-1} & \hat{g} \\ \\ \Rightarrow & & \hat{g} & = & (Q_{o}{'}D^{-1}Q_{o})^{-1}Q_{o}{'}D^{-1}T^{-1} & \hat{u} \end{array}$$

where

$$\begin{array}{lll} T^{-1} &=& (I-1/\!\!{}_2P) & \text{for models with sires and dams: } \textbf{AM, RAM, SDM} \\ \\ T^{-1} &=& (I-1/\!\!{}_2P_s-1/\!\!{}_4P_{mgs}) & \text{for the SMM} \\ \\ T^{-1} &=& (I-1/\!\!{}_2P_s) & \text{for the SM} \\ \\ & \Rightarrow & \hat{\phi} &=& T^{-1} \; \hat{u} &=& \begin{cases} \{\hat{u}_{ij}-\delta_{s_{ij}}\left(^{1}\!\!/_2\;\hat{u}_{s_{ij}}\right)\!\!-\delta_{d_{ij}}\left(^{1}\!\!/_2\;\hat{u}_{d_{ij}}\right)\!\} & \text{for the AM,RAM,SDM} \\ \\ & \{\hat{u}_{ij}-\delta_{s_{ij}}\left(^{1}\!\!/_2\;\hat{u}_{s_{ij}}\right)\!\!-\delta_{mgs_{ij}}\left(^{1}\!\!/_4\;\hat{u}_{mgs_{ij}}\right)\!\} & \text{for the SMM} \end{cases} \\ & & \{\hat{u}_{ij}-\delta_{s_{ij}}\left(^{1}\!\!/_2\;\hat{u}_{s_{ij}}\right)\!\!\} & \text{for the SMM} \end{cases}$$

 $D^{-1}$  = diagonal matrix of group weights, i.e., the  $d_{ij,ij}^{-1}$ , which were defined in the chapters dealing with A and  $A^{-1}$ . Notice that the d of all the animals in a group is the same because of the grouping strategy based on ancestors identified.

 $Q_{o}{}' = \text{incidence matrix that sums the terms } d_{ij,ij}^{-1} \ \hat{\phi}_{ij} \text{ within a group,}$ 

 $(Q_o'D^{-1}Q_o)^{-1}$  = diagonal matrix with the weights for each group.

Notice that because D<sup>-1</sup> is diagonal,

$$(i) \ (Q_o'D^{-1}Q_o)^{-1} \ = \ \left\{\frac{1}{n_i} \ d_{ij,ij}\right\}$$

where

 $n_i$  = number of animals in group i, and

$$\begin{array}{lcl} (ii) \ Q_o{'}D^{-1}T^{-1} \ \hat{u} & = & D^{-1}Q_o{'}T^{-1} \ \hat{u} \\ \\ \Rightarrow & \hat{g} & = & (Q_o{'}Q_o)^{-1}DD^{-1}Q_o{'}T^{-1} \ \hat{u} \\ \\ & = & (Q_o{'}Q_o)^{-1}Q_o{'}T^{-1} \ \hat{u} \\ \\ \Rightarrow & \hat{g}_i & = & \frac{1}{n_i} \sum_{j=1}^{n_i} \left[ \hat{u}_{ij} - \delta_{s_{ij}} \left( \! \frac{1}{2} \hat{u}_{s_{ij}} \right) \! - \delta_{d_{ij}} \left( \! \frac{1}{2} \hat{u}_{d_{ij}} \right) \! - \delta_{mgs_{ij}} \left( \! \frac{1}{2} \hat{u}_{mgs_{ij}} \right) \right] \\ \\ & = & \frac{1}{n_i} \sum_{i=1}^{n_i} \left( \hat{u}_{ij} - PI_{ij} \right) \end{array}$$

where

 $\delta_{k_{ij}}$  =1 if ancestor k of animal ij is known and used in the model and 0 otherwise, and

 $PI_{ij} = pedigree index of animal ij.$ 

# (2) Prediction of additive genetic values

From the second equation of the MME:

$$\begin{split} Z'R^{-1}X\;b^\circ + (Z'R^{-1}Z + A^{-1}\gamma^{-1})\;\hat{u}\; - T^{-T}D^{-1}Q_o\gamma^{-1}\;\hat{g} &=\; Z'R^{-1}y\\ \\ (Z'R^{-1}Z + A^{-1}\gamma^{-1})\;\hat{u}\; &=\; Z'R^{-1}(y - Xb^\circ) + T^{-T}D^{-1}Q_o\gamma^{-1}\;\hat{g}\\ \\ \Rightarrow \;\;\hat{u}\; &=\; (Z'R^{-1}Z + A^{-1}\gamma^{-1})^{-1}Z'R^{-1}(y - Xb^\circ) + (Z'R^{-1}Z + A^{-1}\gamma^{-1})^{-1}T^{-T}D^{-1}Q_o\gamma^{-1}\;\hat{g} \end{split}$$

**As an example**, consider the i<sup>th</sup> nonparent in a **SMM**. In this case, A<sup>-1</sup> has at most 3 nonzero values

in the i<sup>th</sup> row, i.e.,

$$\begin{array}{ccc} -\frac{1}{4}\,d_{ii}^{-1} & & mgs_i \times i \\ \\ -\frac{1}{2}\,d_{ii}^{-1} & & s_i \times i \end{array}$$
 
$$d_{ii}^{-1} & i \times i$$

Also,

$$\begin{split} T^{-T}D^{-1}Q_{o}\gamma^{-1}\;\hat{g} &= \begin{bmatrix} T^{pp} & T^{pn} \\ 0 & I_{n} \end{bmatrix} \begin{bmatrix} D_{p}^{-1} & 0 \\ 0 & D_{n}^{-1} \end{bmatrix} \begin{bmatrix} Q_{p} & 0 \\ 0 & Q_{n} \end{bmatrix} \begin{bmatrix} \hat{g}_{p} \\ \hat{g}_{n} \end{bmatrix} \gamma^{-1} \\ &= \begin{bmatrix} T^{pp}D_{p}^{-1} & T^{pn}D_{n}^{-1} \\ 0 & D_{n}^{-1} \end{bmatrix} \begin{bmatrix} Q_{p} & 0 \\ 0 & Q_{n} \end{bmatrix} \begin{bmatrix} \hat{g}_{p} \\ \hat{g}_{n} \end{bmatrix} \gamma^{-1} \\ &= \begin{bmatrix} T^{pp}D_{p}^{-1}Q_{p}\;\hat{g}_{p} + T^{pn}D_{n}^{-1}Q_{n}\;\hat{g}_{n} \\ &D_{n}^{-1}Q_{n}\;\hat{g}_{n} \end{bmatrix} \end{split}$$

Thus,

$$\begin{split} & \left( \, r_{i \bullet}^* + \gamma^{-1} \, d_{ii}^{-1} \, \right) \, \hat{u}_i - \gamma^{-1} \, {}^{1}\!/_{\!\!\!4} \, d_{ii}^{-1} \, \hat{u}_{mgs_i} - \gamma^{-1} \, {}^{1}\!/_{\!\!\!2} \, d_{ii}^{-1} \, \hat{u}_{s_i} \, = \, y_{i \bullet}^* + \gamma^{-1} \, d_{ii}^{-1} \, \hat{g}_i \\ & \left( \, r_{i \bullet}^* + \gamma^{-1} \, d_{ii}^{-1} \, \right) \hat{u}_i \, = \, y_{i_-}^* + \gamma^{-1} \, d_{ii}^{-1} \left( \, {}^{1}\!/_{\!\!\!2} \, \hat{u}_{s_i} + \, {}^{1}\!/_{\!\!\!4} \, \hat{u}_{mgs_i} \, \right) + \gamma^{-1} \, d_{ii}^{-1} \, \hat{g}_i \\ & \hat{u}_i \, = \, \frac{r_{i \bullet}^*}{r_{i \bullet}^* + \gamma^{-1} d_{ii}^{-1}} \, \overline{y}_{i \bullet}^* + \, \frac{\gamma^{-1} \, d_{ii}^{-1}}{r_{i \bullet}^* + \gamma^{-1} d_{ii}^{-1}} \, \left( \, \hat{g}_i + {}^{1}\!/_{\!\!\!2} \, \hat{u}_{s_i} + {}^{1}\!/_{\!\!\!4} \, \hat{u}_{mgs_i} \, \right) \end{split}$$

Numerical examples for the SAGM and the AAGM

Consider a **SMM** using the example used for the direct effects models.

(1) The **SAGM** is:

$$\begin{bmatrix} 289 \\ 285 \\ 265 \\ 290 \\ 288 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} + \begin{bmatrix} 0 & 1.5 & 0 & 0 \\ 0 & 0.5 & 1 & 0 \\ 0 & 0 & 1.5 & 0 \\ 0 & 0 & 0 & 1.5 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} g_1 \\ g_2 \end{bmatrix} + \begin{bmatrix} 0 & 1.5 & 0 & 0 \\ 0 & 0.5 & 1 & 0 \\ 0 & 0 & 1.5 & 0 \\ 0 & 0 & 0 & 1.5 \end{bmatrix} \begin{bmatrix} s_1 \\ s_3 \\ s_4 \\ s_5 \end{bmatrix} + e$$

The MME and the vector of solutions for the **SAGM** are:

$$\begin{bmatrix} X'R_2^{-1}X & X'R_2^{-1}Z & X'R_2^{-1}ZQ \\ Z'R_2^{-1}X & Z'R_2^{-1}Z + A^{-1}\!\!\left(\frac{4}{\alpha}\right) & Z'R_2^{-1}ZQ \\ Q'Z'R_2^{-1}X & Q'Z'R_2^{-1}Z & Q'Z'R_2^{-1}ZQ \end{bmatrix}\!\!\begin{bmatrix} b \\ s \\ g \end{bmatrix} = \begin{bmatrix} X'R_2^{-1}y \\ Z'R_2^{-1}y \\ Q'Z'R_2^{-1}y \end{bmatrix}$$

$b_1$ °		56.9899
$b_2$ °		45.6285
$\hat{\mathbf{s}}_1$		0.0075
Ŝ <sub>3</sub>	=	0.0263
Ŝ4		-0.7104
$\hat{\mathbf{S}}_5$		0.6954
$g_1^{\circ}$		0
$\lfloor g_2^{\circ} \rfloor$		[153.9236]

## (2) The AAGM is:

$$\begin{bmatrix} 289 \\ 285 \\ 265 \\ 290 \\ 288 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} + \begin{bmatrix} 0 & 1.5 & 0 & 0 \\ 0 & 0.5 & 1 & 0 \\ 0 & 0 & 1.5 & 0 \\ 0 & 0 & 0 & 1.5 \\ 0 & 0 & 0 & 1.5 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ -0.75 & 1 & 0 & 0 \\ -0.5 & 0 & 1 & 0 \\ -0.5 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} g_1 \\ 0 & 1 \\ g_2 \end{bmatrix} + \begin{bmatrix} 0 & 1.5 & 0 & 0 \\ 0 & 0.5 & 1 & 0 \\ 0 & 0 & 1.5 & 0 \\ 0 & 0 & 0 & 1.5 \\ 0 & 0 & 0 & 1.5 \end{bmatrix} \begin{bmatrix} s_1 \\ s_3 \\ s_4 \end{bmatrix} + e$$

The transformed MME (i.e., in terms of g, and  $u = TQ_0 + s$ ) and the vector of solutions for the **AAGM** are:

$$\begin{bmatrix} X'R_{2}^{-1}X & X'R_{2}^{-1}Z & 0 \\ Z'R_{2}^{-1}X & Z'R_{2}^{-1}Z + A^{-1}\left(\frac{4}{\alpha}\right) & -T^{-T}D^{-1}Q_{o}\left(\frac{4}{\alpha}\right) \\ 0 & -Q_{o}'D^{-1}T^{-1}\left(\frac{4}{\alpha}\right) & Q_{o}'D^{-1}Q_{o}\left(\frac{4}{\alpha}\right) \end{bmatrix} \begin{bmatrix} b \\ u \\ g \end{bmatrix} = \begin{bmatrix} X'R_{2}^{-1}y \\ Z'R_{2}^{-1}y \\ 0 \end{bmatrix}$$

$\left[\begin{array}{c}b_1^{\circ}\end{array}\right]$		[ 138.1434 ]
$b_2$ °		127.1044
$\hat{\mathbf{u}}_1$		2.0655
û <sub>3</sub>		100.1134
û <sub>4</sub>		98.8872
û <sub>5</sub>		100.3069
$g_1^{\circ}$		2.0655
$\left[ g_{2}^{\circ} \right]$		98.5643

### Remark

The **AAGM** described here places all animals into genetic groups. A different strategy is to assign only base animals (i.e., animals without identified parents) to genetic groups. For a comprehensive description of this grouping strategy, please refer to Quaas (1988).

#### References

- Quaas, R. L. 1986. Personal Communication. Animal Science 720. Cornell University, Ithaca, N.Y.
- Quaas, R. L. 1988. Additive genetic model with groups and relationships. J. Dairy Sci. 71 (Suppl. 2):1338-1345.
- Thompson, R. 1979. Sire evaluation. Biometrics 35:339-353.
- Van Vleck, L. D. and D. J. Dwyer. 1984. Comparison of iterative procedures for solving equations for sire evaluation. J. Dairy Sci. 68:1006-1014.