

# GENE638 - Exam 2

Daniel Osorio - dcosoriorh@tamu.edu

Department of Veterinary Integrative Biosciences

Texas A&M University

Multiple trait analyses are used to estimate the genetic correlations between traits and to increase the precision of genetic predictions when animals have relatively few observations. They can also be used to model the effects of selection, e.g., when only the genetically superior animals are permitted to go on and have second and third records, etc.

Suppose that the models for trait 1 and trait 2 can be written:

$$\begin{aligned}\underline{y_1} &= X_1\underline{\beta_1} + Z_1\underline{u_1} + \underline{e_1} \\ \underline{y_2} &= X_2\underline{\beta_2} + Z_2\underline{u_2} + \underline{e_2}\end{aligned}\tag{1}$$

Where:

$\underline{y_1}$  and  $\underline{y_2}$  are observations on traits 1 and 2 respectively.  $X_1, X_2, Z_1$  and  $Z_2$  are incidence matrices relating fixed and random effects to observations on each trait.  $\underline{\beta_1}$  and  $\underline{\beta_2}$  are fixed effects for traits 1 and 2 respectively.  $\underline{u_1}$  and  $\underline{u_2}$  are breeding values for traits 1 and 2 respectively, and  $\underline{e_1}$  and  $\underline{e_2}$  are residuals for traits 1 and 2 respectively.

The two models in equation 1 can be combined to form the usual presentation of the mixed model as follows:

$$\begin{bmatrix} \underline{y_1} \\ \underline{y_2} \end{bmatrix} = \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix} \begin{bmatrix} \underline{\beta_1} \\ \underline{\beta_2} \end{bmatrix} + \begin{bmatrix} Z_1 & 0 \\ 0 & Z_2 \end{bmatrix} \begin{bmatrix} \underline{u_1} \\ \underline{u_2} \end{bmatrix} + \begin{bmatrix} \underline{e_1} \\ \underline{e_2} \end{bmatrix}\tag{2}$$

$$\underline{y} = X\underline{\beta} + Z\underline{u} + \underline{e}\tag{3}$$

$$\text{Var}(\underline{u}) = G = \text{Var} \begin{bmatrix} \underline{u_1} \\ \underline{u_2} \end{bmatrix} = \begin{bmatrix} A\sigma_{a_1}^2 & A\sigma_{a_{12}} \\ A\sigma_{a_{12}} & A\sigma_{a_2}^2 \end{bmatrix}\tag{4}$$

For  $A$  the relationship matrix,  $\sigma_{a_1}^2$  and  $\sigma_{a_2}^2$  are the additive genetic variances for trait 1 and 2 respectively and  $\sigma_{a_{12}}^2$  the additive genetic covariance between trait 1 and 2.

$$\text{Var}(\underline{e}) = R = \text{Var} \begin{bmatrix} \underline{e_1} \\ \underline{e_2} \end{bmatrix} = \begin{bmatrix} I\sigma_{e_1}^2 & 0 \\ 0 & I\sigma_{e_2}^2 \end{bmatrix}\tag{5}$$

For  $I$  the identity matrix,  $\sigma_{e_1}^2$  and  $\sigma_{e_2}^2$  are the residual variances for traits 1 and 2 respectively, and assuming no environmental covariance between traits 1 and 2.

$$G^{-1} = \begin{bmatrix} A^{-1}\alpha_{11} & A^{-1}\alpha_{12} \\ A^{-1}\alpha_{21} & A^{-1}\alpha_{22} \end{bmatrix}\tag{6}$$

for

$$\begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} = \begin{bmatrix} \sigma_{a_1}^2 & \sigma_{a_{12}} \\ \sigma_{a_{12}} & \sigma_{a_2}^2 \end{bmatrix}^{-1}\tag{7}$$

1. Assuming  $\sigma_{e_1}^2 = \sigma_{e_2}^2 = 1$  in  $R$  and  $G^{-1}$  as above, write the mixed model equations for the **multiple trait model**. *As the residuals are assumed to be independent in equation 5, I assumed that traits are also independent for each other*

$$\begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix} \begin{bmatrix} \underline{\beta}_1 \\ \underline{\beta}_2 \end{bmatrix} + \begin{bmatrix} Z_1 & 0 \\ 0 & Z_2 \end{bmatrix} \begin{bmatrix} \underline{u}_1 \\ \underline{u}_2 \end{bmatrix} + \begin{bmatrix} \underline{e}_1 \\ \underline{e}_2 \end{bmatrix} = \begin{bmatrix} \underline{y}_1 \\ \underline{y}_2 \end{bmatrix}$$

$$\begin{bmatrix} X_1'X_1 & 0 & X_1'Z_1 & 0 \\ 0 & X_2'X_2 & 0 & X_2'Z_2 \\ Z_1'X_1 & 0 & Z_1'Z_1 + A^{-1}\alpha_{11} & A^{-1}\alpha_{12} \\ 0 & Z_2'X_2 & A^{-1}\alpha_{21} & Z_2'Z_2 + A^{-1}\alpha_{22} \end{bmatrix} \begin{bmatrix} \underline{\hat{\beta}}_1 \\ \underline{\hat{\beta}}_2 \\ \underline{\hat{u}}_1 \\ \underline{\hat{u}}_2 \end{bmatrix} = \begin{bmatrix} X_1'y_1 \\ X_2'y_2 \\ Z_1'y_1 \\ Z_2'y_2 \end{bmatrix} \quad (8)$$

2. Write the two sets of mixed model equations that would have obtained if the two traits had been analyzed separately.

$$\begin{bmatrix} X_1'X_1 & X_1'Z \\ Z_1'X_1 & Z_1'Z + A^{-1}\alpha_{11} \end{bmatrix} \begin{bmatrix} \underline{\hat{\beta}}_1 \\ \underline{\hat{u}} \end{bmatrix} = \begin{bmatrix} X_1'y \\ Z_1'y \end{bmatrix} \quad (9)$$

$$\begin{bmatrix} X_2'X_2 & X_2'Z \\ Z_2'X_2 & Z_2'Z + A^{-1}\alpha_{22} \end{bmatrix} \begin{bmatrix} \underline{\hat{\beta}}_2 \\ \underline{\hat{u}} \end{bmatrix} = \begin{bmatrix} X_2'y \\ Z_2'y \end{bmatrix} \quad (10)$$

3. Compare the systems of equations in questions 1 and 2. These systems could yield identical solutions under what conditions? *If  $\alpha_{12} = \alpha_{21} = 0$  which is only possible if the additive genetic components of variance for both measured traits are independent.*

$$\begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + A^{-1}\lambda \end{bmatrix} \begin{bmatrix} \underline{\hat{\beta}} \\ \underline{\hat{u}} \end{bmatrix} = \begin{bmatrix} X'y \\ Z'y \end{bmatrix} \quad (11)$$

4. In the information from MME like these immediately above, how are the accuracy values of the elements in  $\underline{\hat{u}}$  determined? *Accuracy are computed as the correlation between EBV ( $\underline{\hat{u}}$  values) and the true breeding value, or also as the  $\sqrt{\text{reliability}}$  which is the fraction of the additive genetic variance accounted by the EBV.*

5. Same equations, why is  $\lambda = \frac{\sigma_e^2}{\sigma_a^2}$ ? *Because from MME:*

$$\begin{bmatrix} X'R^{-1}X & X'R^{-1}Z \\ Z'R^{-1}X & Z'R^{-1}Z + A^{-1}\frac{1}{\sigma_a^2} \end{bmatrix} \begin{bmatrix} \underline{\hat{\beta}} \\ \underline{\hat{u}} \end{bmatrix} = \begin{bmatrix} X'R^{-1}y \\ Z'R^{-1}y \end{bmatrix} \quad (12)$$

and taking  $R = I\sigma_e^2$ , we can simplify:

$$\begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + A^{-1}\frac{\sigma_e^2}{\sigma_a^2} \end{bmatrix} \begin{bmatrix} \underline{\hat{\beta}} \\ \underline{\hat{u}} \end{bmatrix} = \begin{bmatrix} X'y \\ Z'y \end{bmatrix} \quad (13)$$

6. Describe briefly the two general ways that genomic information has been included in prediction of genetic merit for livestock.

- (a) *As replacement of the Z matrix in the MME: In this procedure, assuming common residual variance across all marker loci, a matrix of numerically encoded markers (each one represented in a column) with rows equal to individuals (called M) is used as replacement of Z and used to compute the G matrix (assuming that each marker is independent and has their individual variance). Both matrices (M and G) are used then in the computation of the genomic relationship matrix to be used as a replacement of  $A^{-1}\lambda$  in the MME equations as follows:*

$$\begin{bmatrix} X'X & X'I \\ I'X & I + \sigma_e^2[MGM']^{-1} \end{bmatrix} \begin{bmatrix} \underline{\hat{\beta}} \\ \underline{\hat{u}} \end{bmatrix} = \begin{bmatrix} X'y \\ I'y \end{bmatrix} \quad (14)$$

*In this procedure, the computed values of  $\underline{\hat{u}}$  are the expected breeding values for each individual.*

- (b) *As the sum of the effect of informative haplotype markers after independent marker selection: This procedure assume that all markers ( $g_i$ ) are independent each other and have common variance  $\sigma_g^2$ , then the breeding value for each marker is computed as in least squares by  $\hat{g} = (X'X + \lambda I)^{-1}X'y$  where  $\lambda$  is a penalization factor used to select only highly informative markers. In this procedure, the expected breeding values for each individual are equal to  $(\sum_{i=1}^n X_i \hat{g}_i)$  the sum of the breeding values of all markers in a given individual.*