

Linear Mixed Models

Guilherme J. M. Rosa

University of Wisconsin-Madison

Mixed Effects Models

Frequently, linear models contain factors whose levels represent a random sample of a population of all possible factor levels

Models containing both fixed and random effects are called mixed effects models

Linear mixed effects models have been widely used in analysis of data where responses are clustered around some random effects, such that there is a natural dependence between observations in the same cluster

For example, consider repeated measurements taken on each subject in longitudinal data, or observations taken on members of the same family in a genetic study

Linear Mixed Effects Model

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

where:

$$\left\{ \begin{array}{l} \mathbf{y}: \text{response vector; observations} \\ \boldsymbol{\beta}: \text{vector of fixed effects} \\ \mathbf{u}: \text{vector of random effects; } \mathbf{u} \sim N(\mathbf{0}, \mathbf{G}) \\ \mathbf{X} \text{ and } \mathbf{Z}: \text{(known) incidence matrices} \\ \mathbf{e}: \text{residual vector; } \mathbf{e} \sim N(\mathbf{0}, \boldsymbol{\Sigma}) \end{array} \right.$$

Linear Mixed Effects Model

Generally, it is assumed that \mathbf{u} and \mathbf{e} are independent from each other, such that:

$$\begin{bmatrix} \mathbf{u} \\ \mathbf{e} \end{bmatrix} \sim \text{MVN} \left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \mathbf{G} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma} \end{bmatrix} \right)$$

Inferences regarding mixed effects models refer to the estimation of fixed effects, the prediction of random effects, and the estimation of variance and covariance components, which are briefly discussed next

Estimation of Fixed Effects

Let $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, where $\boldsymbol{\varepsilon} = \mathbf{Z}\mathbf{u} + \mathbf{e}$

$$\begin{cases} E[\boldsymbol{\varepsilon}] = E[\mathbf{Z}\mathbf{u} + \mathbf{e}] = \mathbf{Z}E[\mathbf{u}] + E[\mathbf{e}] = \mathbf{0} \\ \text{Var}[\boldsymbol{\varepsilon}] = \text{Var}[\mathbf{Z}\mathbf{u} + \mathbf{e}] = \mathbf{Z}\text{Var}[\mathbf{u}]\mathbf{Z}^T + \text{Var}[\mathbf{e}] = \mathbf{Z}\mathbf{G}\mathbf{Z}^T + \boldsymbol{\Sigma} \end{cases}$$

such that $\mathbf{y} \sim \text{MVN}(\mathbf{X}\boldsymbol{\beta}, \mathbf{V})$, where $\mathbf{V} = \mathbf{Z}\mathbf{G}\mathbf{Z}^T + \boldsymbol{\Sigma}$

Under these circumstances, the MLE for $\boldsymbol{\beta}$ is:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{V}^{-1} \mathbf{y} \sim \text{MVN}(\boldsymbol{\beta}, (\mathbf{X}^T \mathbf{V}^{-1} \mathbf{X})^{-1})$$

Estimation of Fixed Effects

As \mathbf{G} and $\boldsymbol{\Sigma}$ are generally unknown, an estimate of \mathbf{V} is used instead such that the estimator becomes:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \hat{\mathbf{V}}^{-1} \mathbf{X})^{-1} \mathbf{X}^T \hat{\mathbf{V}}^{-1} \mathbf{y}$$

The variance-covariance matrix of $\hat{\boldsymbol{\beta}}$ is now approximated by $(\mathbf{X}^T \hat{\mathbf{V}}^{-1} \mathbf{X})^{-1}$

Note: $(\mathbf{X}^T \hat{\mathbf{V}}^{-1} \mathbf{X})^{-1}$ is biased downwards as a consequence of ignoring the variability introduced by working with estimates of (co)variance components instead of their true (unknown) parameter values

Estimation of Fixed Effects

Approximated confidence regions and test statistics for estimable functions of the type $\mathbf{K}^T \boldsymbol{\beta}$ can be obtained by using the result:

$$\frac{(\mathbf{K}^T \boldsymbol{\beta}^0)^T (\mathbf{K}^T (\mathbf{X}^T \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{K}^T \boldsymbol{\beta}^0)}{\text{rank}(\mathbf{K})} \approx F_{[\varphi_N, \varphi_D]}$$

where $F_{[\varphi_N, \varphi_D]}$ refers to an F-distribution with $\varphi_N = \text{rank}(\mathbf{K})$ degrees of freedom for the numerator, and φ_D degrees of freedom for the denominator, which is generally calculated from the data using, for example, the Satterthwaite's approach

Estimation (Prediction) of Random Effects

In addition to the estimation of fixed effects, very often in genetics interest is also on prediction of random effects.

In linear (Gaussian) models such predictions are given by the conditional expectation of \mathbf{u} given the data, i.e. $E[\mathbf{u} | \mathbf{y}]$

Given the model specifications, the joint distribution of \mathbf{y} and \mathbf{u} is:

$$\begin{bmatrix} \mathbf{y} \\ \mathbf{u} \end{bmatrix} \sim \text{MVN} \left(\begin{bmatrix} \mathbf{X}\boldsymbol{\beta} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \mathbf{V} & \mathbf{Z}\mathbf{G} \\ \mathbf{G}\mathbf{Z}^T & \mathbf{G} \end{bmatrix} \right)$$

Estimation (Prediction) of Random Effects

From the properties of multivariate normal distribution, we have that:

$$\begin{aligned} E[\mathbf{u} | \mathbf{y}] &= E[\mathbf{u}] + \text{Cov}[\mathbf{u}, \mathbf{y}^T] \text{Var}^{-1}[\mathbf{y}](\mathbf{y} - E[\mathbf{y}]) \\ &= \mathbf{GZ}^T \mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) = \mathbf{GZ}^T (\mathbf{ZGZ}^T + \boldsymbol{\Sigma})^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \end{aligned}$$

The fixed effects $\boldsymbol{\beta}$ are typically replaced by their estimates, so that predictions are made based on the following expression:

$$\hat{\mathbf{u}} = \mathbf{GZ}^T (\mathbf{ZGZ}^T + \boldsymbol{\Sigma})^{-1}(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})$$

Mixed Model Equations

The solutions $\hat{\boldsymbol{\beta}}$ and $\hat{\mathbf{u}}$ discussed before require \mathbf{V}^{-1}

As \mathbf{V} can be of huge dimensions, especially in animal breeding applications, its inverse is generally computationally demanding if not unfeasible.

However, Henderson (1950) presented the mixed model equations (MME) to estimate $\boldsymbol{\beta}$ and \mathbf{u} simultaneously, without the need for computing \mathbf{V}^{-1}

The MME were derived by maximizing (for $\boldsymbol{\beta}$ and \mathbf{u}) the joint density of \mathbf{y} and \mathbf{u} , expressed as:

$$\begin{aligned} p(\mathbf{y}, \mathbf{u} | \boldsymbol{\beta}, \mathbf{G}, \boldsymbol{\Sigma}) &\propto |\boldsymbol{\Sigma}|^{-1/2} |\mathbf{G}|^{-1/2} \\ &\times \exp \left\{ -\frac{1}{2} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\mathbf{u})^T \boldsymbol{\Sigma}^{-1} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\mathbf{u}) - \frac{1}{2} \mathbf{u}^T \mathbf{G}^{-1} \mathbf{u} \right\} \end{aligned}$$

Mixed Model Equations

The logarithm of this function is:

$$\begin{aligned}\ell = \log[p(\mathbf{y}, \mathbf{u} | \boldsymbol{\beta}, \mathbf{G}, \boldsymbol{\Sigma})] &\propto |\boldsymbol{\Sigma}| + |\mathbf{G}| + (\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\mathbf{u})^T \boldsymbol{\Sigma}^{-1} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\mathbf{u}) + \mathbf{u}^T \mathbf{G}^{-1} \mathbf{u} \\ &= |\boldsymbol{\Sigma}| + |\mathbf{G}| + \mathbf{y}^T \boldsymbol{\Sigma}^{-1} \mathbf{y} - 2\mathbf{y}^T \boldsymbol{\Sigma}^{-1} \mathbf{X}\boldsymbol{\beta} - 2\mathbf{y}^T \boldsymbol{\Sigma}^{-1} \mathbf{Z}\mathbf{u} \\ &\quad + \boldsymbol{\beta}^T \mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X}\boldsymbol{\beta} + 2\boldsymbol{\beta}^T \mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{Z}\mathbf{u} + \mathbf{u}^T \mathbf{Z}^T \boldsymbol{\Sigma}^{-1} \mathbf{Z}\mathbf{u} + \mathbf{u}^T \mathbf{G}^{-1} \mathbf{u}\end{aligned}$$

The derivatives of ℓ regarding $\boldsymbol{\beta}$ and \mathbf{u} are:

$$\begin{bmatrix} \frac{\partial \ell}{\partial \boldsymbol{\beta}} \\ \frac{\partial \ell}{\partial \mathbf{u}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{y} - \mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X}\hat{\boldsymbol{\beta}} - \mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{Z}\hat{\mathbf{u}} \\ \mathbf{Z}^T \boldsymbol{\Sigma}^{-1} \mathbf{y} - \mathbf{Z}^T \boldsymbol{\Sigma}^{-1} \mathbf{X}\hat{\boldsymbol{\beta}} - \mathbf{Z}^T \boldsymbol{\Sigma}^{-1} \mathbf{Z}\hat{\mathbf{u}} - \mathbf{G}^{-1} \hat{\mathbf{u}} \end{bmatrix}$$

Mixed Model Equations

Equating them to zero gives the following system:

$$\begin{bmatrix} \mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X}\hat{\boldsymbol{\beta}} + \mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{Z}\hat{\mathbf{u}} \\ \mathbf{Z}^T \boldsymbol{\Sigma}^{-1} \mathbf{X}\hat{\boldsymbol{\beta}} + \mathbf{Z}^T \boldsymbol{\Sigma}^{-1} \mathbf{Z}\hat{\mathbf{u}} + \mathbf{G}^{-1} \hat{\mathbf{u}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{y} \\ \mathbf{Z}^T \boldsymbol{\Sigma}^{-1} \mathbf{y} \end{bmatrix}$$

which can be expressed as:

$$\begin{bmatrix} \mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X} & \mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{Z} \\ \mathbf{Z}^T \boldsymbol{\Sigma}^{-1} \mathbf{X} & \mathbf{Z}^T \boldsymbol{\Sigma}^{-1} \mathbf{Z} + \mathbf{G}^{-1} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\mathbf{u}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{y} \\ \mathbf{Z}^T \boldsymbol{\Sigma}^{-1} \mathbf{y} \end{bmatrix}$$

known as the mixed model equations (MME)

BLUE and BLUP

Using the second part of the MME, we have that:

$$\mathbf{Z}^T \boldsymbol{\Sigma}^{-1} \mathbf{X} \hat{\boldsymbol{\beta}} + (\mathbf{Z}^T \boldsymbol{\Sigma}^{-1} \mathbf{Z} + \mathbf{G}^{-1}) \hat{\mathbf{u}} = \mathbf{Z}^T \boldsymbol{\Sigma}^{-1} \mathbf{y}$$

so that:

$$\hat{\mathbf{u}} = (\mathbf{Z}^T \boldsymbol{\Sigma}^{-1} \mathbf{Z} + \mathbf{G}^{-1})^{-1} \mathbf{Z}^T \boldsymbol{\Sigma}^{-1} (\mathbf{y} - \mathbf{X} \hat{\boldsymbol{\beta}})$$

It can be shown that this expression is equivalent to:

$$\hat{\mathbf{u}} = \mathbf{G} \mathbf{Z}^T (\mathbf{Z} \mathbf{G} \mathbf{Z}^T + \boldsymbol{\Sigma})^{-1} (\mathbf{y} - \mathbf{X} \hat{\boldsymbol{\beta}})$$

and, more importantly, that $\hat{\mathbf{u}}$ is the **best linear unbiased predictor (BLUP)** of \mathbf{u}

BLUE and BLUP

Using this result into the first part of the MME, we have that:

$$\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X} \hat{\boldsymbol{\beta}} + \mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{Z} \hat{\mathbf{u}} = \mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{y}$$

$$\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X} \hat{\boldsymbol{\beta}} + \mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{Z} (\mathbf{Z}^T \boldsymbol{\Sigma}^{-1} \mathbf{Z} + \mathbf{G}^{-1})^{-1} \mathbf{Z}^T \boldsymbol{\Sigma}^{-1} (\mathbf{y} - \mathbf{X} \hat{\boldsymbol{\beta}}) = \mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{y}$$

$$\hat{\boldsymbol{\beta}} = \{ \mathbf{X}^T [\boldsymbol{\Sigma}^{-1} - \boldsymbol{\Sigma}^{-1} \mathbf{Z} (\mathbf{Z}^T \boldsymbol{\Sigma}^{-1} \mathbf{Z} + \mathbf{G}^{-1})^{-1} \mathbf{Z}^T \boldsymbol{\Sigma}^{-1}] \mathbf{X} \}^{-1} \mathbf{X}^T [\boldsymbol{\Sigma}^{-1} - \boldsymbol{\Sigma}^{-1} \mathbf{Z} (\mathbf{Z}^T \boldsymbol{\Sigma}^{-1} \mathbf{Z} + \mathbf{G}^{-1})^{-1} \mathbf{Z}^T \boldsymbol{\Sigma}^{-1}] \mathbf{y}$$

Similarly, it is shown that this expression is equivalent to $\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{V}^{-1} \mathbf{y}$, which is the **best linear unbiased estimator (BLUE)** of $\boldsymbol{\beta}$.

BLUE and BLUP

It is important to note that $\hat{\beta}$ and \hat{u} require knowledge of G and Σ . These matrices, however, are rarely known. This is a problem without an exact solution using classical methods.

The practical approach is to replace G and Σ by their estimates (\hat{G} and $\hat{\Sigma}$) into the MME:

$$\begin{bmatrix} \mathbf{X}'\hat{\Sigma}^{-1}\mathbf{X} & \mathbf{X}'\hat{\Sigma}^{-1}\mathbf{Z} \\ \mathbf{Z}'\hat{\Sigma}^{-1}\mathbf{X} & \mathbf{Z}'\hat{\Sigma}^{-1}\mathbf{Z} + \hat{G}^{-1} \end{bmatrix} \begin{bmatrix} \tilde{\beta} \\ \tilde{u} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\hat{\Sigma}^{-1}\mathbf{y} \\ \mathbf{Z}'\hat{\Sigma}^{-1}\mathbf{y} \end{bmatrix}$$

Estimation of Variance Components ANOVA Estimation

Consider the data set below, related to observations of half-sib families of k unrelated sires. The following model can be used to represent these data:

$$y_{ij} = \mu + s_i + e_{ij}$$

Sire			
1	2	...	k
y_{11}	y_{21}	...	y_{k1}
y_{12}	y_{22}	...	y_{k2}
\vdots	\vdots		\vdots
y_{1n_1}	y_{2n_2}	...	y_{kn_k}

where y_{ij} represents the phenotypic trait observation of progeny j ($j = 1, 2, \dots, n_i$) in family i , μ is a mean, s_i is an effect common to all animals having sire i , and e_{ij} is a residual term

Estimation of Variance Components

ANOVA Estimation

The sire effect s_i is equivalent to the transmitting ability (which is equal to one-half additive genetic value) of sire i , as one-half of its genes are (randomly) transmitted to each of its n_i progeny.

The residual terms e_{ij} refer to additional genetics effects (such as the effect of dams) and environmental components.

It is assumed that $S_i \stackrel{\text{ind}}{\sim} (0, \sigma_s^2)$ and $e_{ij} \stackrel{\text{ind}}{\sim} (0, \sigma_e^2)$

From the model settings discussed before we have that

$$E[y_{ij}] = \mu \quad \text{and} \quad \text{Var}[y_{ij}] = \sigma_s^2 + \sigma_e^2$$

The overall sample mean is given by $\bar{y}_{..} = \frac{1}{N} \sum_{i=1}^k \sum_{j=1}^{n_i} y_{ij} = \frac{1}{N} \sum_{i=1}^k y_{i\cdot}$

where $N = \sum_{i=1}^k n_i$, and $\bar{y}_{i\cdot} = \frac{1}{n_i} \sum_{j=1}^{n_i} y_{ij}$ are sire-specific means.

The ANOVA approach consists of an orthogonal decomposition of the total sum of squares (TSS) into between classes (or, in our case, sires) and within classes (or residual) components. The corrected (in terms of the

general mean) TSS is given by:
$$\text{TSS} = \sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{..})^2$$

By adding and subtracting $\bar{y}_{i\cdot}$ within the parentheses, the TSS can be expressed as:

$$\begin{aligned} \text{TSS} &= \sum_{i=1}^k \sum_{j=1}^{n_i} [(y_{ij} - \bar{y}_{i\cdot}) + (\bar{y}_{i\cdot} - \bar{y}_{..})]^2 \\ &= \sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i\cdot})^2 + \sum_{i=1}^k \sum_{j=1}^{n_i} (\bar{y}_{i\cdot} - \bar{y}_{..})^2 + 2 \sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i\cdot})(\bar{y}_{i\cdot} - \bar{y}_{..}) \end{aligned}$$

It is seen that the last part of this expression is equal to zero, so that TSS can be written as two components:

$$\text{SSS} = \sum_{i=1}^k \sum_{j=1}^{n_i} (\bar{y}_{i\cdot} - \bar{y}_{..})^2 \quad \text{and} \quad \text{RSS} = \sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i\cdot})^2$$

which are the sire and the residual sum of squares, respectively. The SSS term measures the variation of each progeny family around the overall mean, while the RSS term measures the extra variation related to each observation around its sire average

It can be shown that the expectation of these sums of squares terms are:

$$E[\text{SSS}] = \left(N - \frac{1}{N} \sum_{i=1}^{n_i} n_i^2 \right) \sigma_s^2 + (k-1) \sigma_e^2 \quad \text{and} \quad E[\text{RSS}] = (N-k) \sigma_e^2$$

so that the ANOVA estimators of the sire and residual variance components are given by:

$$\hat{\sigma}_s^2 = \left(N - \frac{1}{N} \sum_{i=1}^{n_i} n_i^2 \right)^{-1} [\text{SSS} - (k-1) \hat{\sigma}_e^2] \quad \text{and} \quad \hat{\sigma}_e^2 = \frac{1}{(N-k)} \text{RSS}$$

In the specific case of balanced data, i.e. the same progeny size for all sires, $n_i = n = N/k$ and the ANOVA estimators become:

$$\hat{\sigma}_s^2 = \frac{1}{n} \left[\frac{1}{(k-1)} \text{SSS} - \hat{\sigma}_e^2 \right] \quad \text{and} \quad \hat{\sigma}_e^2 = \frac{1}{k(n-1)} \text{RSS}$$

Estimation of Variance Components

ANOVA approach works well for simple models (such as a one-way structure) or balanced data (such as data from designed experiments with no missing data), but they are not indicated for more complex models and data structures

Other proposed methods: *expected mean squares* approach of Henderson (1953), and the *minimum norm quadratic unbiased estimation* (Rao 1971a, 1971b), among others.

However, *maximum likelihood* based methods are currently the most popular, especially the *restricted (or residual) maximum likelihood* (REML) approach, which attempts to correct for the well-known bias in the classical maximum likelihood (ML) estimation of variance components. These two methods are briefly described next.

Estimation of Variance Components

Maximum Likelihood (ML) Estimator

Maximum likelihood estimates of the variance components can be obtained by maximizing the log-likelihood $L(\boldsymbol{\beta}, \mathbf{G}, \boldsymbol{\Sigma})$ with respect to each element of \mathbf{G} and $\boldsymbol{\Sigma}$, after replacing $\boldsymbol{\beta}$ by $\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{V}^{-1} \mathbf{y}$

Alternatively, \mathbf{G} , $\boldsymbol{\Sigma}$, and $\boldsymbol{\beta}$ can be estimated simultaneously by maximizing their joint log-likelihood with respect to the variance components and the fixed effects.

As a simple example of maximum likelihood estimation of variance components, consider the balanced case (i.e., constant progeny sizes) half-sib families data set discussed previously, and the linear model:

$$y_{ij} = \mu + s_i + e_{ij}$$

with the same definitions as before, but with the additional assumption of normality of both the sire and the residual effects, i.e.:

$$s_i \stackrel{\text{ind}}{\sim} N(0, \sigma_s^2) \quad \text{and} \quad e_{ij} \stackrel{\text{ind}}{\sim} N(0, \sigma_e^2)$$

In matrix notation, this model can be expressed as:

$$\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_k \end{bmatrix} = \begin{bmatrix} \mathbf{1}_n \\ \mathbf{1}_n \\ \vdots \\ \mathbf{1}_n \end{bmatrix} \mu + \begin{bmatrix} \mathbf{1}_n & \mathbf{0}_n & \cdots & \mathbf{0}_n \\ \mathbf{0}_n & \mathbf{1}_n & \cdots & \mathbf{0}_n \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_n & \mathbf{0}_n & \cdots & \mathbf{1}_n \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_k \end{bmatrix} + \begin{bmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \\ \vdots \\ \mathbf{e}_k \end{bmatrix}$$

where $\mathbf{y}_i = [y_{i1} \ y_{i2} \ \cdots \ y_{in}]^T$ represents the vector of observations of progeny i (i.e., relative to sire i); $\mathbf{1}_n$ and $\mathbf{0}_n$ represent n -dimensional column vectors of 1's and 0's, respectively; and $\mathbf{e}_i = [e_{i1}, e_{i2}, \dots, e_{in}]^T$ is the vector of residuals associated with progeny i

The vector of observations $\mathbf{y} = [\mathbf{y}_1^T \mathbf{y}_2^T \dots \mathbf{y}_k^T]^T$ has then a multivariate normal distr. with mean vector $\boldsymbol{\mu} = \mathbf{1}_N \boldsymbol{\mu}$ and variance-covariance matrix given by $\mathbf{I}_s \otimes (\mathbf{1}_n \sigma_s^2 \mathbf{1}_n^T) + \mathbf{I}_N \sigma_e^2$, and its density function (from which the likelihood function obtained) can be written as:

$$p(\mathbf{y} | \boldsymbol{\mu}, \sigma_s^2, \sigma_e^2) = \frac{1}{(2\pi)^{N/2} |\mathbf{I}_s \otimes \mathbf{J}_n \sigma_s^2 + \mathbf{I}_N \sigma_e^2|^{1/2}} \\ \times \exp \left\{ -\frac{1}{2} (\mathbf{y} - \mathbf{1}_N \boldsymbol{\mu})^T (\mathbf{J}_n \sigma_s^2 + \mathbf{I}_N \sigma_e^2)^{-1} (\mathbf{y} - \mathbf{1}_N \boldsymbol{\mu}) \right\} \\ = (2\pi)^{\frac{N}{2}} (\sigma_e^2)^{-\frac{(N-k)}{2}} (\sigma_e^2 + n\sigma_s^2)^{-\frac{k}{2}} \exp \left\{ -\frac{1}{2} (\mathbf{y} - \mathbf{1}_N \boldsymbol{\mu})^T \left[\mathbf{I}_s \otimes \mathbf{J}_n \left(\frac{1}{n} \left(\frac{1}{\sigma_e^2 + n\sigma_s^2} - \frac{1}{\sigma_e^2} \right) \right) \right] (\mathbf{y} - \mathbf{1}_N \boldsymbol{\mu}) \right\}$$

where $\mathbf{J}_n = \mathbf{1}_n \mathbf{1}_n^T$ is an $(n \times n)$ matrix of 1's, and \otimes is the Kronecker product

The log-likelihood function can be written then as:

$$l(\boldsymbol{\mu}, \sigma_s^2, \sigma_e^2) \propto -\frac{(N-k)}{2} \log(\sigma_e^2) - \frac{k}{2} \log(\sigma_e^2 + n\sigma_s^2) - \frac{1}{2\sigma_e^2} \sum_{i=1}^k \sum_{j=1}^n (y_{ij} - \bar{y}_{i\cdot})^2 - \frac{1}{2} \sum_{i=1}^k \frac{n(\bar{y}_{i\cdot} - \boldsymbol{\mu})^2}{\sigma_e^2 + n\sigma_s^2}$$

By taking the derivatives and setting them to 0, the following solutions are obtained:

$$\hat{\boldsymbol{\mu}} = \bar{\mathbf{y}}_{..} \quad , \quad \hat{\sigma}_e^2 = \frac{1}{k(n-1)} \text{RSS} \quad \text{and} \quad \hat{\sigma}_s^2 = \frac{1}{n} \left[\frac{\text{SSS}}{k} - \hat{\sigma}_e^2 \right]$$

from which ML estimates of the variance components are obtained, except if $\hat{\sigma}_s^2 < 0$, in which case the estimate is set to zero

ML estimates of variance components are biased downwards as they do not take into account the degrees of freedom used for estimating the fixed effects

Estimation of Variance Components

Residual Maximum Likelihood (REML) Estimator

Restricted (or residual) maximum likelihood approach (REML): **corrects the bias** associated with ML estimates by taking into account the degrees of freedom used for estimating the fixed effects

REML maximizes the likelihood function of a set of **error contrasts** $\mathbf{d} = \mathbf{L}^T \mathbf{y}$, where \mathbf{L} is a $[n \times (n - p)]$ full-rank matrix with columns orthogonal to the columns of the incidence matrix \mathbf{X}

The vector \mathbf{d} follows a multivariate normal distribution with null mean vector and variance-covariance matrix $\mathbf{L}^T \mathbf{V} \mathbf{L} = \mathbf{L}^T (\mathbf{Z} \mathbf{G} \mathbf{Z}^T + \mathbf{\Sigma}) \mathbf{L}$. Note that the distribution of \mathbf{d} does not depend on $\boldsymbol{\beta}$.

The residual likelihood function for the variance components is then:

$$L(\mathbf{G}, \mathbf{\Sigma} | \mathbf{y}) = (2\pi)^{-(n-p)/2} |\mathbf{L}^T \mathbf{V} \mathbf{L}|^{-1/2} \exp \left\{ -\frac{1}{2} \mathbf{d}^T (\mathbf{L}^T \mathbf{V} \mathbf{L})^{-1} \mathbf{d} \right\}$$

Another approach for obtaining the residual likelihood function for the variance components is by **integrating the fixed effects** out of the 'full' likelihood function, i.e.:

$$L(\mathbf{G}, \mathbf{\Sigma} | \mathbf{y}) = \int L(\boldsymbol{\beta}, \mathbf{G}, \mathbf{\Sigma} | \mathbf{y}) d\boldsymbol{\beta}$$

as illustrated in the following example.

Recall the balanced half-sib families data set, and its associated likelihood function:

$$L(\mu, \sigma_s^2, \sigma_e^2) = (2\pi)^{-\frac{N}{2}} (\sigma_e^2)^{-\frac{(N-k)}{2}} (\sigma_e^2 + n\sigma_s^2)^{-\frac{k}{2}} \\ \times \exp\left\{-\frac{1}{2\sigma_e^2} \sum_{i=1}^k \sum_{j=1}^n (y_{ij} - \bar{y}_{i\cdot})^2 - \frac{1}{2} \sum_{i=1}^k \frac{n(\bar{y}_{i\cdot} - \mu)^2}{\sigma_e^2 + n\sigma_s^2}\right\}$$

Its residual likelihood is then:

$$L(\sigma_s^2, \sigma_e^2) = \int L(\mu, \sigma_s^2, \sigma_e^2) d\mu \\ = (2\pi)^{-\frac{N}{2}} (\sigma_e^2)^{-\frac{(N-k)}{2}} (\sigma_e^2 + n\sigma_s^2)^{-\frac{k}{2}} \\ \times \exp\left\{-\frac{1}{2\sigma_e^2} \sum_{i=1}^k \sum_{j=1}^n (y_{ij} - \bar{y}_{i\cdot})^2\right\} \int \exp\left\{-\frac{1}{2} \sum_{i=1}^k \frac{n(\bar{y}_{i\cdot} - \mu)^2}{\sigma_e^2 + n\sigma_s^2}\right\} d\mu$$

which is equal to:

$$L(\sigma_s^2, \sigma_e^2) = (2\pi)^{-\frac{N}{2}} (\sigma_e^2)^{-\frac{(N-k)}{2}} \lambda^{-\frac{k}{2}} \\ \times \exp\left\{-\frac{1}{2\sigma_e^2} \sum_{i=1}^k \sum_{j=1}^n (y_{ij} - \bar{y}_{i\cdot})^2\right\} \exp\left\{-\frac{n}{2\lambda} \sum_{i=1}^k (\bar{y}_{i\cdot} - \mu)^2\right\} \sqrt{2\pi \frac{\lambda}{kn}}$$

where $\lambda = \sigma_e^2 + n\sigma_s^2$.

By taking the derivatives with respect to λ and σ_e^2 , and by using the invariance property of maximum likelihood estimators, the following solutions are obtained:

$$\hat{\sigma}_e^2 = \frac{1}{k(n-1)} \text{RSS} \quad \text{and} \quad \hat{\sigma}_s^2 = \frac{1}{n} \left[\frac{1}{(k-1)} \text{SSS} - \hat{\sigma}_e^2 \right]$$

which are the REML estimates of the variance components, except if $\hat{\sigma}_s^2 < 0$, i.e. if

$$\text{SSS} < \frac{(k-1)}{k(n-1)} \text{RSS}$$

Explicit forms of ML and REML estimators are often not available for more complex mixed effects models

ML and REML estimates are then generally obtained by [iterative approaches](#) such as the expectation-maximization (EM) algorithm and Newton-Raphson-based procedures

Mixed Models in Animal and Plant Breeding

Animal/plant breeding programs are based on the principle that phenotypic observations on related individuals can provide information about their underlying genotypic values

The additive component of genetic variation is the primary determinant of the degree to which offspring resemble their parents, and therefore this is usually the component of interest in artificial selection programs

Mixed Models in Animal and Plant Breeding

Many statistical methods for analysis of genetic data are specific (or more appropriate) for phenotypic measurements obtained from planned experimental designs and with balanced data sets

While such situations may be possible within laboratory or greenhouse experimental settings, data from natural populations and agricultural species are generally highly unbalanced and fragmented by numerous kinds of relationships

Animal Model

Culling of data to accommodate conventional statistical techniques (e.g. ANOVA) may introduce bias and/or lead to a substantial loss of information

The mixed model methodology allows efficient estimation of genetic parameters (such as variance components and heritability) and breeding values while accommodating extended pedigrees, unequal family sizes, overlapping generations, sex-limited traits, assortative mating, and natural or artificial selection

To illustrate such application of mixed models in breeding programs, we consider here the so-called Animal Model in situations with a single trait and a single observation (including missing values) per individual

Animal Model

The animal model can be described as:

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

\mathbf{y} is an $(n \times 1)$ vector of observations (phenotypic scores)

$\boldsymbol{\beta}$ is a $(p \times 1)$ vector of fixed effects (e.g. herd-year-season effects)

$\mathbf{u} \sim N(\mathbf{0}, \mathbf{G})$ is a $(q \times 1)$ vector of breeding values (relative to all individuals with record or in the pedigree file, such that q is in general bigger than n)

$\mathbf{e} \sim N(\mathbf{0}, \mathbf{I}_n\sigma_e^2)$ represents residual effects, where σ_e^2 is the residual variance

The Matrix \mathbf{A}

The matrix \mathbf{G} describing the covariances among the random effects (here the breeding values) follows from standard results for the covariances between relatives

It is seen that the additive genetic covariance between two relatives i and i' is given by $2\theta_{ii'}\sigma_a^2$, where $\theta_{ii'}$ is the **coefficient of coancestry** between individuals i and i' , and σ_a^2 is the additive genetic variance in the base population

Hence, under the animal model, $\mathbf{G} = \mathbf{A}\sigma_a^2$, where \mathbf{A} is the **additive genetic (or numerator) relationship matrix**, having elements given by $a_{ii'} = 2\theta_{ii'}$

The Matrix \mathbf{A}

For each animal i in the pedigree ($i = 1, 2, \dots, n$), going from older to younger animals, compute a_{ii} and a_{ij} ($j = 1, 2, \dots, i-1$) as follows:

If both parents (s and d) of animal i are known:

$$a_{ij} = a_{ji} = (a_{js} + a_{jd})/2 \text{ and } a_{ii} = 1 + a_{sd}/2$$

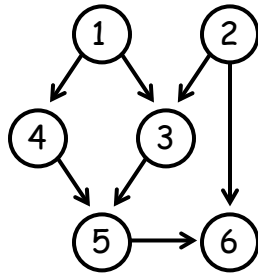
If only one parent (e.g. d) of animal i is known:

$$a_{ij} = a_{ji} = a_{jd}/2 \text{ and } a_{ii} = 1$$

If parents unknown:

$$a_{ij} = a_{ji} = 0 \text{ and } a_{ii} = 1$$

Example



Animal	Sire	Dam
1	-	-
2	-	-
3	1	2
4	1	-
5	4	3
6	5	2



pedigree matrix A

$$A = \begin{bmatrix} 1 & 0 & .5 & .5 & .5 & .25 \\ 0 & 1 & .5 & 0 & .25 & .625 \\ .5 & .5 & 1 & .25 & .625 & .563 \\ .5 & 0 & .25 & 1 & .625 & .313 \\ .5 & .25 & .625 & .625 & 1.125 & .688 \\ .25 & .625 & .563 & .313 & .688 & 1.125 \end{bmatrix}$$

Animal Model

In general, in animal/plant breeding interest is on prediction of breeding values (for selection of superior individuals), and on estimation of variance components and functions thereof, such as heritability

The fixed effects are, in some sense, nuisance factors with no central interest in terms of inferences, but which need to be taken into account (i.e., they need to be corrected for when inferring breeding values)

Animal Model

Since under the animal model $\mathbf{G}^{-1} = \mathbf{A}^{-1}\sigma_a^{-2}$ and $\mathbf{R}^{-1} = \mathbf{I}_n\sigma_e^{-2}$, the mixed model equations can be expressed as:

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

where $\lambda = \frac{\sigma_e^2}{\sigma_a^2} = \frac{1-h^2}{h^2}$, such that:

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

Conditional on the variance components ratio λ , the BLUP of the breeding values are given then by:

$$\hat{\mathbf{u}} = (\mathbf{Z}^T\mathbf{Z} + \lambda\mathbf{A}^{-1})^{-1}\mathbf{Z}^T(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})$$

These are generally referred to as **Estimated Breeding Values (EBV)**

Alternatively, some breeders associations express their results as Predicted Transmitting Abilities (PTA) (or Estimated Transmitting Abilities (ETA) or Expected Progeny Difference (EPD)), which are equal to half the EBV, representing the portion of an animal's breeding values that is passed to its offspring

The amount of information contained in an animal's genetic evaluation depends on the availability of its own record, as well as how many (and how close) relatives it has with phenotypic information

As a measure of amount of information in livestock genetic evaluations, EBVs are typically reported with its associated accuracies

Accuracy of predictions is defined as the correlation between true and estimated breeding values, i.e., $r_i = \rho(\hat{u}_i, u_i)$

Instead of accuracy, some livestock species genetic evaluations use **reliability**, which is the squared correlation of accuracy (r_i^2)

Prediction Accuracy

The calculation of $\rho(\hat{u}_i, u_i)$ requires the diagonal elements of the inverse of the **MME coefficient matrix**, represented as:

$$\mathbf{C} = \begin{bmatrix} \mathbf{X}^T \mathbf{X} & \mathbf{X}^T \mathbf{Z} \\ \mathbf{Z}^T \mathbf{X} & \mathbf{Z}^T \mathbf{Z} + \lambda \mathbf{A}^{-1} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{C}^{\beta\beta} & \mathbf{C}^{\beta u} \\ \mathbf{C}^{u\beta} & \mathbf{C}^{uu} \end{bmatrix}$$

It is shown that the **prediction error variance** of EBV \hat{u}_i is given by:

$$\text{PEV} = \text{Var}(\hat{u}_i - u_i) = c_i^{uu} \sigma_e^2$$

where c_i^{uu} is the i -th diagonal element of \mathbf{C}^{uu} , relative to animal i .

Prediction Accuracy

The PEV can be interpreted as the fraction of additive genetic variance not accounted for by the prediction

Therefore, PEV can be expressed also as:

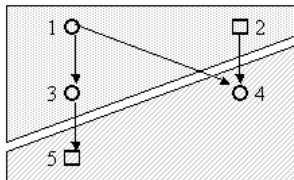
$$\text{PEV} = (1 - r_i^2) \sigma_a^2$$

such that $c_i^{uu} \sigma_e^2 = (1 - r_i^2) \sigma_a^2$, from which the reliability is obtained as:

$$r_i^2 = 1 - c_i^{uu} \sigma_e^2 / \sigma_a^2 = 1 - \lambda c_i^{uu}$$

Animal Model

herd 1



Animal	Sire	Dam	Herd	Observation
1	-	-	h1	310
2	-	-	h1	-
3	-	1	h1	270
4	2	1	h2	350
5	-	3	h2	-

herd 2

$$\begin{bmatrix} 310 \\ 270 \\ 350 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_3 \\ e_4 \end{bmatrix}$$

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

Animal Model

Breeding values: $\mathbf{u} \sim N(\mathbf{0}, \mathbf{A}\sigma_u^2)$, with

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0.5 & 0.5 & 0.25 \\ 0 & 1 & 0 & 0.5 & 0 \\ 0.5 & 0 & 1 & 0.25 & 0.5 \\ 0.5 & 0.5 & 0.25 & 1 & 0.125 \\ 0.25 & 0 & 0.5 & 0.125 & 1 \end{bmatrix}$$

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

$$\lambda = \frac{\sigma_e^2}{\sigma_u^2} = \frac{1-h^2}{h^2}$$

R Code



animal model
toy example

```

y<-matrix(c(310,270,350),nrow=3)
X<-matrix(c(1,1,0,0,0,1),nrow=3)
Z<-matrix(c(1,0,0,0,0,0,0,1,0,0,0,0,1,0),nrow=3, byrow = TRUE)
A<-matrix(c(1,0,0.5,0.5,0.25,
            0,1,0,0.5,0,
            0.5,0,1,0.25,0.5,
            0.5,0.5,0.25,1,0.125,
            0.25,0,0.5,0.125,1),nrow=5)

```

```

h2<-1/3 # heritability
a=(1-h2)/h2

```

```

# crossproducts
XX<-crossprod(X,X)
XZ<-t(X) %*% Z
ZX<-t(Z) %*% X
ZZ<-crossprod(Z,Z)+a*solve(A)

```

```

# mixed model equations
# coefficient matrix and right hand side
C<-rbind(cbind(XX,XZ),cbind(ZX,ZZ))
rhs<-rbind(t(X) %*% y,t(Z) %*% y)

```

```

#solution
theta.hat <- solve(C) %*% rhs

```

$$h^2 = \frac{1}{3} \rightarrow \alpha = 2 \Rightarrow$$

$$\begin{cases} \hat{h}_1 = 290 \\ \hat{h}_2 = 348 \\ \hat{u}_1 = 4.0 \\ \hat{u}_2 = 0.0 \\ \hat{u}_3 = -4.0 \\ \hat{u}_4 = 2.0 \\ \hat{u}_5 = -2.0 \end{cases}$$

Animal Model

The animal model can be extended to model multiple (correlated) traits, multiple random effects (such as maternal effects and common environmental effects), repeated records (e.g. test day models), and so on

Example (Mrode 1996, pp74-76): Weaning weight (kg) of piglets, progeny of three sows mated to two boars:

Piglet	Sire	Dam	Sex	Weight
6	1	2	1	90
7	1	2	2	70
8	1	2	2	65
9	3	4	2	98
10	3	4	1	106
11	3	4	2	60
12	3	4	2	80
13	1	5	1	100
14	1	5	2	85
15	1	5	1	68

A linear model with the (fixed) effect of sex, and the (random) effects of common environment (related to each litter) and breeding values can be expressed as \mathbf{X} :

$$\begin{array}{c}
 \text{Weight} \nearrow \mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{W}\mathbf{c} + \mathbf{e} \\
 \begin{array}{ccc}
 \nearrow \text{Sex} & \nearrow \text{Breeding} & \nearrow \text{Common} \\
 & \text{values} & \text{environment} \\
 & & \nwarrow \text{Residual}
 \end{array}
 \end{array}$$

Assuming that $\sigma_u^2 = 20$, $\sigma_c^2 = 15$ and $\sigma_e^2 = 65$, the MME are as follows:

$$\begin{bmatrix}
 \mathbf{X}^T\mathbf{X} & \mathbf{X}^T\mathbf{Z} & \mathbf{X}^T\mathbf{W} \\
 \mathbf{Z}^T\mathbf{X} & \mathbf{Z}^T\mathbf{Z} + \mathbf{A}^{-1}\lambda_1 & \mathbf{Z}^T\mathbf{W} \\
 \mathbf{W}^T\mathbf{X} & \mathbf{W}^T\mathbf{Z} & \mathbf{W}^T\mathbf{W} + \mathbf{I}\lambda_2
 \end{bmatrix}
 \begin{bmatrix}
 \hat{\boldsymbol{\beta}} \\
 \hat{\mathbf{u}} \\
 \hat{\mathbf{c}}
 \end{bmatrix}
 =
 \begin{bmatrix}
 \mathbf{X}^T\mathbf{y} \\
 \mathbf{Z}^T\mathbf{y} \\
 \mathbf{W}^T\mathbf{y}
 \end{bmatrix}$$

where $\lambda_1 = \frac{\sigma_e^2}{\sigma_u^2} = 3.25$ and $\lambda_2 = \frac{\sigma_e^2}{\sigma_c^2} = 4.3$

The BLUEs and BLUPs
(inverting the numerator
relationship matrix) are:



Mrode example

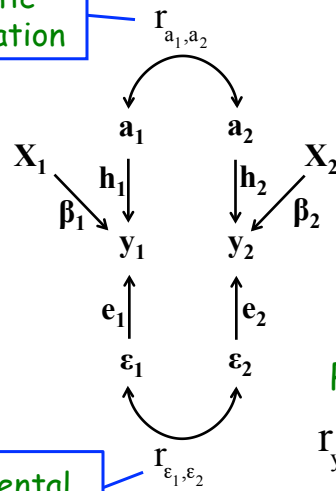
Effects	Solutions
<i>Sex</i>	
1	91.493
2	75.764

<i>Animals</i>	
1	-1.441
2	-1.175
3	1.441
4	1.441
5	-0.266
6	-1.098
7	-1.667
8	-2.334
9	3.925
10	2.895
11	-1.141
12	1.525
13	0.448
14	0.545
15	-3.819

<i>Environ.</i>	
2	-1.762
4	2.161
5	-0.399

Multiple (Correlated) Traits

Genetic correlation



$$\begin{cases} h_j = \sqrt{h_j^2} \\ e_j = \sqrt{1 - h_j^2} \end{cases}$$

Environmental correlation

Phenotypic correlation:

$$r_{y_1, y_2} = h_1 h_2 r_{a_1, a_2} + e_1 e_2 r_{\epsilon_1, \epsilon_2}$$

Multiple (Correlated) Traits

The animal model can be extended for the joint analysis of multiple traits

Let the model for each of k traits be:

$$\mathbf{y}_j = \mathbf{X}_j \boldsymbol{\beta}_j + \mathbf{Z}_j \mathbf{a}_j + \boldsymbol{\varepsilon}_j$$

where j is an index to indicate the trait (j = 1, 2, ..., k).

For the joint analysis of the k trait, the model becomes:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{a} + \boldsymbol{\varepsilon}$$

with design matrices given by:

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{X}_k \end{bmatrix} \quad \mathbf{Z} = \begin{bmatrix} \mathbf{Z}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{Z}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{Z}_k \end{bmatrix}$$

Multiple (Correlated) Traits

In this case it is assumed that:

$$\text{Var} \begin{bmatrix} \mathbf{a} \\ \boldsymbol{\varepsilon} \end{bmatrix} = \begin{bmatrix} \mathbf{G} \otimes \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma} \otimes \mathbf{I} \end{bmatrix}$$

where \mathbf{G} and $\boldsymbol{\Sigma}$ are the genetic and residual variance-covariance matrices, given by:

$$\mathbf{G} = \begin{bmatrix} \sigma_{a_1}^2 & \sigma_{a_1 a_2} & \cdots & \sigma_{a_1 a_k} \\ \sigma_{a_1 a_2} & \sigma_{a_2}^2 & \cdots & \sigma_{a_2 a_k} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{a_1 a_k} & \sigma_{a_2 a_k} & \cdots & \sigma_{a_k}^2 \end{bmatrix} \quad \boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{\varepsilon_1}^2 & \sigma_{\varepsilon_1 \varepsilon_2} & \cdots & \sigma_{\varepsilon_1 \varepsilon_k} \\ \sigma_{\varepsilon_1 \varepsilon_2} & \sigma_{\varepsilon_2}^2 & \cdots & \sigma_{\varepsilon_2 \varepsilon_k} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{\varepsilon_1 \varepsilon_k} & \sigma_{\varepsilon_2 \varepsilon_k} & \cdots & \sigma_{\varepsilon_k}^2 \end{bmatrix}$$

Note: \otimes represents the direct (Kronecker) product

Multiple (Correlated) Traits

The MME for multi-trait analyses are of the same form as before, i.e.:

$$\begin{bmatrix} \mathbf{X}'(\boldsymbol{\Sigma}^{-1} \otimes \mathbf{I})\mathbf{X} & \mathbf{X}'(\boldsymbol{\Sigma}^{-1} \otimes \mathbf{I})\mathbf{Z} \\ \mathbf{Z}'(\boldsymbol{\Sigma}^{-1} \otimes \mathbf{I})\mathbf{X} & \mathbf{Z}'(\boldsymbol{\Sigma}^{-1} \otimes \mathbf{I})\mathbf{Z} + \mathbf{G}^{-1} \otimes \mathbf{A}^{-1} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\mathbf{a}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'(\boldsymbol{\Sigma}^{-1} \otimes \mathbf{I})\mathbf{y} \\ \mathbf{Z}'(\boldsymbol{\Sigma}^{-1} \otimes \mathbf{I})\mathbf{y} \end{bmatrix}$$

from which the BLUEs and BLUPs of $\boldsymbol{\beta}$ and \mathbf{a} can be obtained.

Multiple (Correlated) Traits

The dimensionality of multi-trait MME, however, can become a hurdle for solving it when more than two or three traits are considered

An alternative for the analysis of multiple traits is to use a [canonical transformation](#) of the traits, which consists of transforming the vectors of correlated traits into a new vector of uncorrelated variables

In such case, each transformed variable can be analyzed independently using standard single trait models, and subsequently the estimated breeding values are transformed back to the original scale of measurement