Lecture 3
Overview and Derivation of the Mixed Model

Guilherme J. M. Rosa

University of Wisconsin-Madison
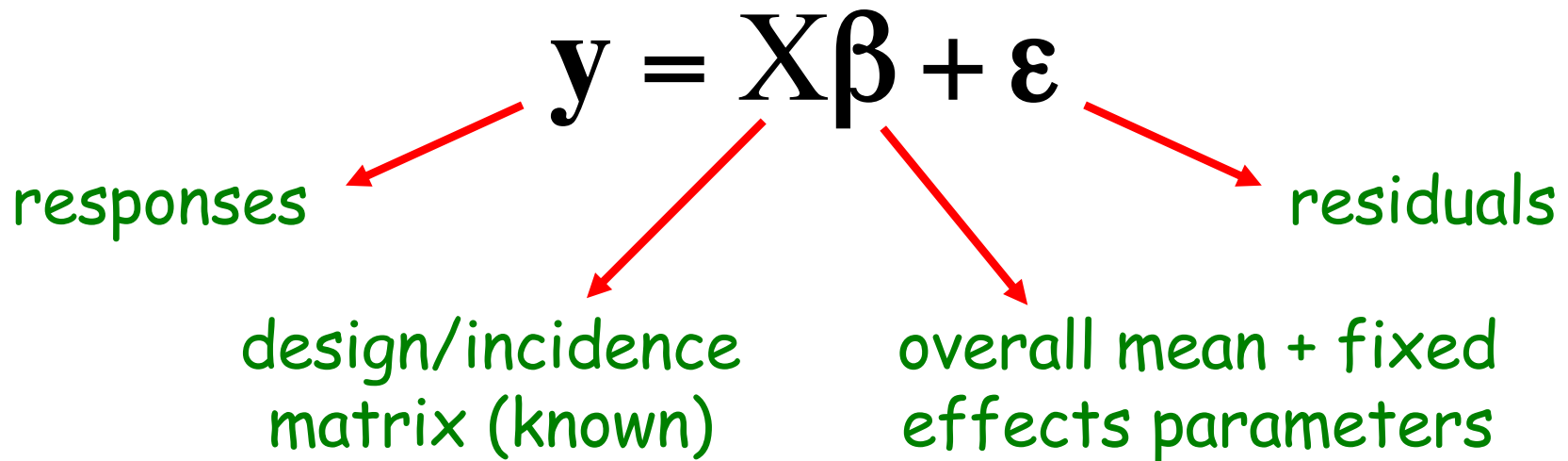
Mixed Models in Quantitative Genetics
SISG, Seattle
22 – 24 July 2020

General Linear Model
(Fixed Effects Model)

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

responses

residuals

design/incidence matrix (known)

overall mean + fixed effects parameters

$$\boldsymbol{\varepsilon} \sim N(\mathbf{0}, I_n \sigma^2) \quad \rightarrow \quad \varepsilon_i \overset{iid}{\sim} N(0, \sigma^2)$$

⇨ **Fixed effect:** levels included in the study represent all levels about which inference is to be made. **Fixed effects models:** models containing only fixed effects

Least-Squares Estimation

$$y = X\beta + \varepsilon$$

$$\varepsilon \sim (0, I_n \sigma^2) \quad \rightarrow \quad \varepsilon_i \overset{iid}{\sim} (0, \sigma^2)$$

An estimate $(\hat{\beta})$ of the vector β can be obtained by the method of least-squares, which aims to minimize the residual sum of squares, given (in matrix notation) by:

$$RSS = \sum_{i=1}^{n} (\hat{\varepsilon}_i)^2 = \hat{\varepsilon}^T \hat{\varepsilon} = (y - X\hat{\beta})^T (y - X\hat{\beta})$$

Taking the derivatives and equating to zero, it can be shown that the least-squares estimator of β is:

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

➲ It is shown that $E[\hat{\beta}] = \beta$ and $\text{Var}[\hat{\beta}] = (X^T X)^{-1} \sigma^2$

3

Example

Flowering time (days, log scale) of *Brassica napus* according to genotype in specific locus, such as a candidate gene

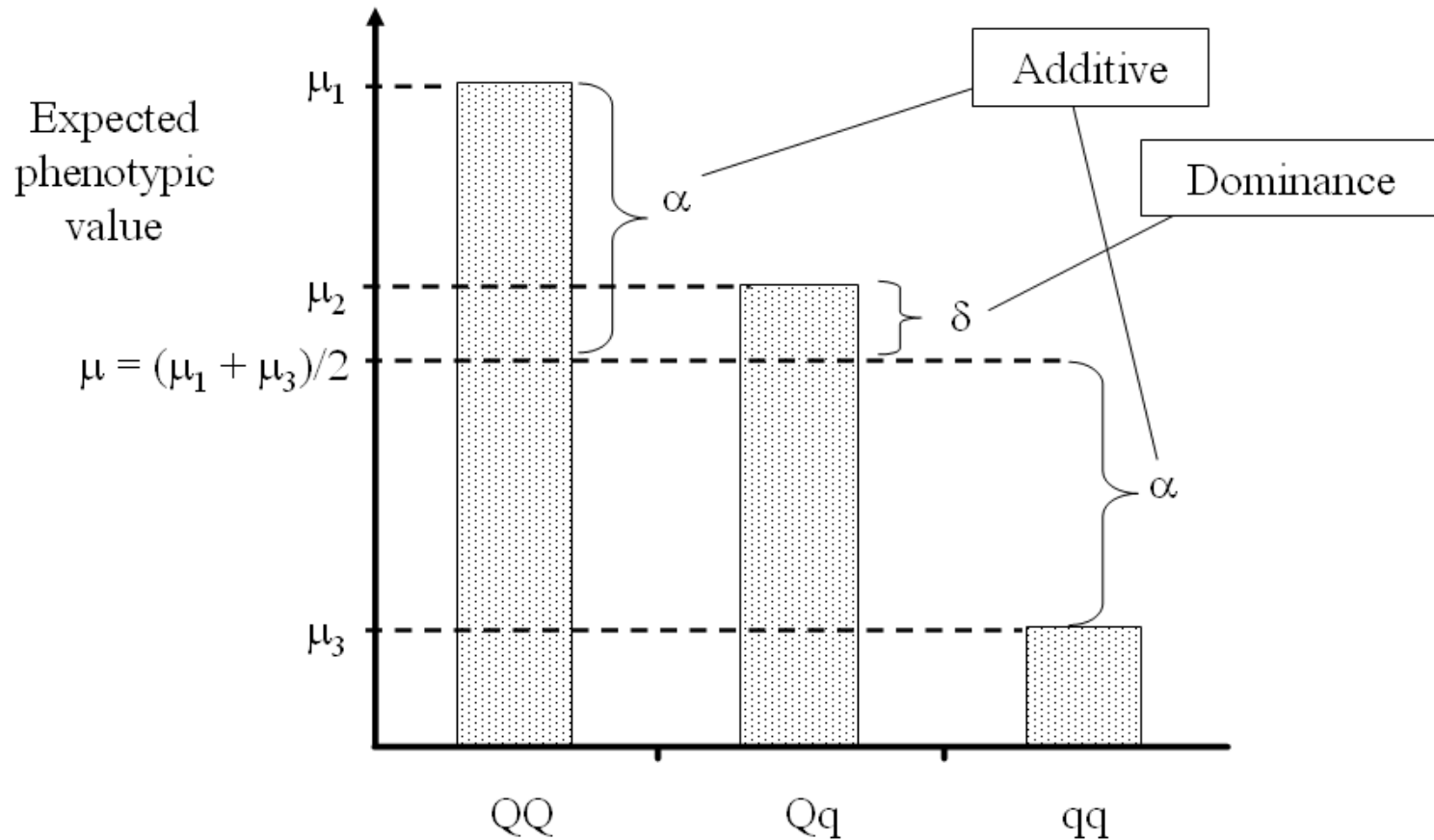| Genotype		
qq	Qq	QQ
3.4	2.9	3.1
3.7	2.5	2.6
3.2		

Model: $y_{ij} = \mu_i + e_{ij}$

y_{ij}: flowering time of replication j (j = 1,..., n_i) of genotype i (i = qq, Qq and QQ)

μ_i: expected flowering time of plants of genotype i

e_{ij}: residual (environment and polygenic effects)

4

⇨ The expected phenotypic values μ_i, however, can be expressed as a function of the additive and dominant effects

Expected phenotypic value according to the genotype on a specific locus.

The model can be written then as:

$$y_{ij} = \mu + x_{ij}\alpha + (1 - |x_{ij}|)\delta + e_{ij}$$

μ: constant (mid-point flowering time between homozygous genotypes)

x_{ij}: indicator variable (genotype), coded as -1, 0 and 1 for genotypes qq, Qq and QQ

α and β: additive and dominance effects

In matrix notation:

$$
\begin{bmatrix} y_{11} \\ y_{12} \\ y_{13} \\ y_{21} \\ y_{22} \\ y_{31} \\ y_{32} \end{bmatrix} =
\begin{bmatrix} 3.4 \\ 3.7 \\ 3.2 \\ 2.9 \\ 2.5 \\ 3.1 \\ 2.6 \end{bmatrix} =
\begin{bmatrix} 1 & -1 & 0 \\ 1 & -1 & 0 \\ 1 & -1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix}
\begin{bmatrix} \mu \\ \alpha \\ \delta \end{bmatrix} +
\begin{bmatrix} e_{11} \\ e_{12} \\ e_{13} \\ e_{21} \\ e_{22} \\ e_{31} \\ e_{32} \end{bmatrix}
$$

6

Analysis of Longitudinal Data

Supposed a series of longitudinal data (e.g., repeated measurements on time) on n individuals.
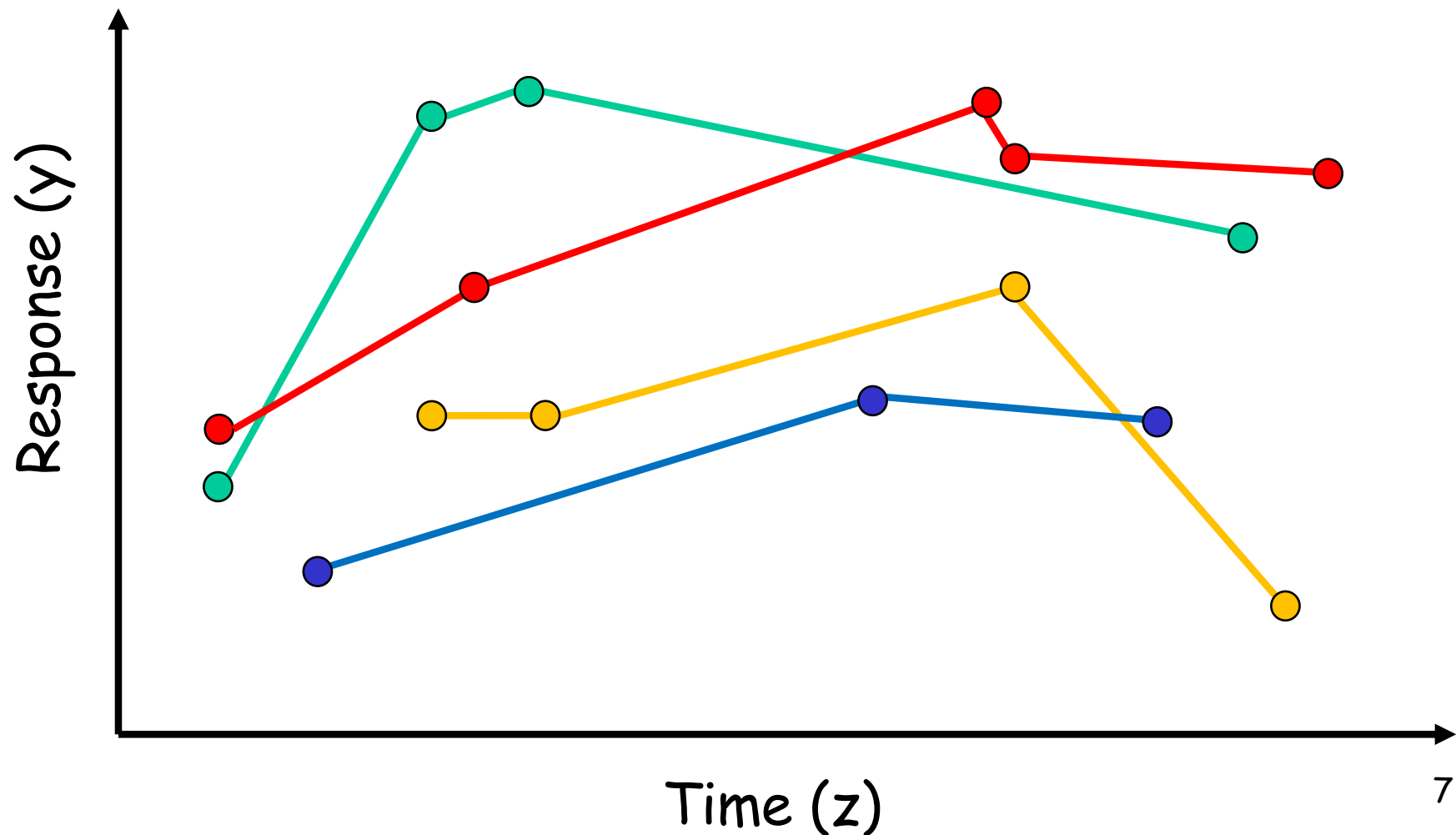
Two-stage Analysis of Longitudinal Data
Step 1

Let y_{ij} represent the observation j ($j = 1,2,...,n_i$) on individual i ($i = 1,2,...,n$), and the following quadratic regression of measurements on time (z_{ij}) for each individual:

$$y_{ij} = \beta_{0i} + \beta_{1i} z_{ij} + \beta_{2i} z_{ij}^2 + \varepsilon_{ij}$$

where β_{0i}, β_{1i} and β_{2i} are subject-specific regression parameters, and ε_{ij} are residual terms, assumed normally distributed with mean zero and variance σ_ε^2

8

In matrix notation such subject-specific regressions can be expressed as:

$$\mathbf{y}_i = \mathbf{Z}_i \boldsymbol{\beta}_i + \boldsymbol{\varepsilon}_i \qquad (1)$$

where $\mathbf{y}_i = (y_{i1}, y_{i2}, \ldots, y_{in_i})^T$, $\boldsymbol{\beta}_i = (\beta_{0i}, \beta_{1i}, \beta_{2i})^T$,

$\boldsymbol{\varepsilon}_i = (\varepsilon_{i1}, \varepsilon_{i2}, \ldots, \varepsilon_{in_i})^T \sim N(\mathbf{0}, \mathbf{I}\sigma_\varepsilon^2)$ and

$$\mathbf{Z}_i = \begin{bmatrix} 1 & z_{i1} & z_{i1}^2 \\ 1 & z_{i2} & z_{i2}^2 \\ \vdots & \vdots & \vdots \\ 1 & z_{in_i} & z_{in_i}^2 \end{bmatrix}$$

Under these specifications, it is shown that the least-squares estimate of β_i is:

$$\hat{\boldsymbol{\beta}}_i = (\mathbf{Z}_i^{\mathrm{T}}\mathbf{Z}_i)^{-1}\mathbf{Z}_i^{\mathrm{T}}\mathbf{y}_i$$

Note that this is also the maximum likelihood estimate of β_i

Such estimates can be viewed as summary statistics for the longitudinal data, the same way one could use area under the curve (AUC), or peak (maximum value of y_{ij}), or mean response.

Two-stage Analysis of Longitudinal Data
Step 2

Supposed now we are interested on the effect of some other variables (such as gender, treatment, year, etc.) on the values of β_i

Such effects could be studied using a model as:

$$\hat{\boldsymbol{\beta}}_i = \mathbf{W}_i\boldsymbol{\beta} + \mathbf{u}_i$$

where $\mathbf{u}_i \sim N(\mathbf{0},\mathbf{D})$, which is an approximation for the model:

$$\boldsymbol{\beta}_i = \mathbf{W}_i\boldsymbol{\beta} + \mathbf{u}_i \qquad (2)$$

Single-stage Analysis of Longitudinal Data

The two step-analysis described here can be merged into a single stage approach by substituting (2) in (1):

$$\mathbf{y}_i = \mathbf{Z}_i[\mathbf{W}_i\boldsymbol{\beta} + \mathbf{u}_i] + \boldsymbol{\varepsilon}_i$$

which can be expressed as:

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

where $\mathbf{X}_i = \mathbf{Z}_i\mathbf{W}_i$. By concatenating observations from multiple individuals, we have the following mixed model:

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

PAUSE

⇨ General Linear Model

⇨ Least Squares Estimation

⇨ Random Effects (Motivation with two-stage longitudinal analysis)

⇨ Next: Formalization of Mixed Models

Next PAUSE, slide 22

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

14

Linear Mixed Effects Model

$$\mathbf{y} = \mathbf{X\beta} + \mathbf{Zu} + \mathbf{e}$$

where:

- $\mathbf{y}$: response vector; observations
- $\mathbf{\beta}$: vector of fixed effects
- $\mathbf{u}$: vector of random effects; $\mathbf{u} \sim N(\mathbf{0}, \boldsymbol{G})$
- $\mathbf{X}$ and $\mathbf{Z}$: (known) incidence matrices
- $\boldsymbol{e}$: residual vector; $\boldsymbol{e} \sim N(\mathbf{0}, \Sigma)$

Linear Mixed Effects Model

Generally, it is assumed that **u** and **e** are independent from each other, such that:

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

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 $\boldsymbol{\varepsilon} = \mathbf{Z}\mathbf{u} + \mathbf{e}$

$$\begin{cases} \mathrm{E}[\boldsymbol{\varepsilon}] = \mathrm{E}[\mathbf{Z}\mathbf{u} + \mathbf{e}] = \mathbf{Z}\mathrm{E}[\mathbf{u}] + \mathrm{E}[\mathbf{e}] = \mathbf{0} \\[2mm] \mathrm{Var}[\boldsymbol{\varepsilon}] = \mathrm{Var}[\mathbf{Z}\mathbf{u} + \mathbf{e}] = \mathbf{Z}\mathrm{Var}[\mathbf{u}]\mathbf{Z}^{\mathrm{T}} + \mathrm{Var}[\mathbf{e}] = \mathbf{Z}\mathbf{G}\mathbf{Z}^{\mathrm{T}} + \boldsymbol{\Sigma} \end{cases}$$

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

Under these circumstances, the MLE for β is:

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

Estimation of Fixed Effects

As **G** and Σ are generally unknown, an estimate of **V** is used instead such that the estimator becomes:

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

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

Note: $(\mathbf{X}^{\mathrm{T}}\hat{\mathbf{V}}^{-1}\mathbf{X})^{-1}$ 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 $\mathbf{K}^{\mathrm{T}}\boldsymbol{\beta}$ can be obtained by using the result:

$$\frac{(\mathbf{K}^{\mathrm{T}}\boldsymbol{\beta}^{0})^{\mathrm{T}}(\mathbf{K}^{\mathrm{T}}(\mathbf{X}^{\mathrm{T}}\mathbf{V}^{-1}\mathbf{X})^{-}\mathbf{K})^{-1}(\mathbf{K}^{\mathrm{T}}\boldsymbol{\beta}^{0})}{\operatorname{rank}(\mathbf{K})} \approx \mathrm{F}_{[\varphi_{\mathrm{N}},\varphi_{\mathrm{D}}]}$$

where $\mathrm{F}_{[\varphi_{\mathrm{N}},\varphi_{\mathrm{D}}]}$ refers to an F-distribution with $\varphi_{\mathrm{N}} = \operatorname{rank}(\mathbf{K})$ 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 **u** given the data, i.e.

$$E[\mathbf{u}\,|\,\mathbf{y}]$$

Given the model specifications, the joint distribution of **y** and **u** is:

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

Estimation (Prediction) of Random Effects

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

$$E[\mathbf{u} \mid \mathbf{y}] = E[\mathbf{u}] + \text{Cov}[\mathbf{u}, \mathbf{y}^T]\text{Var}^{-1}[\mathbf{y}](\mathbf{y} - E[\mathbf{y}])$$

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

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

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

PAUSE

⇨ Mixed Models

⇨ Fixed vs. Random Effects

⇨ BLUE and BLUP

⇨ Next: Another way to get BLUE and BLUP

Next PAUSE, slide 29 (end)

Mixed Model Equations

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

As **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 **β** and **u** simultaneously, without the need for computing $\mathbf{V}^{-1}$

The MME were derived by maximizing (for **β** and **u**) the joint density of **y** and **u**, expressed as:

$$p(\mathbf{y},\mathbf{u}\,|\,\boldsymbol{\beta},\mathbf{G},\boldsymbol{\Sigma}) \propto \;|\boldsymbol{\Sigma}|^{-1/2}|\mathbf{G}|^{-1/2}$$

$$\times \exp\left\{-\frac{1}{2}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta}-\mathbf{Z}\mathbf{u})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta}-\mathbf{Z}\mathbf{u})-\frac{1}{2}\mathbf{u}^{\mathrm{T}}\mathbf{G}^{-1}\mathbf{u}\right\}$$

23

Mixed Model Equations

The logarithm of this function is:

$$\ell = \log[\mathrm{p}(\mathbf{y}, \mathbf{u} \mid \boldsymbol{\beta}, \mathbf{G}, \boldsymbol{\Sigma})] \propto |\boldsymbol{\Sigma}| + |\mathbf{G}| + (\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\mathbf{u})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\mathbf{u}) + \mathbf{u}^{\mathrm{T}} \mathbf{G}^{-1} \mathbf{u}$$

$$= |\boldsymbol{\Sigma}| + |\mathbf{G}| + \mathbf{y}^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} \mathbf{y} - 2\mathbf{y}^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} \mathbf{X}\boldsymbol{\beta} - 2\mathbf{y}^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} \mathbf{Z}\mathbf{u}$$

$$+ \boldsymbol{\beta}^{\mathrm{T}} \mathbf{X}^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} \mathbf{X}\boldsymbol{\beta} + 2\boldsymbol{\beta}^{\mathrm{T}} \mathbf{X}^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} \mathbf{Z}\mathbf{u} + \mathbf{u}^{\mathrm{T}} \mathbf{Z}^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} \mathbf{Z}\mathbf{u} + \mathbf{u}^{\mathrm{T}} \mathbf{G}^{-1} \mathbf{u}$$

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

$$\begin{bmatrix} \dfrac{\partial \ell}{\partial \boldsymbol{\beta}} \\[2ex] \dfrac{\partial \ell}{\partial \mathbf{u}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} \mathbf{y} - \mathbf{X}^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} \mathbf{X}\hat{\boldsymbol{\beta}} - \mathbf{X}^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} \mathbf{Z}\hat{\mathbf{u}} \\[2ex] \mathbf{Z}^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} \mathbf{y} - \mathbf{Z}^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} \mathbf{X}\hat{\boldsymbol{\beta}} - \mathbf{Z}^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} \mathbf{Z}\hat{\mathbf{u}} - \mathbf{G}^{-1}\hat{\mathbf{u}} \end{bmatrix}$$

Mixed Model Equations

Equating them to zero gives the following system:

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

which can be expressed as:

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

known as the mixed model equations (MME)

BLUE and BLUP

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

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

so that:

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

It can be shown that this expression is equivalent to:

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

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:

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

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

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

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

BLUE and BLUP

It is important to note that $\hat{\boldsymbol{\beta}}$ and $\hat{\mathbf{u}}$ 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 ($\hat{\mathbf{G}}$ and $\hat{\Sigma}$) into the MME:

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

Estimation of Variance Components

BLUE and BLUP require knowledge of G and Σ

These matrices, however, are rarely known and must be estimated

Variance and covariance components estimation:

- Analysis of Variance (ANOVA)

- Maximum likelihood

- Restricted maximum likelihood (REML)

- Bayesian approach