# Genetic Covariances Between Animals

Peter von Rohr

2021-11-05

# BLUP Animal Model

- Traditional prediction of breeding values
  - i.e. no genomic information available
  - de-facto standard method
- Model recap

$$y = X\beta + Zu + e$$

where

| | |
|---|---|
| $y$ | vector of length $n$ of phenotypic information |
| $\beta$ | vector of length $p$ of unknown fixed effects |
| $X$ | $n \times p$ incidence matrix |
| $u$ | vector of length $q$ of unknown random breeding values |
| $Z$ | $n \times q$ incidence matrix |
| $e$ | vector of length $n$ of unknown random residuals |

# Example

| Calf | Sire | Dam | Herd | WWG |
|---:|---:|---:|---:|---:|
| 3 | 1 | 2 | 1 | 4.5 |
| 4 | 1 | NA | 2 | 2.9 |
| 5 | 4 | 3 | 2 | 3.9 |
| 6 | 5 | 2 | 1 | 3.5 |

# Animal Model Setup

- Convert information from data into model
- Put observations in vector $y$

$$y = \begin{bmatrix} 4.5 \\ 2.9 \\ 3.9 \\ 3.5 \end{bmatrix}$$

- Fixed effects such as `herd` go in vector $\beta$ and $X$ links observations to components in $\beta$

$$X = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}, \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}$$

# Breeding Values As Random Effects and Residuals

$$Z = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix},\ u = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \\ u_6 \end{bmatrix},\ e = \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \end{bmatrix}$$

# Putting Everything Together

$$\begin{bmatrix} 4.5 \\ 2.9 \\ 3.9 \\ 3.5 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \\ u_6 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \end{bmatrix}$$

# Solution with Mixed Model Equations

- General form

$$\left[\begin{array}{lr} X^T R^{-1} X & X^T R^{-1} Z \\ Z^T R^{-1} X & Z^T R^{-1} Z + G^{-1} \end{array}\right] \left[\begin{array}{c} \hat{\beta} \\ \hat{u} \end{array}\right] = \left[\begin{array}{c} X^T R^{-1} y \\ Z^T R^{-1} y \end{array}\right]$$

- Assume $R = I * \sigma_e^2$ and we set $G = A * \sigma_u^2$

$$\left[\begin{array}{lr} X^T X & X^T Z \\ Z^T X & Z^T Z + \lambda * A^{-1} \end{array}\right] \left[\begin{array}{c} \hat{\beta} \\ \hat{u} \end{array}\right] = \left[\begin{array}{c} X^T y \\ Z^T y \end{array}\right]$$

with $\lambda = \sigma_e^2 / \sigma_u^2$

# Components of Mixed Model Equations

$$X^TX = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, X^TZ = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 \end{bmatrix}$$

$$Z^TZ = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, X^Ty = \begin{bmatrix} 8 \\ 6.8 \end{bmatrix}, Z^Ty = \begin{bmatrix} 0 \\ 0 \\ 4.5 \\ 2.9 \\ 3.9 \\ 3.5 \end{bmatrix}$$

# Numerator Relationship Matrix

- Only component missing in mixed model equations: matrix $A$
- $A$ is called the **numerator relationship matrix**
- Properties of $A$

$$var(u) = G = A * \sigma_u^2$$

# Meaning of *var*() for scalar variable $x$

- Depending on type of argument to *var*(), the result is different
- Scalar argument $x$, that means $x$ is just a single number

$$var(x) = \sum (x - E[x])^2\, f(x)$$

for a discrete random variable $x$, e.g. genotypic values $V$ in single locus model.

$$var(x) = \int (x - E[x])^2\, f(x) dx$$

for a continuous random variable $x$.

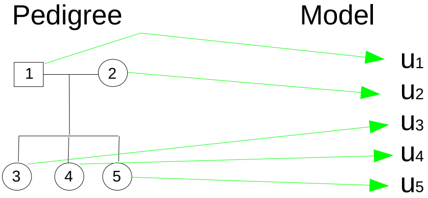
# Meaning of *var*() for a vector $u$

- Assume that $u$ is a vector with $q$ elements

$$u = \begin{bmatrix} u_1 \\ u_2 \\ \dots \\ u_q \end{bmatrix}$$

- Then $var(u)$ is a $q \times q$ variance-covariance matrix with a structure shown on the next slide

# Meaning of $A$



Variance-Covariance

| | $u_1$ | $u_2$ | $u_3$ | $u_4$ | $u_5$ |
|---|---|---|---|---|---|
| $u_1$ | $\mathrm{var}(u_1)$ | $\mathrm{cov}(u_1,u_2)$ | $\mathrm{cov}(u_1,u_3)$ | $\mathrm{cov}(u_1,u_4)$ | $\mathrm{cov}(u_1,u_5)$ |
| $u_2$ | | $\mathrm{var}(u_2)$ | $\mathrm{cov}(u_2,u_3)$ | $\mathrm{cov}(u_2,u_4)$ | $\mathrm{cov}(u_2,u_5)$ |
| $u_3$ | | | $\mathrm{var}(u_3)$ | $\mathrm{cov}(u_3,u_4)$ | $\mathrm{cov}(u_3,u_5)$ |
| $u_4$ | | | | $\mathrm{var}(u_4)$ | $\mathrm{cov}(u_4,u_5)$ |
| $u_5$ | | | | | $\mathrm{var}(u_5)$ |

# Elements of $A$

- Remember: $var(u) = G = A * \sigma_u^2$
- Diagonal elements $(A)_{ii} = 1 + F_i$ where $F_i$: inbreeding coefficient of $i$
- Offdiagonal elements $(A)_{ij} = cov(u_i, u_j)/\sigma_u^2$
- Symmetric
- Definition leads to recursive algorithm for computing components of $A$

# Example

Elements of $G$ are computed as

- Animal 1 has unknown parents and is assumed to show no inbreeding (parents are not related)

$$var(u_1) = (1 + F_1) * \sigma_u^2 = \sigma_u^2$$

- Animal 2 has unknown parents and is not related to animal 1

$$cov(u_1, u_2) = 0$$

- Animal 3 has parents 1 and 2

$$cov(u_1, u_3) = cov\left(u_1, \left[\frac{1}{2}(u_1 + u_2) + m_3\right]\right) = \frac{1}{2}\sigma_u^2$$