



Centre for
Tropical Livestock
Genetics and Health

Pedigree-based genetic evaluations

Chrissy Rochus

Roslin Institute, University of Edinburgh

Slides adapted from Gregor Gorjanc, Gabriela Mafra Fortuna and Ivan Pocrnić





Learning objectives

Understand how to combine phenotype information from all relatives connected via pedigree

Familiarise yourself with linear mixed models and equations

Practice inference of breeding values with the pedigree- based model

- simple cases using R matrix algebra
- using other packages



Centre for
Tropical Livestock
Genetics and Health

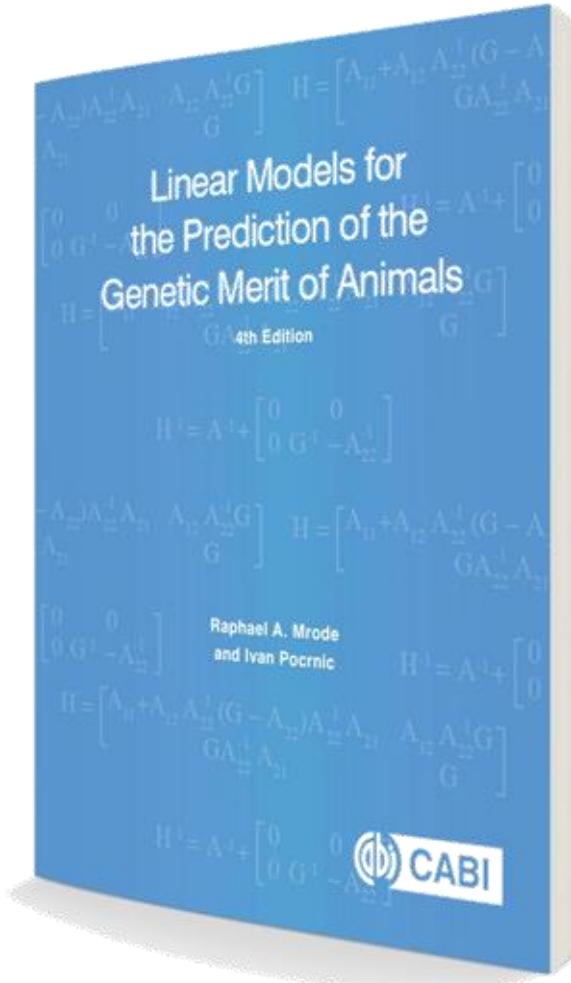
Linear models for the prediction of the genetic merit of animals

CABI Biotechnology Series

September 2023 | 412pp

Raphael A Mrode
Ivan Pocrnic

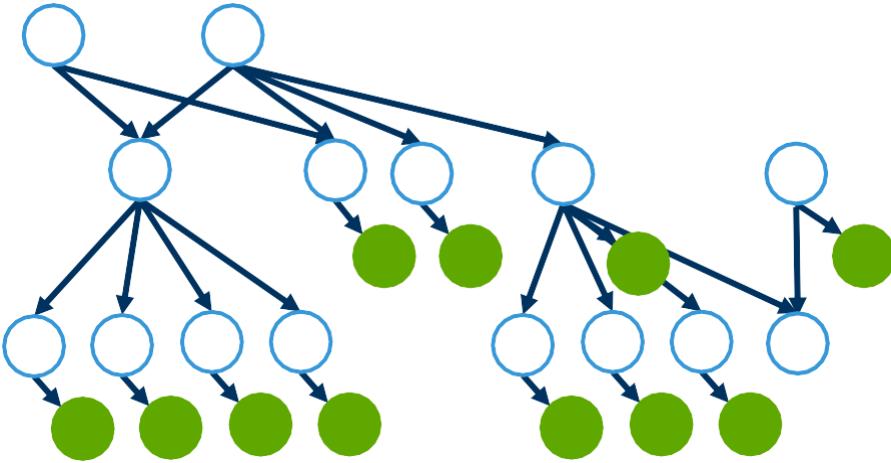
Robin Thompson
Gregor Gorjanc



See chapters
3 & 4!



General setting

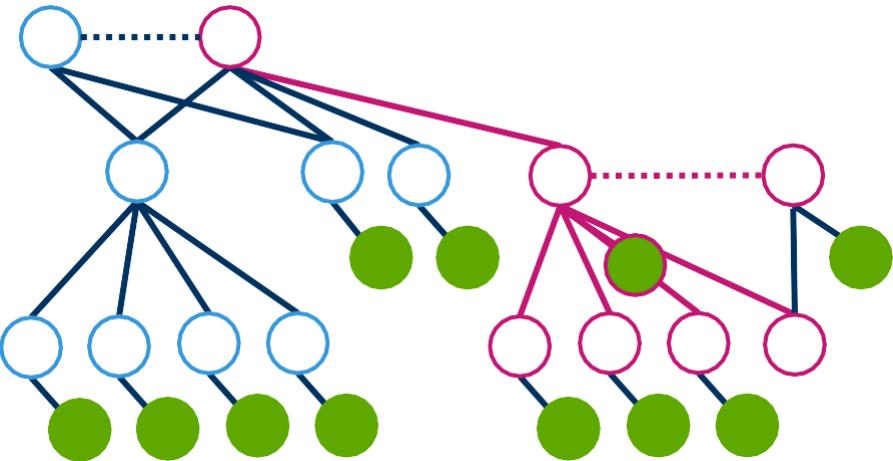


- Evaluation/Estimation
 - ancestors phenotypes
 - own phenotypes
 - sib... phenotypes
 - descendants phenotypes
- Prediction
 - progeny
- Population/group means



Centre for
Tropical Livestock
Genetics and Health

Information for an individual

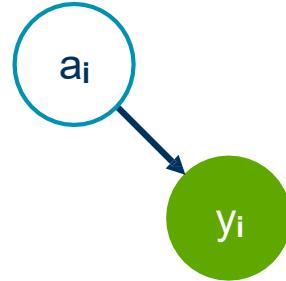




Pedigree-based model

Phenotype model

$$y_i = \mu + a_i + e_i$$
$$e_i \sim N(0, \sigma_e^2)$$



Pedigree model

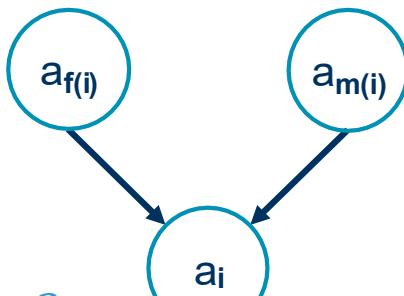
$$a_i \sim N(0, \sigma_a^2)$$

$$a_i = \frac{1}{2}a_{f(i)} + \frac{1}{2}a_{m(i)} + r_i$$

$$a_i | a_{f(i)}, a_{m(i)} \sim N\left(\frac{1}{2}a_{f(i)} + \frac{1}{2}a_{m(i)}, k\sigma_a^2\right)$$

$$r_i \sim N(0, k\sigma_a^2)$$

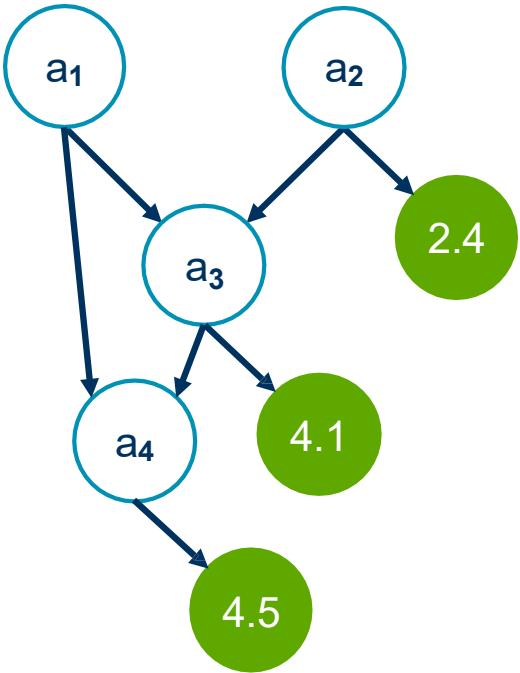
$$k = \frac{1}{2} - \frac{1}{4}(F_{f(i)} + F_{m(i)})$$



Galton (1886)
Wright (1920+)



Pedigree-based model - example



Phenotype model

$$y_2 = 2.4 = \mu + a_2 + e_2$$

$$y_3 = 4.1 = \mu + a_3 + e_3 \quad e_i \sim N(0, \sigma_e^2)$$

$$y_4 = 4.5 = \mu + a_4 + e_4$$

Pedigree model

$$a_1 = r_1$$

$$r_i \sim N(0, k\sigma_a^2)$$

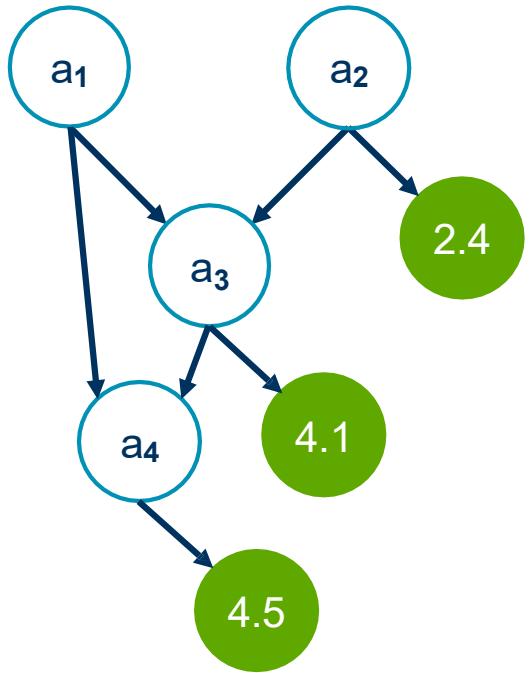
$$a_2 = r_2$$

$$a_3 = \frac{1}{2}a_1 + \frac{1}{2}a_2 + r_3 = \frac{1}{2}r_1 + \frac{1}{2}r_2 + r_3$$

$$\begin{aligned} a_4 &= \frac{1}{2}a_1 + \frac{1}{2}a_3 + r_4 = \frac{1}{2}r_1 + \frac{1}{2}\left(\frac{1}{2}r_1 + \frac{1}{2}r_2 + r_3\right) + r_4 \\ &= \frac{3}{4}r_1 + \frac{1}{4}r_2 + \frac{1}{2}r_3 + r_4 \end{aligned}$$



Pedigree-based model - example



Phenotype model

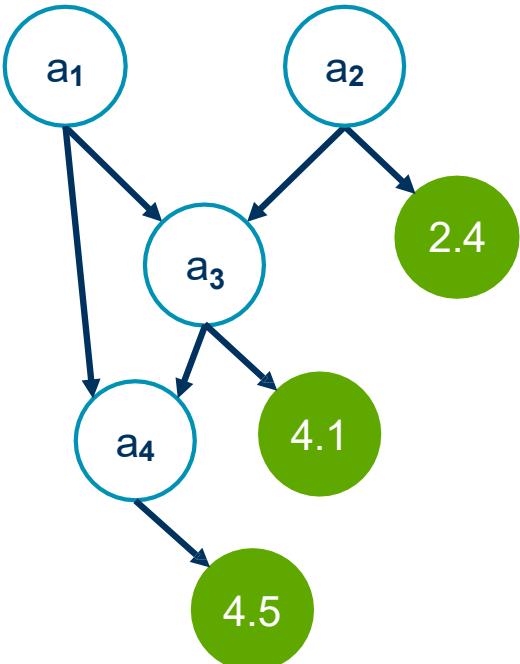
$$\begin{pmatrix} y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} 2.4 \\ 4.1 \\ 4.5 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}(\mu) + \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix} + \begin{pmatrix} e_2 \\ e_3 \\ e_4 \end{pmatrix}$$

$$y = Xb + Za + e$$

$$e \sim N(\mathbf{0}, E\sigma_e^2)$$



Pedigree-based model - example



Phenotype model

$$\begin{pmatrix} y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} 2.4 \\ 4.1 \\ 4.5 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}(\mu) + \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix} + \begin{pmatrix} e_2 \\ e_3 \\ e_4 \end{pmatrix}$$

$$y = Xb + Za + e$$

$$e \sim N(\mathbf{0}, E\sigma_e^2)$$

Pedigree model

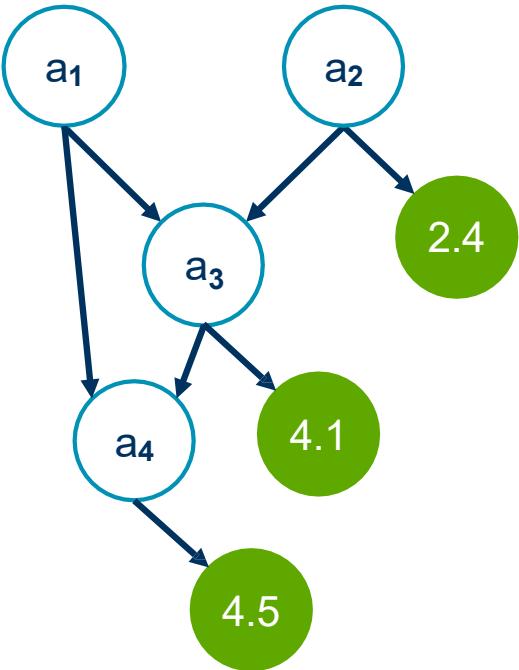
$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 1 & 0 \\ \frac{3}{4} & \frac{1}{4} & \frac{1}{2} & 1 \end{pmatrix} \begin{pmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \end{pmatrix}$$

$$a = Tr$$

$$r \sim N(\mathbf{0}, R\sigma_a^2)$$



Pedigree-based model - example



Phenotype model

$$\begin{pmatrix} y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} 2.4 \\ 4.1 \\ 4.5 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}(\mu) + \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix} + \begin{pmatrix} e_2 \\ e_3 \\ e_4 \end{pmatrix}$$

$$y = Xb + Za + e$$

$$e \sim N(0, E\sigma_e^2)$$

Pedigree model

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 1 & 0 \\ \frac{3}{4} & \frac{1}{4} & \frac{1}{2} & 1 \end{pmatrix} \begin{pmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \end{pmatrix}$$

$$a = Tr$$

$$r \sim N(\mathbf{0}, R\sigma_a^2)$$

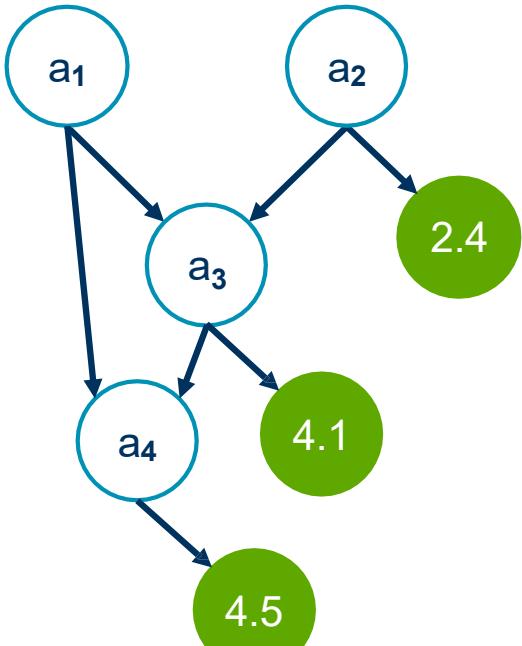
PHENO. POPULATION
(often not clearly definable)

$$\sigma^2$$

BASE POPULATION!!!!
(often not clearly definable)



Pedigree-based model - example



Phenotype model

$$\begin{pmatrix} y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} 2.4 \\ 4.1 \\ 4.5 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}(\mu) + \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix} + \begin{pmatrix} e_2 \\ e_3 \\ e_4 \end{pmatrix}$$

$$y = Xb + Za + e$$

$$e \sim N(\mathbf{0}, E\sigma_e^2)$$

Pedigree model

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 1 & 0 \\ \frac{3}{4} & \frac{1}{4} & \frac{1}{2} & 1 \end{pmatrix} \begin{pmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \end{pmatrix}$$

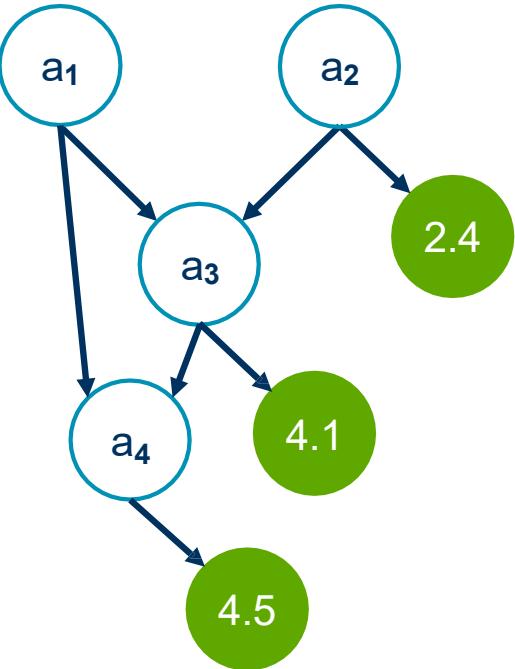
$$a = Tr$$

$$r \sim N(\mathbf{0}, R\sigma_a^2)$$

$$\begin{aligned} Var(a) &= Var(Tr) \\ &= TVar(r)T^T \\ &= TRT^T\sigma_a^2 \\ &= A\sigma_a^2 \end{aligned}$$



Pedigree-based relationship matrix



$$Var(\mathbf{a}) = \mathbf{A}\sigma_a^2$$

- Elements of \mathbf{A} are covariance coefficients!
- Wright's relationships were correlations
 - $R_{i,j} = A_{1,2}/\sqrt{A_{i,i} A_{j,j}}$
 - $\mathbf{A} \rightarrow$ Numerator Relationship Matrix (NRM)
- $A_{i,i}=1+F_i$, F_i – **inbreeding coefficient**
(are alleles of the individual i IBD)
- $F_i=K_{f(i),m(i)}=1/2A_{f(i),m(i)}$
- $A_{i,j}=2K_{i,j}$, $K_{i,j}$ – **kinship coefficient**
(are alleles of the individuals i and j IBD)
- $A_{i,j}=1/2A_{i,f(j)}+1/2A_{i,m(j)}$
- $\mathbf{A}=2\mathbf{K}$



Pedigree-based model

- The usual notation for pedigree-based linear mixed model

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{Z}\mathbf{a} + \mathbf{e}$$

$$\mathbf{e} \sim N(\mathbf{0}, E\sigma_e^2)$$

$$\mathbf{a} \sim N(\mathbf{0}, A\sigma_a^2)$$

- Estimator/Predictor (summarising the conditional distribution):

$$E(\mathbf{a}|\mathbf{y}) = \hat{\mathbf{a}} = Cov(\mathbf{a}, \mathbf{y})Var(\mathbf{y})^{-1}(\mathbf{y} - E(\mathbf{y}))$$

$$V(\mathbf{a}|\mathbf{y}) = Var(\mathbf{a}) - Cov(\mathbf{a}, \mathbf{y})Var(\mathbf{y})^{-1}Cov(\mathbf{y}, \mathbf{a})$$

$$Cor(\mathbf{a}, \hat{\mathbf{a}})^2 = \mathbf{1} - Var(\mathbf{a}|\mathbf{y})/Var(\mathbf{a})$$



Pedigree-based model

- The usual notation for pedigree-based linear mixed model

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{Z}\mathbf{a} + \mathbf{e}$$

$$\mathbf{e} \sim N(\mathbf{0}, \mathbf{E}\sigma_e^2)$$

$$\mathbf{a} \sim N(\mathbf{0}, \mathbf{A}\sigma_a^2)$$

- Estimator/Predictor (summarising the conditional distribution):

$$\begin{pmatrix} \mathbf{X}^T \mathbf{E}^{-1} \mathbf{X} & \mathbf{X}^T \mathbf{E}^{-1} \mathbf{Z} \\ \mathbf{Z}^T \mathbf{E}^{-1} \mathbf{X} & \mathbf{Z}^T \mathbf{E}^{-1} \mathbf{Z} + \mathbf{A}^{-1} \frac{\sigma_e^2}{\sigma_a^2} \end{pmatrix} \begin{pmatrix} \hat{\mathbf{b}} \\ \hat{\mathbf{a}} \end{pmatrix} = \begin{pmatrix} \mathbf{X}^T \mathbf{E}^{-1} \mathbf{y} \\ \mathbf{Z}^T \mathbf{E}^{-1} \mathbf{y} \end{pmatrix}$$

$$\mathbf{C} \begin{pmatrix} \hat{\mathbf{b}} \\ \hat{\mathbf{a}} \end{pmatrix} = \begin{pmatrix} \mathbf{X}^T \mathbf{E}^{-1} \mathbf{y} \\ \mathbf{Z}^T \mathbf{E}^{-1} \mathbf{y} \end{pmatrix}$$

$$Var(\mathbf{a}|\mathbf{y}) = diag(\mathbf{C}^{-1})_{\mathbf{a}} \sigma_e^2$$



Henderson's mixed model equations

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{Z}\mathbf{a} + \mathbf{e}$$

$$\begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{b}} \\ \hat{\mathbf{a}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \end{bmatrix}$$

$$\hat{\mathbf{b}} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y}$$

$$\begin{aligned} \hat{\mathbf{a}} &= \mathbf{G}\mathbf{Z}'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\hat{\mathbf{b}}) \\ &= \mathbf{G}\mathbf{Z}'(\mathbf{Z}\mathbf{G}\mathbf{Z}' + \mathbf{R})^{-1}(\mathbf{y} - \mathbf{X}\hat{\mathbf{b}}) \end{aligned}$$



Henderson's mixed model equations (simplification)

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{Z}\mathbf{a} + \mathbf{e}$$

$$Var(\mathbf{a}) = \mathbf{G} = \sigma_a^2 \mathbf{A}$$
$$Var(\mathbf{e}) = \mathbf{R} = \sigma_e^2 \mathbf{I}$$

$$\begin{bmatrix} \mathbf{X}' \mathbf{R}^{-1} \mathbf{X} & \mathbf{X}' \mathbf{R}^{-1} \mathbf{Z} \\ \mathbf{Z}' \mathbf{R}^{-1} \mathbf{X} & \mathbf{Z}' \mathbf{R}^{-1} \mathbf{Z} + \mathbf{G}^{-1} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{b}} \\ \hat{\mathbf{a}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}' \mathbf{R}^{-1} \mathbf{y} \\ \mathbf{Z}' \mathbf{R}^{-1} \mathbf{y} \end{bmatrix}$$

$$\begin{bmatrix} \mathbf{X}' \frac{1}{\sigma_e^2} \mathbf{I} \mathbf{X} & \mathbf{X}' \frac{1}{\sigma_e^2} \mathbf{I} \mathbf{Z} \\ \mathbf{Z}' \frac{1}{\sigma_e^2} \mathbf{I} \mathbf{X} & \mathbf{Z}' \frac{1}{\sigma_e^2} \mathbf{I} \mathbf{Z} + \frac{1}{\sigma_a^2} \mathbf{A}^{-1} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{b}} \\ \hat{\mathbf{a}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}' \frac{1}{\sigma_e^2} \mathbf{I} \mathbf{y} \\ \mathbf{Z}' \frac{1}{\sigma_e^2} \mathbf{I} \mathbf{y} \end{bmatrix}$$

$$\begin{bmatrix} \mathbf{X}' \mathbf{X} & \mathbf{X}' \mathbf{Z} \\ \mathbf{Z}' \mathbf{X} & \mathbf{Z}' \mathbf{Z} + \frac{\sigma_e^2}{\sigma_a^2} \mathbf{A}^{-1} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{b}} \\ \hat{\mathbf{a}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}' \mathbf{y} \\ \mathbf{Z}' \mathbf{y} \end{bmatrix}$$



Properties of BLUP

- Best: Maximises $\text{Cor}(a, \hat{a})$ or minimises $\text{Var}(a - \hat{a})$
- Linear: Predictors are linear functions of observations
- Unbiased: $E(\hat{a}) = a$
- Prediction: Predicts true random effects (e.g., breeding value)



Main benefits of BLUP

- Handles unbalanced data
- Combines information from all relatives
- We obtain EBV for all individuals
- Estimates fixed and random effects simultaneously
- Easily expandable to various biological assumptions and quantitative genetics models
- But ...



Main benefits of BLUP

- Handles unbalanced data
- Combines information from all relatives
- We obtain EBV for all individuals
- Estimates fixed and random effects simultaneously
- Easily expandable to various biological assumptions and quantitative genetics models
- But ...
- BLUP requires genetic connectivity between the management units



Computational intricacies

- In this course, we will solve MME by direct inversion
- In real-world applications, due to the size of data, specialised algorithms and computational techniques are used
- Direct solvers: Gaussian elimination, LU decomposition, Cholesky decomposition, ...
- Iterative solvers: Gauss-Seidel iteration, successive over-relaxation (SOR), Jacobi iteration, preconditioned conjugate gradient (PCG), ...



$$y = Xb + Za + e$$

Known additive variance $a \sim N(0, \sigma_a^2 A)$ Pedigree-based numerator relationship matrix

Known residual variance $e \sim N(0, \sigma_e^2 I)$

$$\begin{bmatrix} X' X & X' Z \\ Z' X & Z' Z + \alpha_1 A^{-1} \end{bmatrix} \begin{bmatrix} \hat{b} \\ \hat{a} \end{bmatrix} = \begin{bmatrix} X' y \\ Z' y \end{bmatrix}$$

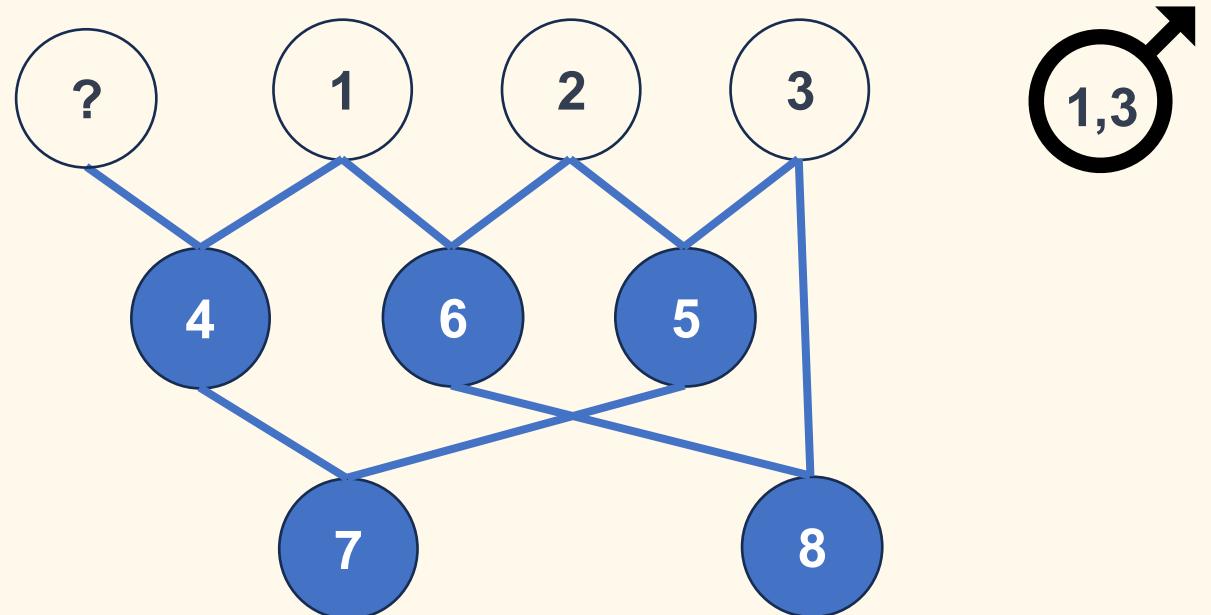
$$\alpha_1 = \frac{\sigma_e^2}{\sigma_a^2} = \frac{1 - h^2}{h^2}$$



Calf	Sex	WWG (kg)
4	Male	4.5
5	Female	2.9
6	Female	3.9
7	Male	3.5
8	Male	5.0

$$\alpha_1 = \frac{\sigma_e^2}{\sigma_a^2} = \frac{40}{20} = 2$$

Animal model (example)





Animal model (example)

$$y = Xb + Za + e$$

$$\begin{bmatrix} 4.5 \\ 2.9 \\ 3.9 \\ 3.5 \\ 5.0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 1 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \widehat{b}_1 \\ \widehat{b}_2 \end{bmatrix}$$

y X

Calf	Sex	WWG (kg)
4	Male	4.5
5	Female	2.9
6	Female	3.9
7	Male	3.5
8	Male	5.0



Animal model (example)

$$y = Xb + Za + e$$

$$\begin{bmatrix} 4.5 \\ 2.9 \\ 3.9 \\ 3.5 \\ 5.0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 1 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \widehat{b}_1 \\ \widehat{b}_2 \end{bmatrix} + \begin{bmatrix} a_1 & a_2 & a_3 & a_4 & a_5 & a_6 & a_7 & a_8 \end{bmatrix} \begin{bmatrix} \widehat{a}_1 \\ \widehat{a}_2 \\ \widehat{a}_3 \\ \widehat{a}_4 \\ \widehat{a}_5 \\ \widehat{a}_6 \\ \widehat{a}_7 \\ \widehat{a}_8 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \end{bmatrix}$$

y X

Z

Calf	WWG (kg)
4	4.5
5	2.9
6	3.9
7	3.5
8	5.0



Animal model (example)

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{Z}\mathbf{a} + \mathbf{e}$$

$$\begin{bmatrix} 4.5 \\ 2.9 \\ 3.9 \\ 3.5 \\ 5.0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 1 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \widehat{\mathbf{b}}_1 \\ \widehat{\mathbf{b}}_2 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \widehat{\mathbf{a}}_1 \\ \widehat{\mathbf{a}}_2 \\ \widehat{\mathbf{a}}_3 \\ \widehat{\mathbf{a}}_4 \\ \widehat{\mathbf{a}}_5 \\ \widehat{\mathbf{a}}_6 \\ \widehat{\mathbf{a}}_7 \\ \widehat{\mathbf{a}}_8 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \end{bmatrix}$$

\mathbf{y} \mathbf{X} \mathbf{Z}

$$\begin{bmatrix} \mathbf{X}' \mathbf{X} & \mathbf{X}' \mathbf{Z} \\ \mathbf{Z}' \mathbf{X} & \mathbf{Z}' \mathbf{Z} + \alpha_1 \mathbf{A}^{-1} \end{bmatrix} \begin{bmatrix} \widehat{\mathbf{b}} \\ \widehat{\mathbf{a}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}' \mathbf{y} \\ \mathbf{Z}' \mathbf{y} \end{bmatrix}$$

$$\begin{bmatrix} \widehat{\mathbf{b}} \\ \widehat{\mathbf{a}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}' \mathbf{X} & \mathbf{X}' \mathbf{Z} \\ \mathbf{Z}' \mathbf{X} & \mathbf{Z}' \mathbf{Z} + \alpha_1 \mathbf{A}^{-1} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{X}' \mathbf{y} \\ \mathbf{Z}' \mathbf{y} \end{bmatrix}$$



Animal model (example)

$$y = Xb + \mathbf{Z}a + e$$

$$\begin{aligned} X'X &= \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix} & X'Z &= \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \end{bmatrix} \\ Z'X &= \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 1 & 0 \\ 1 & 0 \end{bmatrix} & Z'Z &= \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \\ Z'y &= \begin{bmatrix} 0 \\ 0 \\ 0 \\ 4.5 \\ 2.9 \\ 3.9 \\ 3.5 \\ 5.0 \end{bmatrix} & X'y &= \begin{bmatrix} 13.0 \\ 6.8 \end{bmatrix} \\ \begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + \alpha_1 A^{-1} \end{bmatrix} \begin{bmatrix} \hat{b} \\ \hat{a} \end{bmatrix} &= \begin{bmatrix} X'y \\ Z'y \end{bmatrix} \end{aligned}$$



Animal model (example)

$$y = Xb + Za + e$$

$$\begin{bmatrix} \mathbf{X}' \mathbf{X} & \mathbf{X}' \mathbf{Z} \\ \mathbf{Z}' \mathbf{X} & \mathbf{Z}' \mathbf{Z} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{b}} \\ \hat{\mathbf{a}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}' \mathbf{y} \\ \mathbf{Z}' \mathbf{y} \end{bmatrix}$$

$$\begin{bmatrix} 3 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 2 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \hat{b}_1 \\ \hat{b}_2 \\ \hat{a}_1 \\ \hat{a}_2 \\ \hat{a}_3 \\ \hat{a}_4 \\ \hat{a}_5 \\ \hat{a}_6 \\ \hat{a}_7 \\ \hat{a}_8 \end{bmatrix} = \begin{bmatrix} 13.0 \\ 6.8 \\ 0 \\ 0 \\ 0 \\ 4.5 \\ 2.9 \\ 3.9 \\ 3.5 \\ 5.0 \end{bmatrix}$$

Left hand side (LHS)

Right hand side
(RHS)



Animal model (example)

$$y = Xb + Za + e$$

$$\begin{bmatrix} X' X & X' Z \\ Z' X & Z' Z + \alpha_1 A^{-1} \end{bmatrix} \begin{bmatrix} \hat{b} \\ \hat{a} \end{bmatrix} = \begin{bmatrix} X' y \\ Z' y \end{bmatrix}$$

3	0	0	0	1	0	1	1
0	2	0	0	0	1	1	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
1	0	0	0	1	0	0	0
0	1	0	0	0	1	0	0
0	1	0	0	0	0	1	0
1	0	0	0	0	0	0	1

Left hand side (LHS)

$$\begin{bmatrix} \hat{b}_1 \\ \hat{b}_2 \\ \hat{a}_1 \\ \hat{a}_2 \\ \hat{a}_3 \\ \hat{a}_4 \\ \hat{a}_5 \\ \hat{a}_6 \\ \hat{a}_7 \\ \hat{a}_8 \end{bmatrix} = \begin{bmatrix} 13.0 \\ 6.8 \\ 0 \\ 0 \\ 0 \\ 4.5 \\ 2.9 \\ 3.9 \\ 3.5 \\ 5.0 \end{bmatrix}$$

Right hand side
(RHS)



Animal model (example)

$$y = Xb + Za + e$$

$$\begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + \alpha_1 A^{-1} \end{bmatrix} \begin{bmatrix} \hat{b} \\ \hat{a} \end{bmatrix} = \begin{bmatrix} X'y \\ Z'y \end{bmatrix}$$

$$\begin{bmatrix} \hat{b} \\ \hat{a} \end{bmatrix} = \begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + \alpha_1 A^{-1} \end{bmatrix}^{-1} \begin{bmatrix} X'y \\ Z'y \end{bmatrix}$$

$$\begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \hat{b}_1 \\ \hat{b}_2 \\ \hat{a}_1 \\ \hat{a}_2 \\ \hat{a}_3 \\ \hat{a}_4 \\ \hat{a}_5 \\ \hat{a}_6 \\ \hat{a}_7 \\ \hat{a}_8 \end{bmatrix} = \begin{bmatrix} 13.0 \\ 6.8 \\ 0 \\ 0 \\ 0 \\ 4.5 \\ 2.9 \\ 3.9 \\ 3.5 \\ 5.0 \end{bmatrix}$$

Left hand side (LHS)

Right hand side (RHS)



Calf	Sex	WWG (kg)
4	Male	4.5
5	Female	2.9
6	Female	3.9
7	Male	3.5
8	Male	5.0

Animal model (example)

$$\widehat{\text{male}} = 4.3585$$
$$\widehat{\text{female}} = 3.4044$$

$$\begin{aligned}\hat{a}_1 &= +0.0984 \\ \hat{a}_2 &= -0.0188 \\ \hat{a}_3 &= -0.0411 \\ \hat{a}_4 &= -0.0087 \\ \hat{a}_5 &= -0.1857 \\ \hat{a}_6 &= +0.1769 \\ \hat{a}_7 &= -0.2495 \\ \hat{a}_8 &= +0.1826\end{aligned}$$



Pedigree-based evaluations use A matrix to represent how animals are related to each other (relative genetic variances and covariances among individuals)

Can be calculated using tabular method, but we will mostly use software

An example of tabular method

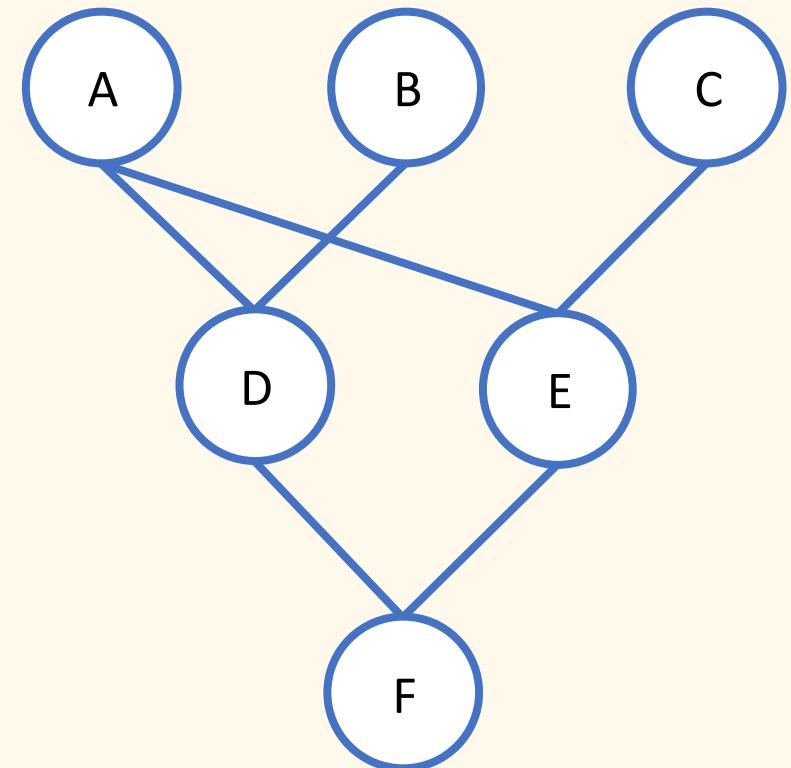


Make a list of all
animals with
observations, include
sire and dam if known

Animals need to be in
order

Animal	Sire	Dam
A		
B		
C		
D	A	B
E	A	C
F	E	D

Genetic relationships





Centre for
Tropical Livestock
Genetics and Health

Genetic relationships

6 animals, 6 by 6 table

	A	B	C	A,B	A,C	E,D
A						
B						
C						
D						
E						
F						



Genetic relationships

For animals with unknown parents, assume unrelated

	A	B	C	A,B	A,C	E,D	F
A	1	0	0				
B	0	1	0				
C	0	0	1				
D							
E							
F							



Genetic relationships

Compute relationships between A
and D, E, F

$$a_{AD} = 0.5(a_{AA} + a_{AB}) = 0.5(1 + 0) = 0.5$$

	A	B	C	A,B	A,C	E,D	F
A	1	0	0				
B	0	1	0				
C	0	0	1				
D							
E							
F							



Genetic relationships

Compute relationships between A
and D, E, F

$$a_{AD} = 0.5(a_{AA} + a_{AB}) = 0.5(1 + 0) = 0.5$$

$$a_{AE} = 0.5(a_{AA} + a_{AC}) = 0.5(1 + 0) = 0.5$$

	A	B	C	A,B	A,C	E,D
A	1	0	0	0.5		
B	0	1	0			
C	0	0	1			
D	0.5					
E						
F						



Genetic relationships

Compute relationships between A
and D, E, F

$$a_{AD} = 0.5(a_{AA} + a_{AB}) = 0.5(1 + 0) = 0.5$$

$$a_{AE} = 0.5(a_{AA} + a_{AC}) = 0.5(1 + 0) = 0.5$$

$$a_{AF} = 0.5(a_{AC} + a_{AB}) = 0.5(0.5 + 0.5) = 0.5$$

	A	B	C	A,B	A,C	E,D
A	1	0	0	0.5	0.5	
B	0	1	0			
C	0	0	1			
D	0.5					
E	0.5					
F						



Genetic relationships

Compute relationships between A
and D, E, F

$$a_{AD} = 0.5(a_{AA} + a_{AB}) = 0.5(1 + 0) = 0.5$$

$$a_{AE} = 0.5(a_{AA} + a_{AC}) = 0.5(1 + 0) = 0.5$$

$$a_{AF} = 0.5(a_{AC} + a_{AB}) = 0.5(0.5 + 0.5) = 0.5$$

	A	B	C	A,B	A,C	E,D
A	1	0	0	0.5	0.5	0.5
B	0	1	0			
C	0	0	1			
D	0.5					
E	0.5					
F	0.5					



Compute relationships between B and D, E, F

$$a_{BD} = 0.5(a_{BA} + a_{BB}) = 0.5(0 + 1) = 0.5$$

$$a_{BE} = 0.5(a_{BA} + a_{BC}) = 0.5(0 + 0) = 0$$

$$a_{BF} = 0.5(a_{BE} + a_{BD}) = 0.5(0 + 0.5) = 0.25$$

Compute relationships between C and D, E, F

$$a_{CD} = 0.5(a_{CA} + a_{CB}) = 0.5(0 + 0) = 0$$

$$a_{CE} = 0.5(a_{CA} + a_{CC}) = 0.5(0 + 1) = 0.5$$

$$a_{CF} = 0.5(a_{CE} + a_{CD}) = 0.5(0.5 + 0) = 0.25$$

	A	B	C	A,B	A,C	E,D
A	1	0	0	0.5	0.5	0.5
B	0	1	0	0.5	0	0.25
C	0	0	1	0	0.5	0.25
D	0.5	0.5	0			
E	0.5	0	0.5			
F	0.5	0.25	0.25			



Genetic relationships

Compute the diagonal for D

$$a_{DD} = 1 + F_D$$
$$F_D = 0.5a_{AB} = 0.5(0) = 0$$

$$a_{DD} = 1 + 0 = 1$$

Compute relationships between D
and E, F

$$a_{DE} = 0.5(a_{DA} + a_{DC}) = 0.5(0.5 + 0) = 0.25$$
$$a_{DF} = 0.5(a_{DE} + a_{DD}) = 0.5(0.25 + 1) = 0.5$$

	A	B	C	A,B	A,C	E,D
A	1	0	0	0.5	0.5	0.5
B	0	1	0	0.5	0	0.25
C	0	0	1	0	0.5	0.25
D	0.5	0.5	0	1	0.25	0.625
E	0.5	0	0.5	0.25		
F	0.5	0.25	0.25	0.625		



Genetic relationships

Compute the diagonal for E

$$a_{EE} = 1 + F_E = 1 + 0.5a_{AC} = 1 + 0 = 1$$

Compute relationship between E and F

$$a_{ED} = 0.5(a_{EE} + a_{ED}) = 0.5(1 + 0.25) = 0.625$$

Compute the diagonal for F

$$\begin{aligned}a_{FF} &= 1 + F_F = 1 + 0.5a_{ED} = 1 + 0.5(0.25) \\&= 1.125\end{aligned}$$

	A	B	C	A,B	A,C	E,D
A	1	0	0	0.5	0.5	0.5
B	0	1	0	0.5	0	0.25
C	0	0	1	0	0.5	0.25
D	0.5	0.5	0	1	0.25	0.625
E	0.5	0	0.5	0.25	1	0.625
F	0.5	0.25	0.25	0.625	0.625	1.125



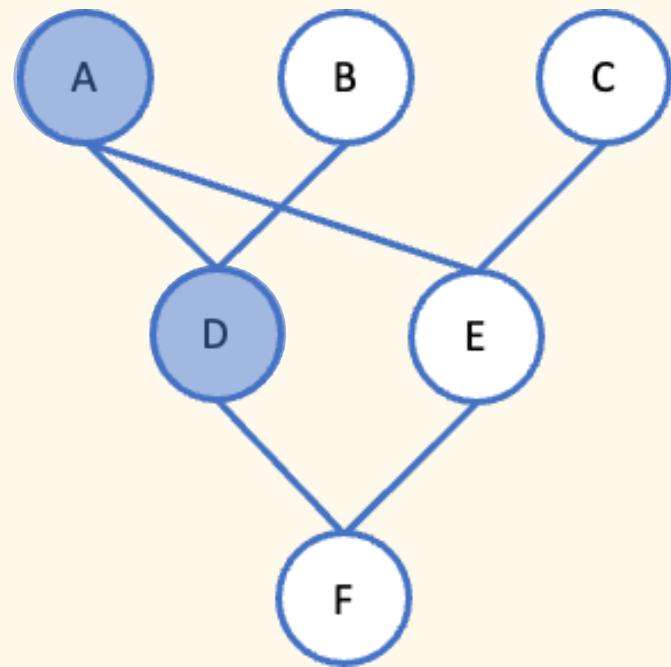
	A	B	C	D	A,B	A,C	E,D
A	1	0	0	0.5	0.5	0.5	0.5
B	0	1	0	0.5	0	0.25	
C	0	0	1	0	0.5	0.25	
D	0.5	0.5	0	1	0.25	0.625	
E	0.5	0	0.5	0.25	1	0.625	
F	0.5	0.25	0.25	0.625	0.625	0.625	1.125

$a_{AE} = 0.5$ parent-offspring

$a_{CF} = 0.25$
grandparent-offspring

$a_{DE} = 0.25$ half
siblings

Genetic relationships





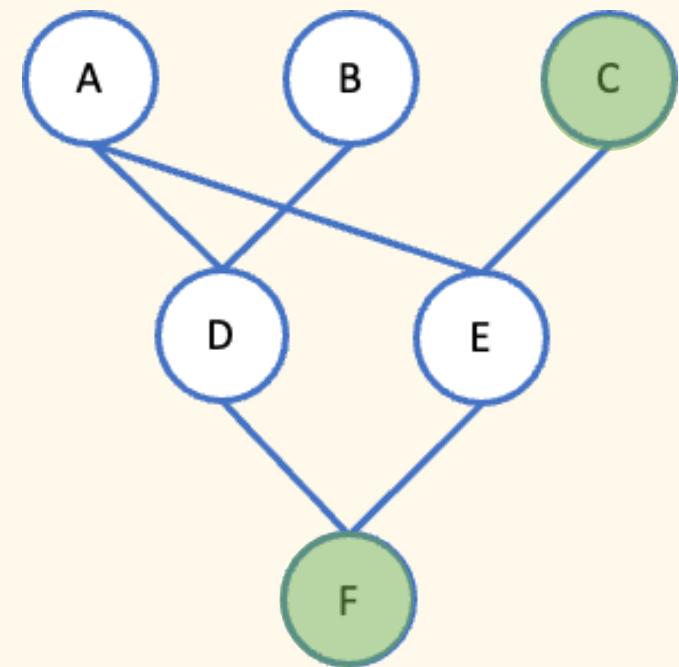
	A	B	C	D	A,B	A,C	E,D
A	1	0	0	0.5	0.5	0.5	0.5
B	0	1	0	0.5	0	0.25	
C	0	0	1	0	0.5	0.25	
D	0.5	0.5	0	1	0.25	0.625	
E	0.5	0	0.5	0.25	1	0.625	
F	0.5	0.25	0.25	0.625	0.625	0.625	1.125

$a_{AE} = 0.5$ parent-offspring

$a_{CF} = 0.25$
grandparent-offspring

$a_{DE} = 0.25$ half
siblings

Genetic relationships





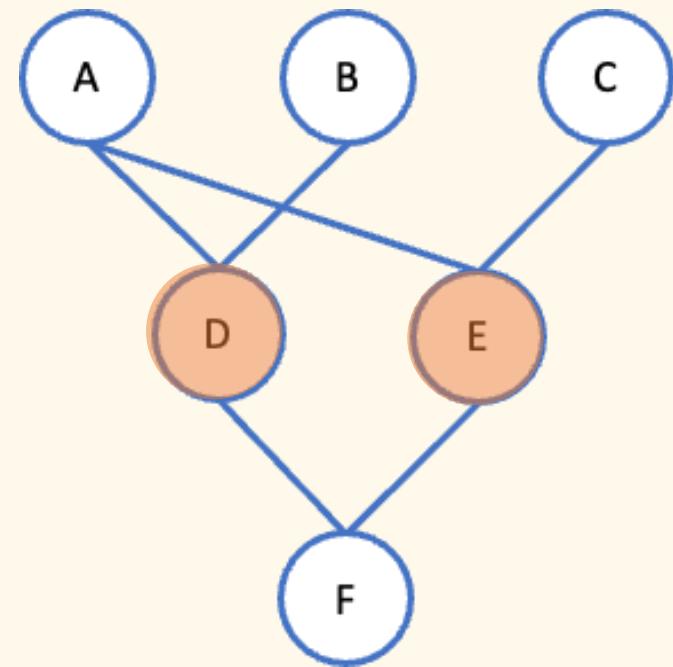
	A	B	C	D	A,B	A,C	E,D
A	1	0	0	0.5	0.5	0.5	0.5
B	0	1	0	0.5	0	0.25	
C	0	0	1	0	0.5	0.25	
D	0.5	0.5	0	1	0.25	0.625	
E	0.5	0	0.5	0.25	1	0.625	
F	0.5	0.25	0.25	0.625	0.625	0.625	1.125

$a_{AE} = 0.5$ parent-offspring

$a_{CF} = 0.25$
grandparent-offspring

$a_{DE} = 0.25$ half
siblings

Genetic relationships

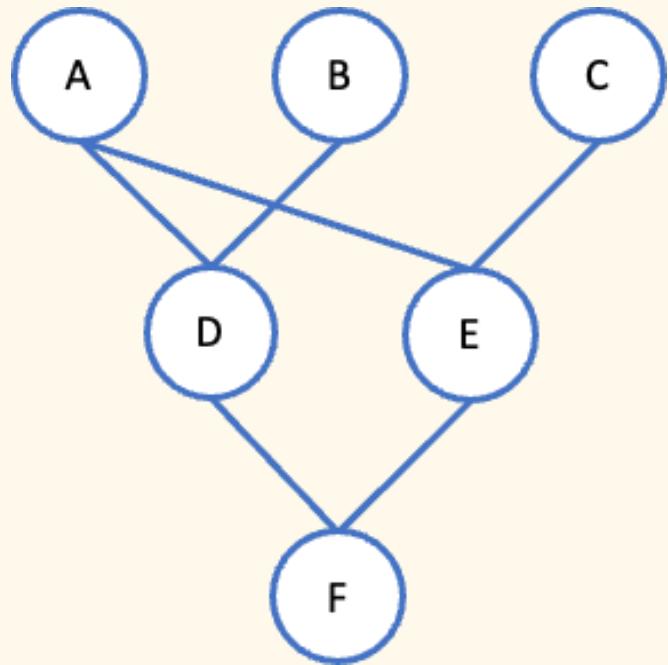




	A	B	C	A,B	A,C	E,D
A	1	0	0	0.5	0.5	0.5
B	0	1	0	0.5	0	0.25
C	0	0	1	0	0.5	0.25
D	0.5	0.5	0	1	0.25	0.625
E	0.5	0	0.5	0.25	1	0.625
F	0.5	0.25	0.25	0.625	0.625	1.125

A matrix

Genetic relationships



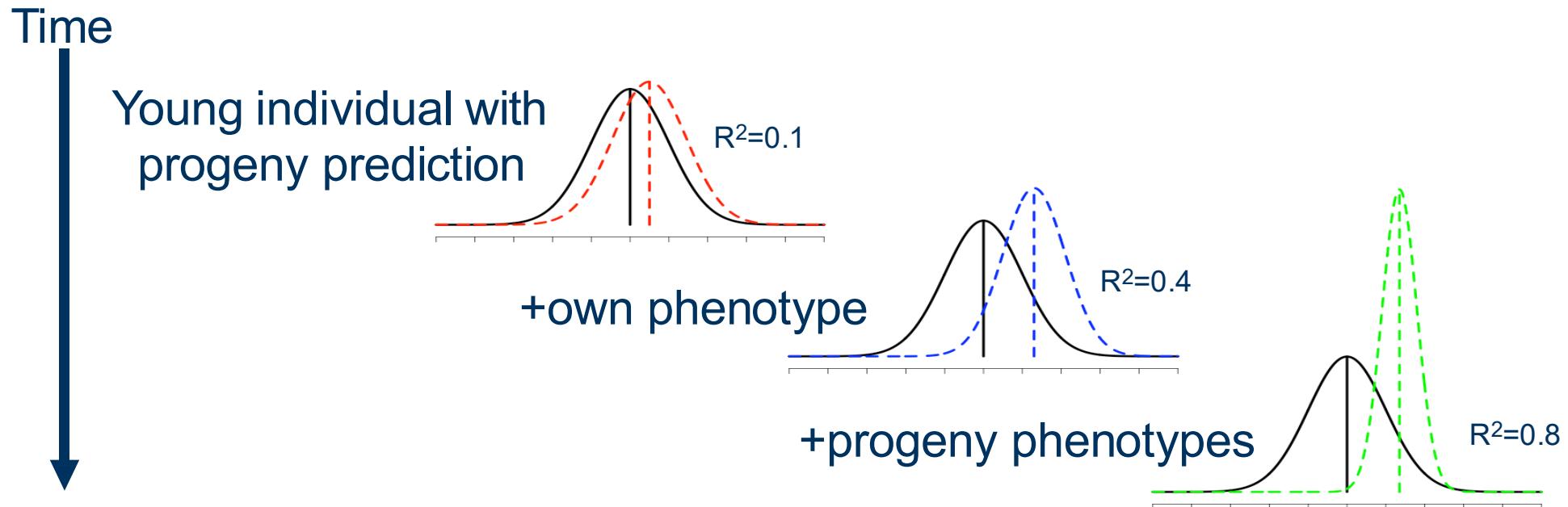


Centre for
Tropical Livestock
Genetics and Health

Demonstrate calculation in R with vectors & matrices



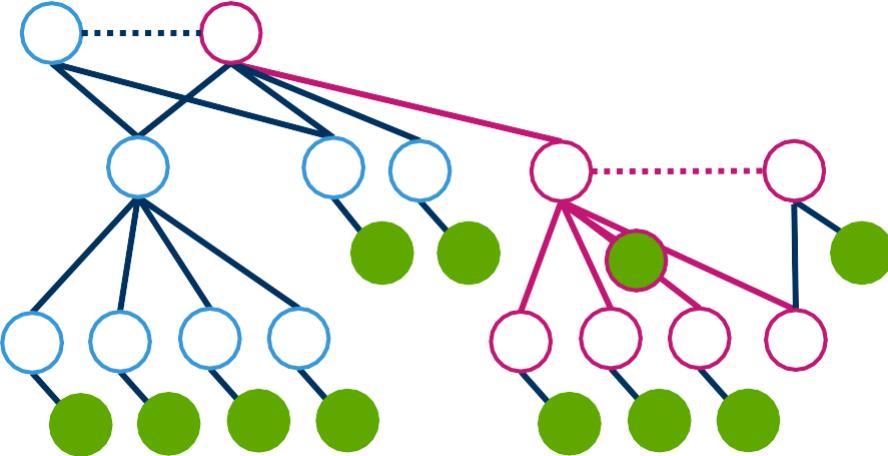
Accumulation of information over time



- Response = $f(\text{accuracy}, \text{intensity}, \text{diversity}, \text{time})$
- Accuracy vs. Time conflict



Information for an individual



- We take all information into account, though recursively
- Markov blanket for an individual breeding value
 - graphical “parents”, “progeny”, and “mates”
(=non-zero elements in C)



Pedigree-based linear-mixed model

- Phenotype extensions
 - multiple traits
 - different distributions
 - (binary, threshold, ordinal, survival, ...)
- Genetic extensions
 - sire/father, sire/father-dam/mother, animal/individual model
 - genetic groups, uncertain parentage
 - gametic & imprinting effects
 - sex & cytoplasmic inheritance
 - “social” genetic effects
 - individual QTL & genes



Computational aspects

- Standard programs (ASReml, blupf90, Mix99, WOMBAT, ...)
- Write your own! :(
- Sparse vs. dense matrices (\mathbf{C} and \mathbf{A}^{-1} are sparse!)
- Tasks
 - Estimate “fixed” and “random” effects (=location parameters)
 $\rightarrow p(\mathbf{b}, \mathbf{a} | \mathbf{y}, \sigma_a^2, \sigma_e^2)$
 - Estimate variance components (=dispersion/hyper-parameters)
 $\rightarrow p(\sigma_a^2, \sigma_e^2 | \mathbf{y})$ MAP/REML or full distribution
("Hill-climbing" (EM, NR, AI, ...) or "Hill-exploring" (MCMC, MC-EM) algorithms)



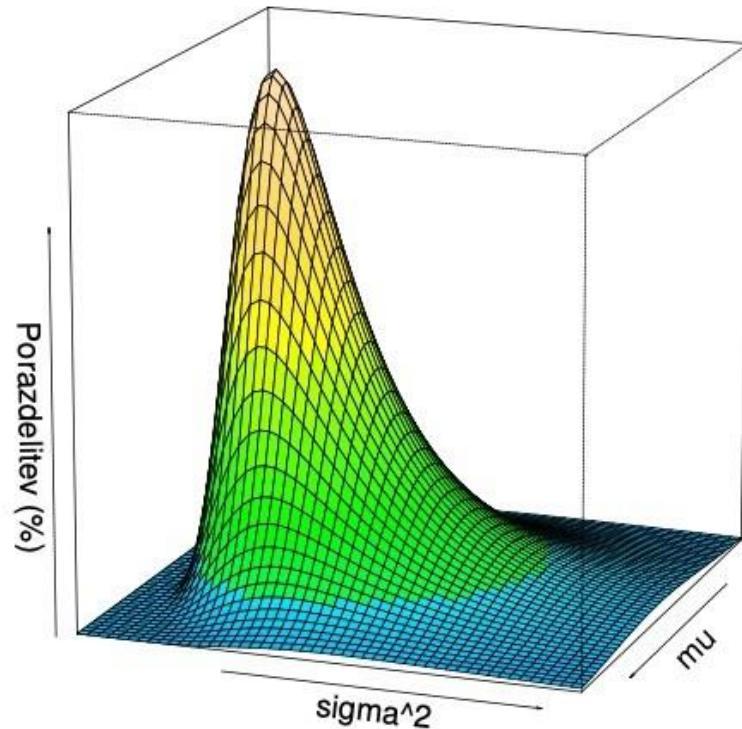
$$p(y|\mu, \sigma^2) = N(\mu, \sigma^2)$$

$$p(\mu) = N(\dots)$$

$$p(\sigma^2) = IG(\dots)$$

$$p(\mu, \sigma^2 | y) \propto p(y|\mu, \sigma^2) p(\mu) p(\sigma^2)$$

Simple mean & variance problem





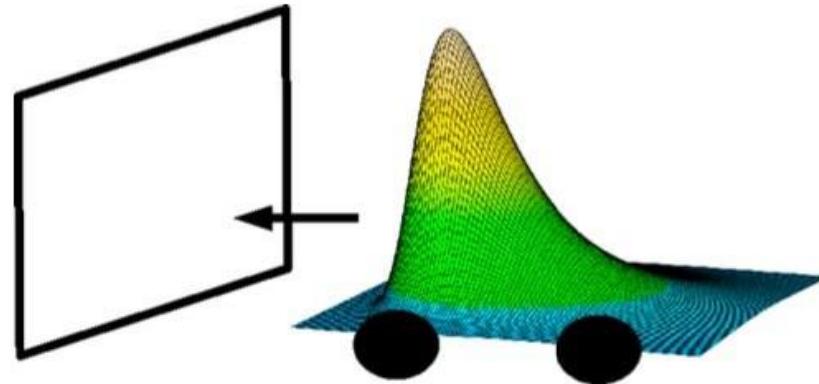
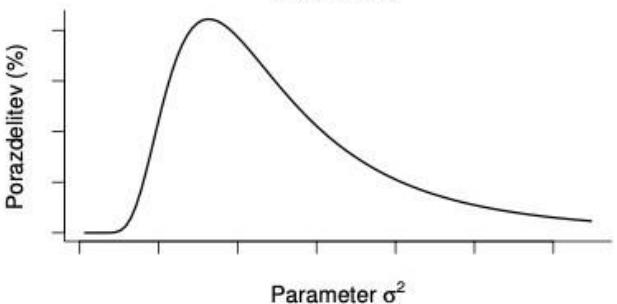
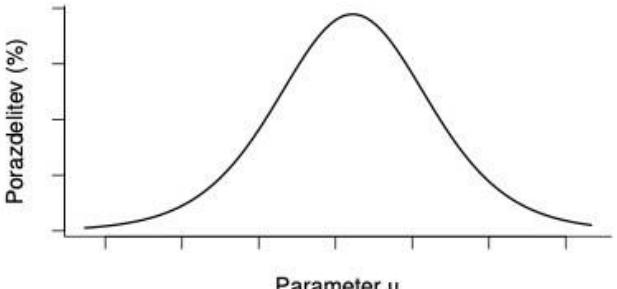
Simple mean & variance problem

$$p(\mu|y) = \int_0^\infty p(\mu, \sigma^2|y) d\sigma^2$$

$$\hat{\mu} = \sum_{i=1}^n y_i / n$$

$$p(\sigma^2|y) = \int_{-\infty}^{+\infty} p(\mu, \sigma^2|y) d\mu$$

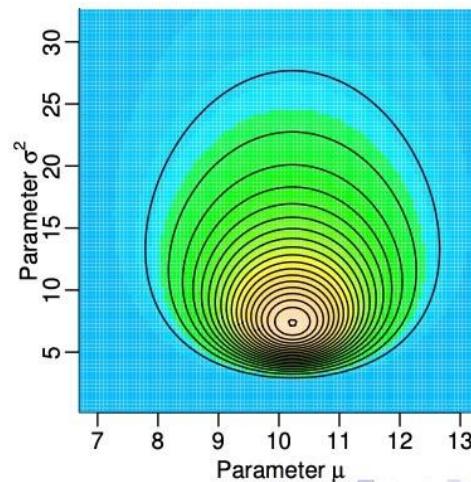
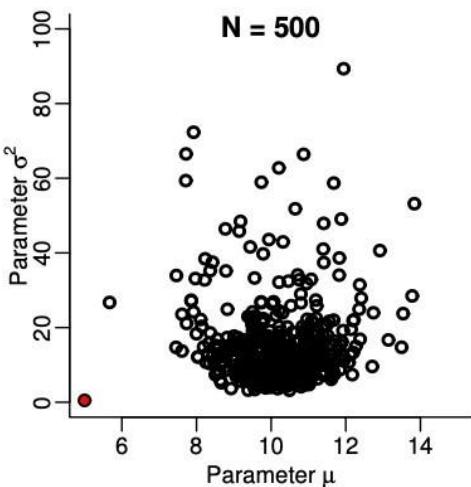
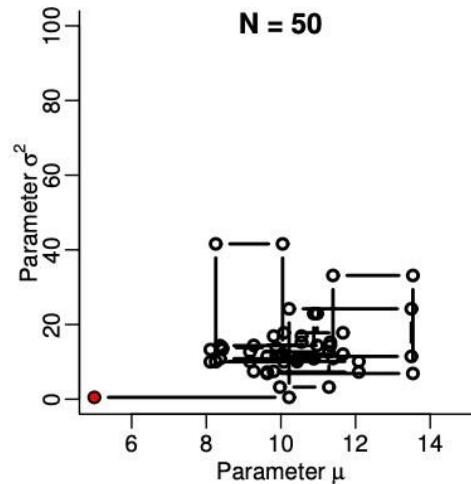
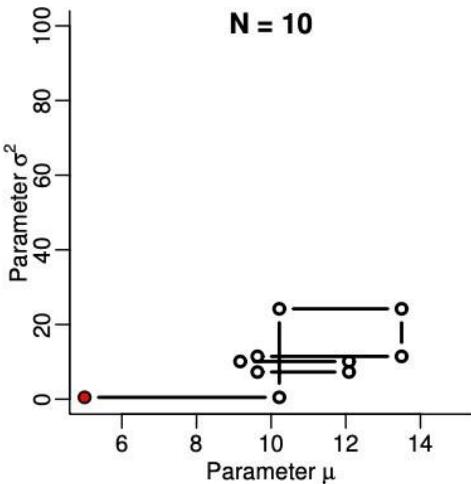
$$\sigma^2 = \sum_{i=1}^n (y_i - \hat{\mu})^2 / (n - 1)$$





Centre for
Tropical Livestock
Genetics and Health

Simple mean & variance problem

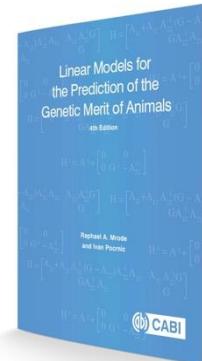




Computational aspects

- Sparse vs. dense matrices (C and A^{-1} are sparse!)
- Tasks
 - Estimate “fixed” and “random” effects (=location parameters)
 $\rightarrow p(\mathbf{b}, \mathbf{a} | \mathbf{y}, \sigma_a^2, \sigma_e^2)$
 - Estimate variance components (=dispersion/hyper-parameters)
 $\rightarrow p(\sigma_a^2, \sigma_e^2 | \mathbf{y})$ MAP/REML or full distribution
("Hill-climbing" (EM, NR, AI, ...) or "Hill-exploring" (MCMC, MC-EM) algorithms)

See chapters
3, 17-18, &
Appendix A!





Learning objectives

Understand how to combine phenotype information from all relatives connected via pedigree

Familiarise yourself with linear mixed models and equations

Practice inference of breeding values with the pedigree- based model

- simple cases using R matrix algebra
- using other packages