Pedigree BLUP - Sire Model and Animal Model

Peter von Rohr

2023-05-08

Independence Assumption

So far

- random residuals: $var(e) = I * \sigma_e^2$ and
- random sire effects: $var(s) = I * \sigma_s^2$

In real livestock datasets, this is not realistic, because sires are related

Example Dataset

Animal	Sire	Sex	WWG
4	1	M	4.5
5	3	F	2.9
6	1	F	3.9
7	4	M	3.5
8	3	M	5.0

Relationship

- For son *i* and sire *k* of *i*

$$cov(s_i, s_k) = 1/2 * \sigma_s^2$$

Sire Relationship Matrix

```
library(pedigreemm)
ped_sire <- pedigree(sire = c(rep(NA,2), 1),
                     dam = rep(NA,3),
                     label = as.character(c(1,3,4)))
mat_A <- getA(ped = ped_sire)
```

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

Sire Model

- pedigreemm

```
lmem_sire <- pedigreemm(
  formula = WWG ~ Sex + (1 | Sire),
  data = tbl_sire_model,
  pedigree = list(Sire = ped_sire)
)
```

```
## boundary (singular) fit: see help('isSingular')
```

```
summary(lmem_sire)
```

```
## Linear mixed model fit by REML ['lmerpedigreemm']
## Formula: WWG ~ Sex + (1 | Sire)
##    Data: tbl_sire_model
##
## REML criterion at convergence: 8.5
##
## Scaled residuals:
```

Mixed model equations

- `pedigreemm`: Cannot specify assumed variance components

$$\begin{bmatrix} X^T X & X^T Z \\ Z^T X & Z^T Z + \lambda A_s^{-1} \end{bmatrix} \begin{bmatrix} \hat{b} \\ \hat{s} \end{bmatrix} = \begin{bmatrix} X^T y \\ Z^T y \end{bmatrix}$$

Animal Model

- Extension of sire model
- Include dams in pedigree
- Predict breeding values for all animals in the pedigree

Dataset

Animal	Sire	Dam	Sex	WWG
4	1	NA	M	4.5
5	3	2	F	2.9
6	1	2	F	3.9
7	4	5	M	3.5
8	3	6	M	5.0

where `NA` stands for unknown

The Model

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Zu} + \mathbf{e}$$

- Random effects **u** and **e**

$$E(\mathbf{e}) = \mathbf{0}$$

$$var(\mathbf{e}) = \mathbf{I} * \sigma_e^2$$

$$E(\mathbf{u}) = \mathbf{0}$$

$$var(\mathbf{u}) = \mathbf{A} * \sigma_u^2$$

with numerator relationship matrix **A**

Numerator Relationship Matrix **A**

- Diagonal elements

$$(\mathbf{A})_{ii} = 1 + F_i$$

where F_i is the inbreeding coefficient

$$F_i = \frac{1}{2} * (\mathbf{A})_{sd}$$

- Off-diagonal elements

$$cov(u_i, u_j) = (\mathbf{A})_{ij} * \sigma_u^2$$

Example

```
library(pedigreemm)
ped_ani <- pedigree(sire = c(rep(NA, n_nr_founder),
                             1,3,1,4,3),
                    dam  = c(rep(NA, n_nr_founder),
                             NA,2,2,5,6),
                    label = as.character(1:n_nr_animal))
mat_A <- as.matrix(getA(ped = ped_ani))
```

$$A = \begin{bmatrix} 1 & 0 & 0 & 0.5 & 0 & 0.5 & 0.25 & 0.25 \\ 0 & 1 & 0 & 0 & 0.5 & 0.5 & 0.25 & 0.25 \\ 0 & 0 & 1 & 0 & 0.5 & 0 & 0.25 & 0.5 \\ 0.5 & 0 & 0 & 1 & 0 & 0.25 & 0.5 & 0.125 \\ 0 & 0.5 & 0.5 & 0 & 1 & 0.25 & 0.5 & 0.375 \\ 0.5 & 0.5 & 0 & 0.25 & 0.25 & 1 & 0.25 & 0.5 \\ 0.25 & 0.25 & 0.25 & 0.5 & 0.5 & 0.25 & 1 & 0.25 \\ 0.25 & 0.25 & 0.5 & 0.125 & 0.375 & 0.5 & 0.25 & 1 \end{bmatrix}$$

Solution

- `pedigreemm` cannot handle such small datasets with only one observation per animal
- Mixed model equations

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

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

Genomic BIUP

1. Marker effect models (MEM): Linear mixed effects models with marker effects as random effects
2. Breeding-value based models (BVM): Genomic breeding values as random effects

Marker Effect Models

- Model

$$y = 1_n\mu + Wq + e$$

- Solution

$$\left[\begin{array}{cc} 1_n^T 1_n & 1_n^T W \\ W^T 1_n & W^T W + \lambda_q * I \end{array}\right] \left[\begin{array}{c} \hat{\mu} \\ \hat{q} \end{array}\right] = \left[\begin{array}{c} 1_n^T y \\ W^T y \end{array}\right]$$

with $\lambda_q = \sigma_e^2/\sigma_q^2$.

Breeding Value Models

- Model

$$y = Xb + Zg + e$$

- Solution

$$\left[\begin{array}{cc} X^TX & X^TZ \\ Z^TX & Z^TZ + \lambda_g * G^{-1} \end{array} \right] \left[\begin{array}{c} \hat{b} \\ \hat{g} \end{array} \right] = \left[\begin{array}{c} X^Ty \\ Z^Ty \end{array} \right]$$

with $\lambda_g = \sigma_e^2/\sigma_g^2$.

Genomic Relationship Matrix

$$g = U \cdot q$$

with $U = W - P$ and P has columns $2p_j - 1$ with p_j being the frequency of the positive allele at locus j.

$$var(g) = G * \sigma_g^2$$

$$var(g) = UU^T * \sigma_q^2$$

$$\sigma_g^2 = 2 \sum_{j=1}^{m} p_j(1 - p_j)\sigma_q^2$$

Genomic Relationship Matrix II

$$var(g) = G * \sigma_g^2 = UU^T \sigma_q^2$$

$$G = \frac{UU^T}{2\sum_{j=1}^{m} p_j(1 - p_j)}$$