Additional Aspects of BLUP

Peter von Rohr

30 November 2018

Aspects

- Accurracy
  - Results from MME are estimates of fixed effects and predictions of breeding values
  - Need statement about quality of estimates and predictions
- Confidence Intervals
- Decomposition of Predicted Breeding values

Accurracy

- One property of BLUP was that variance of prediction error is minimal
- How can we measure the variance of the prediction error
- Fixed effects

$$var(\beta - \hat{\beta}) = var(\hat{\beta})$$

- Random effects

$$var(a - \hat{a}) = var(a) - 2 * cov(a, \hat{a}) + var(\hat{a}) = var(a) - var(\hat{a}) = PEV(\hat{a})$$

because with BLUP: $cov(a, \hat{a}) = var(\hat{a})$

PEV

- PEV depends on inverse of coefficient matrix of MME

$$\left[\begin{array}{cc} X^TR^{-1}X & X^TR^{-1}Z \\ Z^TR^{-1}X & Z^TR^{-1}Z + G^{-1} \end{array}\right]^{-1} = \left[\begin{array}{cc} C^{11} & C^{12} \\ C^{21} & C^{22} \end{array}\right]$$

- For predicted breeding values $\hat{a}$

$$PEV(\hat{a}) = var(a) - var(\hat{a}) = C^{22}$$

Single Animal i

$$PEV(\hat{a}_i) = (C)^{22}_{ii}$$

where $(C)^{22}_{ii}$ is the i-th diagonal of C^{22}

- Accuracy measured by correlation

$$r_{a_i,\hat{a}_i} = \frac{cov(a_i, \hat{a}_i)}{\sqrt{var(a_i) * var(\hat{a}_i)}} = \sqrt{\frac{var(\hat{a}_i)}{var(a_i)}}$$

- Combining

$$PEV(\hat{a}_i) = (C)^{22}_{ii} = var(a_i) - var(\hat{a}_i) = var(a_i) - r^2_{a_i,\hat{a}_i} var(a_i)$$

Accuracy B_i

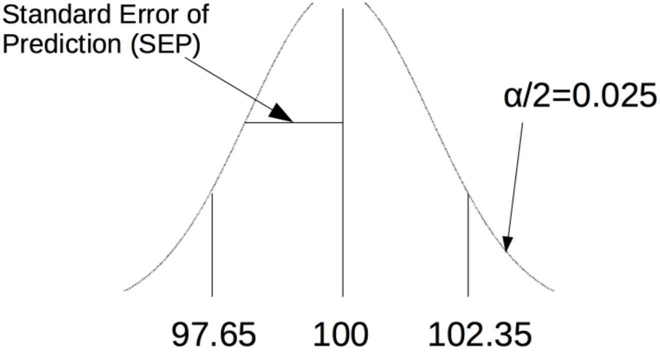
$$B_i = r^2_{a_i,\hat{a}_i} = \frac{var(a_i) - (C)^{22}_{ii}}{var(a_i)} = 1 - \frac{PEV(\hat{a}_i)}{var(a_i)} = 1 - \frac{(C)^{22}_{ii}}{var(a_i)}$$

- B_i is large for small $PEV(\hat{a}_i)$
- In the limit $B_i \rightarrow 1$ for $PEV(\hat{a}_i) \rightarrow 0$
- For $PEV(\hat{a}_i) \rightarrow 0$ we must have $var(\hat{a}_i) \rightarrow var(a_i)$
- Therefore, the closer $var(\hat{a}_i)$ is to $var(a_i)$, the more accurate the predicted breeding value

Confidence Intervals of $\hat{a}_i$

- Predicted breeding value ($\hat{a}_i$) is a function of the data (y)
- Hence $\hat{a}_i$ is a random variable with a distribution

Distribution



$$SEP(\hat{a}_i) = \sqrt{PEV(\hat{a}_i)} = \sqrt{(1 - r^2_{a_i,\hat{a}_i}) * var(a_i)}$$

Widths Of Confidence Intervals

Table 1: Widths of Confidence Intervals for Given Accuracies

Accurracy	Interval Width
0.40	36.44
0.50	33.26
0.60	29.75
0.70	25.76
0.80	21.04
0.90	14.88
0.95	10.52
0.99	4.70

with $\hat{a}_i = 100$, $var(a_i) = 144$ and $\alpha = 0.05$

Decomposition of Predicted Breeding Value

- Write MME as

$$M \cdot s = r$$

with

$$s = \left[\begin{array}{c} \hat{\beta} \\ \hat{a} \end{array} \right]$$

- $\hat{\beta}$ has length p
- $\hat{a}$ has length q

Simplified Model

$$y_i = \mu + a_i + e_i$$

where

y_i	Observation for animal i
a_i	breeding value of animal i with a variance of $(1 + F_i)\sigma_a^2$
e_i	random residual effect with variance σ_e^2
μ	single fixed effect

Data

- all animals have an observation
- animal i has
  - parents s and d
  - n progeny k_j (with $j = 1, \ldots, n$)
  - n mates l_j (with $j = 1, \ldots, n$).
- progeny k_j has parents i and l_j.

Example

Animal	Sire	Dam	WWG
1	NA	NA	4.5
2	NA	NA	2.9
3	NA	NA	3.9
4	1	2	3.5
5	4	3	5.0

Variance components $\sigma_e^2 = 40$ and $\sigma_a^2 = 20$.

Model Components

$$X = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \; Z = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$X^T X = \begin{bmatrix} 5 \end{bmatrix}, \; X^T Z = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

$$Z^T Z = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Right-hand Side

$$X^T y = \left[\sum_{j=1}^{n} y_i \right]$$

$$Z^T y = \begin{bmatrix} y1 \\ y2 \\ y3 \\ y4 \\ y5 \end{bmatrix}$$

A^{-1}

$$A^{-1} = \begin{bmatrix} 1.50 & 0.50 & 0.00 & -1.00 & 0.00 \\ 0.50 & 1.50 & 0.00 & -1.00 & 0.00 \\ 0.00 & 0.00 & 1.50 & 0.50 & -1.00 \\ -1.00 & -1.00 & 0.50 & 2.50 & -1.00 \\ 0.00 & 0.00 & -1.00 & -1.00 & 2.00 \end{bmatrix}$$

MME

$$\begin{bmatrix} X^T X & X^T Z \\ Z^T X & Z^T Z + A^{-1} * \lambda \end{bmatrix} \begin{bmatrix} \hat{\mu} \\ \hat{a} \end{bmatrix} = \begin{bmatrix} X^T y \\ Z^T y \end{bmatrix}$$

Insert Data

$$\begin{bmatrix} 5.00 & 1.00 & 1.00 & 1.00 & 1.00 & 1.00 \\ 1.00 & 4.00 & 1.00 & 0.00 & -2.00 & 0.00 \\ 1.00 & 1.00 & 4.00 & 0.00 & -2.00 & 0.00 \\ 1.00 & 0.00 & 0.00 & 4.00 & 1.00 & -2.00 \\ 1.00 & -2.00 & -2.00 & 1.00 & 6.00 & -2.00 \\ 1.00 & 0.00 & 0.00 & -2.00 & -2.00 & 5.00 \end{bmatrix} \begin{bmatrix} \mu \\ \hat{a}1 \\ \hat{a}2 \\ \hat{a}3 \\ \hat{a}4 \\ \hat{a}5 \end{bmatrix} = \begin{bmatrix} 19.8 \\ 4.5 \\ 2.9 \\ 3.9 \\ 3.5 \\ 5.0 \end{bmatrix}$$

Animal 4

- parents 1 and 2
- progeny 5
- mate 3
- inspection of second but last equation in MME where y_4 and $\hat{a}_4$ occur
- Remember from construction of A^{-1}, the variable d^{ii} can assume the following values

$$d^{ii} = \begin{cases} 2 & \text{both parents known} \\ \frac{4}{3} & \text{one parent known} \\ 1 & \text{both parents unknown} \end{cases}$$

Extract Equation

$$y_4 = 3.5 = 1 * \hat{\mu} - 2 * \hat{a}_1 - 2 * \hat{a}_2 + 1 * \hat{a}_3 + 6 * \hat{a}_4 - 2 * \hat{a}_5$$

- Solving for $\hat{a}_4$

$$\hat{a}_4 = \frac{1}{6}\left[y_4 - \hat{\mu} + 2 * (\hat{a}_1 + \hat{a}_2) - \hat{a}_3 + 2\hat{a}_5\right]$$

- $\hat{a}_4$ depends on
  - own performance record y_4
  - estimate of fixed effect $\hat{\mu}$ - environment
  - predicted breeding value of parents 1 and 2, mate 3 and progeny 5