LfL

Bavarian State Research Center
for Agriculture

66th EAAP Annual Meeting

Equivalence of genomic breeding values and reliabilities estimated with SNP-BLUP and GBLUP

Laura Plieschke[1], Christian Edel[1], Eduardo CG Pimentel[1], Reiner Emmerling[1], Jörn Bennewitz[2] and Kay-Uwe Götz[1]

[1]Bavarian State Research Center for Agriculture, Institute of Animal Breeding
[2]Institute of Animal Husbandry and Breeding, University Hohenheim

2 September 2015

Introduction

- GBLUP (VanRaden, 2008) is one of the most common procedures to estimate genomic breeding values
- An alternative to estimate genomic breeding values is the so called SNP-BLUP (Meuwissen et al., 2001)
- The equivalence of both model has been shown (e.g. Goddard, 2008; VanRaden, 2008)
- But: there have been some irritations in the literature (Strandén and Christensen, 2011)
- Demonstrate the identity of DGVs and reliabilities by a practical example

- Why did we do that?
  - Check the practicability and feasibility in real life applications
  - Preliminary investigation: methods
  - Current studies: enhance the reference population by a large number of genotyped cows

Data

- Phenotypes and genotypes of 11 852 Fleckvieh sires
- Genotyped with the Illumina BovineSNP50 BeadChip
- DGVs and their reliabilities for milk yield
- Total of 41 266 SNPs retained after filtering
- Genotype coding: 2q, (q-p), -2p

Models

G-BLUP

$$y = Xb + Za + e$$

$$\begin{pmatrix} X'R^{-1}X & X'R^{-1}Z \\ Z'R^{-1}X & Z'R^{-1}Z+G^{-1}/\sigma_a^2 \end{pmatrix} \begin{pmatrix} \hat{b} \\ \hat{a} \end{pmatrix} = \begin{pmatrix} X'R^{-1}y \\ Z'R^{-1}y \end{pmatrix}$$

SNP-BLUP

$$y = Xb + Mg + e$$

$$\begin{pmatrix} X'R^{-1}X & X'R^{-1}M \\ M'R^{-1}X & M'R^{-1}M+I/\sigma_g^2 \end{pmatrix} \begin{pmatrix} \hat{b} \\ \hat{g} \end{pmatrix} = \begin{pmatrix} X'R^{-1}y \\ M'R^{-1}y \end{pmatrix}$$

y	= vector of observations (DYD)
b	= vector of fixed effects (mean)
a	= vector of random animal effects
X, Z	= design matrices
R	= residual co-variance matrix
G	= genomic relationship matrix
M	= coefficient matrix of marker genotypes
g	= vector of random marker effects
I	= identity matrix
e	= vector of residual effects

G-BLUP

$$y = Xb + Za + e$$

$$\begin{pmatrix} X'R^{-1}X & X'R^{-1}Z \\ Z'R^{-1}X & Z'R^{-1}Z+G^{-1}/\sigma_a^2 \end{pmatrix} \begin{pmatrix} \hat{b} \\ \hat{a} \end{pmatrix} = \begin{pmatrix} X'R^{-1}y \\ Z'R^{-1}y \end{pmatrix}$$

$$G = \frac{MM'}{\sum_{j=1}^{m}(2p_jq_j)}$$

SNP-BLUP

$$y = Xb + Mg + e$$

$$\begin{pmatrix} X'R^{-1}X & X'R^{-1}M \\ M'R^{-1}X & M'R^{-1}M+I/\sigma_g^2 \end{pmatrix} \begin{pmatrix} \hat{b} \\ \hat{g} \end{pmatrix} = \begin{pmatrix} X'R^{-1}y \\ M'R^{-1}y \end{pmatrix}$$

$$\sigma_g^2 = \frac{\sigma_a^2}{\sum_{j=1}^{m}(2p_jq_j)}$$

G-BLUP

$$DGV_i = \hat{\mathbf{b}} + \hat{\mathbf{a}}_i$$

$$pev(DGV) = \mathbf{Z}^{*}\mathbf{C}_{\mathbf{g}}^{-\mathbf{1}}\mathbf{Z}^{*\prime}$$

SNP-BLUP

$$DGV_i = \hat{\mathbf{b}} + \mathbf{M}_{[i,:]}\hat{\mathbf{g}}$$

$$pev(DGV) = \mathbf{M}^{*}\mathbf{C}_{\mathbf{s}}^{-\mathbf{1}}\mathbf{M}^{*\prime}$$

$$\Rightarrow \quad r^2_i = 1 - \frac{diag(pev(DGV))_i}{diag(\mathbf{G})_i \sigma^2_a}$$

C^{-1} = Inverse of the left-hand-side of the MME

G-BLUP

$$DGV_i = \hat{\mathbf{b}} + \hat{\mathbf{a}}_i$$

$$pev(DGV) = \mathbf{Z}^{*}\mathbf{C}_{\mathbf{g}}^{-\mathbf{1}}\mathbf{Z}^{*\prime}$$

SNP-BLUP

$$DGV = \hat{\mathbf{b}} + \mathbf{M}_{[i,:]}\hat{\mathbf{g}}$$

$$pev(DGV) = \mathbf{M}^{*}\mathbf{C}_{\mathbf{S}}^{-\mathbf{1}}\mathbf{M}^{*\prime}$$

$$\Rightarrow \quad r^2_i = 1 - \frac{diag(pev(DGV))_i}{diag(\mathbf{G})_i \sigma^2_a}$$

C^{-1} = Inverse of the left-hand-side of the MME

G-BLUP

$$DGV_i = \hat{\mathbf{b}} + \hat{\mathbf{a}}_i$$

$$pev(DGV) = \mathbf{Z}^{*}\mathbf{C}_{\mathbf{G}}^{-1}\mathbf{Z}^{*\prime}$$

SNP-BLUP

$$DGV = \hat{\mathbf{b}} + \mathbf{M}_{[i,:]}\hat{\mathbf{g}}$$

$$pev(DGV) = \mathbf{M}^{*}\mathbf{C}_{\mathbf{S}}^{-1}\mathbf{M}^{*\prime}$$

$$r^2_i = 1 - \frac{diag(pev(DGV))_i}{diag(\mathbf{G})_i \sigma^2_a}$$

C^{-1} = Inverse of the left-hand-side of the MME

Results

- DGVs obtained with both models were the same

  ⟹ models are **equivalent in DGVs**

- Reliabilities of DGVs from both models were also the same, when genomic inbreeding was taken into account

  ⟹ models are **equivalent in reliabilities of DGVs**

- For both methods we can imagine advantages and disadvantages in different scenarios
- This is primarily due to the structure of data:

  - If number of markers **>>** number of animals: GBLUP would be preferable
  - If number of markers **<<** number of animals: SNP-BLUP would be preferable

Conclusion

- SNP-BLUP and GBLUP lead to equivalent results for DGVs and their reliabilities

- Prerequisites for identical and meaningful reliabilities in both cases:
  - Error variance of the intercept
  - Genomic inbreeding coefficient
  - Genotype coding

- The model of choice should mainly depend on the structure of the dataset

Thank you for your attention

We gratefully acknowledge:

- Arbeitsgemeinschaft Süddeutscher Rinderzucht- und Besamungsorganisationen
- Contributors of the genotype pool Germany-Austria