

An Equivalent (animal) Model for Genomic Prediction

Usual Single Trait Pedigree Model

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

$$\begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + \lambda A^{-1} \end{bmatrix} \begin{bmatrix} b^0 \\ \widehat{\mathbf{u}} \end{bmatrix} = \begin{bmatrix} X'y \\ Z'y \end{bmatrix}$$

Often known as PBLUP

Single Trait Marker Effects Model

$$y = Xb + ZM\alpha + e$$

$$\begin{bmatrix} X'X & X'ZM \\ M'Z'X & M'Z'ZM + \frac{\sigma_e^2}{\sigma_\alpha^2}I \end{bmatrix} \begin{bmatrix} b^0 \\ \hat{\alpha} \end{bmatrix} = \begin{bmatrix} X'y \\ M'Z'y \end{bmatrix}$$

$$\begin{bmatrix} X'X & X'M \\ M'X & M'M + \frac{\sigma_e^2}{\sigma_\alpha^2}I \end{bmatrix} \begin{bmatrix} b^0 \\ \hat{\alpha} \end{bmatrix} = \begin{bmatrix} X'y \\ M'y \end{bmatrix}$$

$$\text{Then } \hat{u} = M\hat{\alpha}$$

This is MHG "BLUP" or is sometimes known as (Ridge-regression) RR-BLUP with known $\frac{\sigma_e^2}{\sigma_\alpha^2}$

More loci than animals

But for selection we are more interested in animal (not allelic) merit

$$y = Xb + ZM\alpha + e$$

$$y = Xb + M\alpha + e$$

$$y = Xb + IM\alpha + e$$

$$y = Xb + "Z" "u" + e$$

Order of MME is number of fixed effects plus number of animals
Consider the implications for 100-1,000 animals with 50,000 loci

Mixed Model Equations

$$y = Xb + IM\alpha + e$$

$$y = Xb + "Z" "u" + e$$

$$\begin{bmatrix} X'X & X' \\ X & I + \sigma_e^2 [\text{var}(M\alpha)]^{-1} \end{bmatrix} \begin{bmatrix} b^0 \\ \widehat{M\alpha} \end{bmatrix} = \begin{bmatrix} X'y \\ y \end{bmatrix}$$

What is $\text{var}(M\alpha)$?

$$\begin{aligned} \text{var}(M\alpha) &= M \text{var}(\alpha) M' \\ &= MIM' \sigma_\alpha^2 = MM' \sigma_\alpha^2 \end{aligned}$$

Homogenous locus variance

What is $\text{var}(M\alpha)$?

$$\begin{aligned}\text{var}(M\alpha) &= M\text{var}(\alpha)M' \\ &= MIM'\sigma_{\alpha}^2 = MM'\sigma_{\alpha}^2\end{aligned}$$

Homogeneous locus variance

$$\begin{aligned}\text{var}[u] &= \text{var}\left[\sum_{loci} m_i \alpha_i\right] = \sum_{loci} \text{var}(m_i \alpha_i) \\ &= \sum_{loci} m_i \text{var}(\alpha_i) m_i' = \sum_{loci} m_i m_i' \sigma_{\alpha_i}^2\end{aligned}$$

Heterogeneous locus variance

Genomic Relationship Matrix

$$G = MM'$$

$$G = \sum_{loci} m_i m_i' \sigma_{\alpha_i}^2$$

Trait specific

Genomic Relationship Matrix

$M = k$ columns of $(0, 1, 2)$ marker covariates

$$\textcolor{red}{G} = [MM' + (2 - M)(2 - M)'] / k$$

$$\text{var}[\textcolor{red}{u}] = \textcolor{red}{G}\sigma_a^2$$

Nejati-Javaremi, Smith & Gibson JAS (1997)

GBLUP

$$\begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + \lambda \textcolor{blue}{G}^{-1} \end{bmatrix} \begin{bmatrix} b^0 \\ \widehat{\textcolor{blue}{u}} \end{bmatrix} = \begin{bmatrix} X'y \\ Z'y \end{bmatrix}$$

$$G = MM'$$

$$\widehat{\textcolor{blue}{u}} = M\widehat{\alpha}$$

GBLUP

- If the variance parameters are assumed known and the inverse of the genomic relationship matrix is multiplied by (known) λ , the system is known as GBLUP, as opposed to conventional pedigree or PBLUP
 - It is effectively weighting all the loci equally
 - It is similar to BayesC0 except that in that method we estimate the variance components after including a prior distribution for them

Lack of Equivalence

- The GBLUP and Marker Effects Models (MEM) such as BayesC0 with high df for the prior variances will give the same EBV for the genotyped animals
 - This is true regardless of
 - whether the models fit the A allele at every locus, the B allele at every locus, or both alleles at every locus
 - how the alleles are centered (coded 0,1,2 or -1,0,1 etc)
 - However, the PEV (and reliability) for GBLUP are not invariant to these alternative models