

Quantitative Genetics, Prediction and Selection Theory

SELECTION INDICES

BLUP and REML

BAYESIAN INFERENCE

GENOMIC SELECTION

Genomic selection

- 6.1. The model with molecular markers
- 6.2. Genomic selection with the infinitesimal model
 - RR-BLUP
 - G-BLUP
 - Single-step G-BLUP
- 6.3. Gene detection
 - GWAS
 - The Bayesian alphabet

The model with markers

THE SIMPLIFIED MODEL

- The markers are SNPs: A/C, A/G, A/T, C/G, C/T, G/T
- Three possible genotypes per SNP; e.g., AA, AG, GG
- The chips contain thousands of SNPs (e.g., 50,000 SNPs)
- The objective is to associate productivity with SNPs via regression
 - A **reference population** with productive data is required
 - Animals must be genotyped to predict their production
 - The equations must be recalculated every two or three generations
 - In the simplified model, fixed effects are ignored
 - Predictions are made solely with the SNPs, without considering an additive effect
- It is assumed that there are SNPs statistically associated (linked) with genes that contribute to the additive value

The SNPs ARE NOT GENES; they are associated markers

The greater the number of SNPs, the smaller the effect of each one

The model with markers

THE SIMPLIFIED MODEL

Values of the genotypes corresponding to an A/G SNP

AA	AG	GG
2	1	0

Chip [A/G, C/G, A/T, ..., C/T]

Arbitrary reference allele [A, C, A, ..., C]

[AA, CC, AA, ..., CC] → 2

[AG, CG, AT, ..., CT] → 1

[GG, GG, TT, ..., TT] → 0

The model with markers

THE SIMPLIFIED MODEL

$$y = \mu + \alpha_1 m_1 + \alpha_2 m_2 + \alpha_3 m_3 + \cdots + \alpha_{50,000} m_{50,000} + e = \mu + a + e$$

$$\begin{bmatrix} 10,602 \\ 11,440 \\ 9,103 \\ \dots \\ \dots \\ 9,657 \end{bmatrix} = \begin{bmatrix} \mu \\ \mu \\ \mu \\ \dots \\ \dots \\ \mu \end{bmatrix} + \alpha_1 \begin{bmatrix} 2 \\ 0 \\ 2 \\ \dots \\ \dots \\ 1 \end{bmatrix} + \alpha_2 \begin{bmatrix} 1 \\ 1 \\ 0 \\ \dots \\ \dots \\ 2 \end{bmatrix} + \alpha_3 \begin{bmatrix} 0 \\ 2 \\ 1 \\ \dots \\ \dots \\ 0 \end{bmatrix} + \cdots + \alpha_{50,000} \begin{bmatrix} 1 \\ 0 \\ 0 \\ \dots \\ \dots \\ 2 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ \dots \\ \dots \\ e_{4,000} \end{bmatrix}$$

$$\mathbf{y} = \mu + \alpha_1 \mathbf{m}_1 + \alpha_2 \mathbf{m}_2 + \alpha_3 \mathbf{m}_3 + \cdots + \alpha_{50,000} \mathbf{m}_{50,000} + \mathbf{e}$$

$$\mathbf{y} = \mathbf{M} \cdot \boldsymbol{\alpha} + \mathbf{e}$$

$$\hat{a} = \mu + \hat{\alpha}_1 m_1 + \hat{\alpha}_2 m_2 + \hat{\alpha}_3 m_3 + \cdots + \hat{\alpha}_{50,000} m_{50,000}$$

The model with markers

THE SIMPLIFIED MODEL

$$y = \mu + \alpha_1 m_1 + \alpha_2 m_2 + \alpha_3 m_3 + \cdots + \alpha_{50,000} m_{50,000} + e$$

$$\begin{bmatrix} 10,602 \\ 11,440 \\ 9,103 \\ \dots \\ 9,657 \end{bmatrix} = \begin{bmatrix} \mu \\ \mu \\ \mu \\ \dots \\ \mu \end{bmatrix} + \alpha_1 \begin{bmatrix} 2 \\ 0 \\ 2 \\ \dots \\ 1 \end{bmatrix} + \alpha_2 \begin{bmatrix} 1 \\ 1 \\ 0 \\ \dots \\ 2 \end{bmatrix} + \alpha_3 \begin{bmatrix} 0 \\ 2 \\ 1 \\ \dots \\ 0 \end{bmatrix} + \cdots + \alpha_{50,000} \begin{bmatrix} 1 \\ 0 \\ 0 \\ \dots \\ 2 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ \dots \\ e_{4,000} \end{bmatrix}$$

$$\mathbf{y} = \mu + \alpha_1 \mathbf{m}_1 + \alpha_2 \mathbf{m}_2 + \alpha_3 \mathbf{m}_3 + \cdots + \alpha_{50,000} \mathbf{m}_{50,000} + \mathbf{e}$$

10,602	AA	CG	TT	CT
--------	----	----	----	----

11,440	GG	CG	AA	TT
--------	----	----	----	----

The model with markers

THE SIMPLIFIED MODEL

$$m_y = 10,000$$

$$\begin{bmatrix} 10,602 \\ 11,440 \\ 9,103 \\ \dots \\ \dots \\ 9,657 \end{bmatrix} = \begin{bmatrix} \mu \\ \mu \\ \mu \\ \dots \\ \dots \\ \mu \end{bmatrix} + \alpha_1 \begin{bmatrix} 2 \\ 0 \\ 2 \\ \dots \\ \dots \\ 1 \end{bmatrix} + \alpha_2 \begin{bmatrix} 1 \\ 1 \\ 0 \\ \dots \\ \dots \\ 2 \end{bmatrix} + \alpha_3 \begin{bmatrix} 0 \\ 2 \\ 1 \\ \dots \\ \dots \\ 0 \end{bmatrix} + \dots + \alpha_{50,000} \begin{bmatrix} 1 \\ 0 \\ 0 \\ \dots \\ \dots \\ 2 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ \dots \\ \dots \\ e_{4,000} \end{bmatrix}$$

$$\mathbf{y} = \mu + \alpha_1 \mathbf{m}_1 + \alpha_2 \mathbf{m}_2 + \alpha_3 \mathbf{m}_3 + \dots + \alpha_{50,000} \mathbf{m}_{50,000} + \mathbf{e}$$

$$\bar{m}_1 = \frac{2+0+2+\dots+1}{n} = \frac{\sum n_k}{n} \quad \begin{array}{l} n_k: \text{n}^\circ \text{ alleles A in row } k \\ n: \text{n}^\circ \text{ of rows (n}^\circ \text{ of data)} \end{array}$$

$$p_1 = \frac{\text{n}^\circ \text{ A}}{\text{total}} = \frac{2+0+2+\dots+1}{2n} = \frac{\sum n_k}{2n}$$

$$\bar{m}_1 = 2p_1$$

Genomic selection

The SNP-BLUP or RR-BLUP

$$m_y = 10,000$$

$$\begin{bmatrix} 602 \\ -440 \\ 103 \\ \dots \\ \dots \\ -657 \end{bmatrix} = \alpha_1 \begin{bmatrix} 2-2p_1 \\ 0-2p_1 \\ 2-2p_1 \\ \dots \\ \dots \\ 1-2p_1 \end{bmatrix} + \alpha_2 \begin{bmatrix} 1-2p_2 \\ 1-2p_2 \\ 0-2p_2 \\ \dots \\ \dots \\ 2-2p_2 \end{bmatrix} + \alpha_3 \begin{bmatrix} 1-2p_3 \\ 2-2p_3 \\ 1-2p_3 \\ \dots \\ \dots \\ 0-2p_3 \end{bmatrix} + \dots + \alpha_{50,000} \begin{bmatrix} 1-2p_{50,000} \\ 0-2p_{50,000} \\ 0-2p_{50,000} \\ \dots \\ \dots \\ 2-2p_{50,000} \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ \dots \\ \dots \\ e_{4,000} \end{bmatrix}$$

$$\mathbf{y} = [\mathbf{M} - 2\mathbf{p}']\boldsymbol{\alpha} + \mathbf{e} = \mathbf{W}\boldsymbol{\alpha} + \mathbf{e}$$

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{W}\boldsymbol{\alpha} + \mathbf{e}$$

$$\begin{bmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{W} \\ \mathbf{W}'\mathbf{X} & \mathbf{W}'\mathbf{W} + \mathbf{I} \frac{\sigma_e^2}{\sigma_\alpha^2} \end{bmatrix} \begin{bmatrix} \mathbf{b} \\ \boldsymbol{\alpha} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{W}'\mathbf{y} \end{bmatrix}$$

$$\sigma_\alpha^2 = \frac{\sigma_A^2}{50,000}$$

Genomic selection

The SNP-BLUP or RR-BLUP

$$m_y = 10,000$$

$$\begin{bmatrix} 602 \\ -440 \\ 103 \\ \dots \\ \dots \\ -657 \end{bmatrix} = \alpha_1 \begin{bmatrix} 2-2p_1 \\ 0-2p_1 \\ 2-2p_1 \\ \dots \\ \dots \\ 1-2p_1 \end{bmatrix} + \alpha_2 \begin{bmatrix} 1-2p_2 \\ 1-2p_2 \\ 0-2p_2 \\ \dots \\ \dots \\ 2-2p_2 \end{bmatrix} + \alpha_3 \begin{bmatrix} 1-2p_3 \\ 2-2p_3 \\ 1-2p_3 \\ \dots \\ \dots \\ 0-2p_3 \end{bmatrix} + \dots + \alpha_{50,000} \begin{bmatrix} 1-2p_{50,000} \\ 0-2p_{50,000} \\ 0-2p_{50,000} \\ \dots \\ \dots \\ 2-2p_{50,000} \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ \dots \\ \dots \\ e_{4,000} \end{bmatrix}$$

$$\mathbf{y} = [\mathbf{M} - 2\mathbf{p}']\boldsymbol{\alpha} + \mathbf{e} = \mathbf{W}\boldsymbol{\alpha} + \mathbf{e}$$

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{W}\boldsymbol{\alpha} + \mathbf{e}$$

$$\begin{bmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{W} \\ \mathbf{W}'\mathbf{X} & \mathbf{W}'\mathbf{W} + \mathbf{I} \frac{\sigma_e^2}{\sigma_\alpha^2} \end{bmatrix} \begin{bmatrix} \mathbf{b} \\ \boldsymbol{\alpha} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{W}'\mathbf{y} \end{bmatrix}$$

$$\sigma_\alpha^2 = \frac{\sigma_A^2}{\sum_{i=1}^{50,000} 2p_i q_i}$$

Genomic selection

The SNP-BLUP or RR-BLUP

$$m_y = 10,000$$

$$\begin{bmatrix} 602 \\ -440 \\ 103 \\ \dots \\ \dots \\ -657 \end{bmatrix} = \alpha_1 \begin{bmatrix} 2-2p_1 \\ 0-2p_1 \\ 2-2p_1 \\ \dots \\ \dots \\ 1-2p_1 \end{bmatrix} + \alpha_2 \begin{bmatrix} 1-2p_2 \\ 1-2p_2 \\ 0-2p_2 \\ \dots \\ \dots \\ 2-2p_2 \end{bmatrix} + \alpha_3 \begin{bmatrix} 1-2p_3 \\ 2-2p_3 \\ 1-2p_3 \\ \dots \\ \dots \\ 0-2p_3 \end{bmatrix} + \dots + \alpha_{50,000} \begin{bmatrix} 1-2p_{50,000} \\ 0-2p_{50,000} \\ 0-2p_{50,000} \\ \dots \\ \dots \\ 2-2p_{50,000} \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ \dots \\ \dots \\ e_{4,000} \end{bmatrix}$$

$$\sigma_\alpha^2 = E(\alpha^2 - 0) = E(\alpha^2) \rightarrow \hat{\sigma}_\alpha^2 = \alpha^2 \quad \sigma_A^2 = \sum_1^{50,000} 2p_i q_i \alpha^2 = \sum_1^{50,000} 2p_i q_i \sigma_\alpha^2$$

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

$$\sigma_\alpha^2 = \frac{\sigma_A^2}{\sum_1^{50,000} 2p_i q_i}$$

Genomic selection

The SNP-BLUP or RR-BLUP

$$\begin{bmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{W} \\ \mathbf{W}'\mathbf{X} & \mathbf{W}'\mathbf{W} + \mathbf{I} \frac{\sigma_e^2}{\sigma_\alpha^2} \end{bmatrix} \begin{bmatrix} \mathbf{b} \\ \boldsymbol{\alpha} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{W}'\mathbf{y} \end{bmatrix}$$

$$\hat{a} = \mu + \hat{\alpha}_1 w_1 + \hat{\alpha}_2 w_2 + \hat{\alpha}_3 w_3 + \cdots + \hat{\alpha}_{50,000} w_{50,000}$$

$$\hat{\alpha}_1 = 0.10, \hat{\alpha}_2 = -0.35, \hat{\alpha}_3 = 2.11, \cdots, \hat{\alpha}_{50,000} = 0.22$$

[AG, AA, AG,..., TT] with or without data

$$\hat{a} = 0.10 \cdot (1 - 2p_1) - 0.35 \cdot (2 - 2p_2) + 2.11 \cdot (1 - 2p_3) + \cdots + 0.22 \cdot (0 - 2p_{50,000}) = -160$$

Genomic selection

The G-BLUP

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

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

$$\hat{a} = \mu + \hat{\alpha}_1 \mathbf{w}_1 + \hat{\alpha}_2 \mathbf{w}_2 + \hat{\alpha}_3 \mathbf{w}_3 + \cdots + \hat{\alpha}_{50,000} \mathbf{w}_{50,000}$$

$$\mathbf{a} = \mathbf{W}\boldsymbol{\alpha} = \alpha_1 \mathbf{w}_1 + \alpha_2 \mathbf{w}_2 + \alpha_3 \mathbf{w}_3 + \cdots + \alpha_{50,000} \mathbf{w}_{50,000}$$

$$\text{var}(\mathbf{a}) = \text{var}(\mathbf{W}\boldsymbol{\alpha}) = \mathbf{W} \text{var}(\boldsymbol{\alpha}) \mathbf{W}' = \mathbf{W}\mathbf{W}' \sigma_{\alpha}^2 = \mathbf{W}\mathbf{W}' \frac{\sigma_A^2}{\sum_{i=1}^n 2p_i q_i} = \frac{\mathbf{W}\mathbf{W}'}{\sum_{i=1}^n 2p_i q_i} \cdot \sigma_A^2 = \mathbf{G} \cdot \sigma_A^2$$

Genomic selection

The G-BLUP

	SNP1	SNP2	SNP3	SNP4	SNP5	SNP6	SNP7	SNP8
Animal 1	AA	CC	AA	TT	GT	GA	GG	TT
Animal 2	GG	CC	AA	TT	TT	GA	GG	CT
Animal 3	GG	GG	AC	TC	GG	GG	AA	CC
Animal 4	AA	CG	AC	TC	GG	GG	AA	CC

Major allele frequency: value 2

SNP1			SNP2			SNP3			SNP4		
AA	AG	GG	CC	CG	GG	AA	AC	CC	TT	TC	CC
2	1	0	2	1	0	2	1	0	2	1	1
SNP5			SNP6			SNP7			SNP8		
GG	GA	AA	GG	GA	AA	GG	GA	AA	CC	CT	TT
2	1	0	2	1	0	2	1	0	2	1	0

Genomic selection

The G-BLUP

$$\begin{bmatrix} 10,602 \\ 11,440 \\ 9,103 \\ 9,657 \end{bmatrix} = \begin{bmatrix} \mu \\ \mu \\ \mu \\ \mu \end{bmatrix} + \alpha_1 \begin{bmatrix} 2 \\ 0 \\ 0 \\ 2 \end{bmatrix} + \alpha_2 \begin{bmatrix} 2 \\ 2 \\ 0 \\ 1 \end{bmatrix} + \alpha_3 \begin{bmatrix} 2 \\ 2 \\ 1 \\ 1 \end{bmatrix} + \alpha_4 \begin{bmatrix} 2 \\ 2 \\ 1 \\ 1 \end{bmatrix} + \alpha_5 \begin{bmatrix} 1 \\ 0 \\ 2 \\ 2 \end{bmatrix} + \alpha_6 \begin{bmatrix} 1 \\ 1 \\ 2 \\ 2 \end{bmatrix} + \alpha_7 \begin{bmatrix} 2 \\ 2 \\ 0 \\ 0 \end{bmatrix} + \alpha_8 \begin{bmatrix} 0 \\ 0 \\ 2 \\ 2 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \end{bmatrix}$$

$$\begin{bmatrix} 10,602 \\ 11,440 \\ 9,103 \\ 9,657 \end{bmatrix} = \begin{bmatrix} \mu \\ \mu \\ \mu \\ \mu \end{bmatrix} + \begin{bmatrix} 2 & 2 & 2 & 2 & 1 & 1 & 2 & 0 \\ 0 & 2 & 2 & 2 & 0 & 1 & 2 & 0 \\ 0 & 0 & 1 & 1 & 2 & 2 & 0 & 2 \\ 2 & 1 & 1 & 1 & 2 & 2 & 0 & 2 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \\ \alpha_5 \\ \alpha_6 \\ \alpha_7 \\ \alpha_8 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \end{bmatrix} = \mu + \mathbf{M}\alpha + \mathbf{e}$$

Genomic selection

The G-BLUP

$$\mathbf{y} = [\mathbf{M} - 2\mathbf{p}']\boldsymbol{\alpha} + \mathbf{e} = \mathbf{W}\boldsymbol{\alpha} + \mathbf{e}$$

$$2 \cdot \mathbf{p}' = \begin{bmatrix} 2 \\ 2 \\ 2 \\ 2 \\ 2 \\ 2 \\ 2 \\ 2 \end{bmatrix} [0.50 \quad 0.625 \quad 0.75 \quad 0.75 \quad 0.625 \quad 0.75 \quad 0.50 \quad 0.50] = \begin{bmatrix} 1 & 1.25 & 1.5 & 1.5 & 1.25 & 1.5 & 1 & 1 \\ 1 & 1.25 & 1.5 & 1.5 & 1.25 & 1.5 & 1 & 1 \\ 1 & 1.25 & 1.5 & 1.5 & 1.25 & 1.5 & 1 & 1 \\ 1 & 1.25 & 1.5 & 1.5 & 1.25 & 1.5 & 1 & 1 \end{bmatrix}$$

$$\mathbf{W} = \mathbf{M} - 2\mathbf{p}' = \begin{bmatrix} 2 & 2 & 2 & 2 & 1 & 1 & 2 & 0 \\ 0 & 2 & 2 & 2 & 0 & 1 & 2 & 0 \\ 0 & 0 & 1 & 1 & 2 & 2 & 0 & 2 \\ 2 & 1 & 1 & 1 & 2 & 2 & 0 & 2 \end{bmatrix} - \begin{bmatrix} 1 & 1.25 & 1.5 & 1.5 & 1.25 & 1.5 & 1 & 1 \\ 1 & 1.25 & 1.5 & 1.5 & 1.25 & 1.5 & 1 & 1 \\ 1 & 1.25 & 1.5 & 1.5 & 1.25 & 1.5 & 1 & 1 \\ 1 & 1.25 & 1.5 & 1.5 & 1.25 & 1.5 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0.75 & 0.5 & 0.5 & -0.25 & -0.5 & 1 & -1 \\ -1 & 0.75 & 0.5 & 0.5 & -1.25 & -0.5 & 1 & -1 \\ -1 & -1.25 & -0.5 & -0.5 & 0.75 & 0.5 & -1 & 1 \\ 1 & -0.25 & -0.5 & -0.5 & 0.75 & 0.5 & -1 & 1 \end{bmatrix}$$

Genomic selection

The G-BLUP

$$\sum_1^8 2p_i q_i = \sum_1^8 2p_i (1 - p_i) = 2 \cdot (0.50 \cdot 0.50 + 0.375 \cdot 0.625 + 0.25 \cdot 0.75 + 0.25 \cdot 0.75 + 0.375 \cdot 0.625 + 0.25 \cdot 0.75 + 0.50 \cdot 0.50 + 0.50 \cdot 0.50) = 3.56$$

$$\mathbf{G} = \mathbf{W}\mathbf{W}' \frac{1}{3.56} = \begin{bmatrix} 4.38 & 2.63 & -4.88 & -2.13 \\ & 5.88 & -3.625 & -4.88 \\ & & 5.88 & 2.63 \\ & & & 4.38 \end{bmatrix} \frac{1}{3.56} = \begin{bmatrix} 1.23 & 0.74 & -1.37 & -0.60 \\ & 1.65 & -1.02 & -1.37 \\ & & 1.65 & 0.74 \\ & & & 1.23 \end{bmatrix}$$

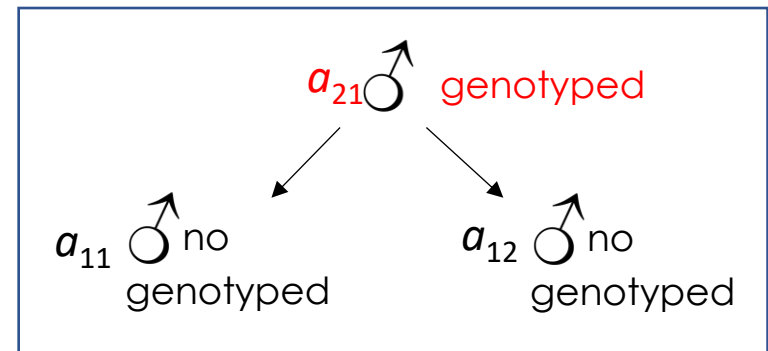
Genomic selection

Single step G-BLUP

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Za} + \mathbf{e} = \mathbf{Xb} + \mathbf{Z} \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \end{bmatrix} + \mathbf{e}$$

$\mathbf{a}_1$ no-genotyped
 $\mathbf{a}_2$ genotyped

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} = \left[\begin{array}{cc|c} 1 & 0.25 & 0.5 \\ 0.25 & 1 & 0.5 \\ \hline 0.5 & 0.5 & 1 \end{array} \right]$$



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

$$\mathbf{H}^{-1} = \mathbf{A}^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & \mathbf{G}^{-1} - \mathbf{A}_{22}^{-1} \end{bmatrix}$$

Genomic selection

- The values of the regression coefficients are not the additive gene values, since introducing new markers or removing some present in the equation changes the regression coefficients.
- The SNPs may be associated with gene effects or not (in the latter case, the value of their coefficient would be zero), and an SNP may be correlated with the effects of multiple genes. We estimate regression coefficients, not gene effects.

The Bayesian alphabet

PRIOR INFORMATION

$$f(\mathbf{u}) \sim N(\mathbf{0}, I\sigma_u^2) \rightarrow \sigma_i^2 = \sigma_u^2 \longrightarrow \text{SNP-BLUP}$$

$$y = F + z_1 u_1 + z_2 u_2 + z_3 u_3 + \cdots + z_{50,000} u_{50,000}$$

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

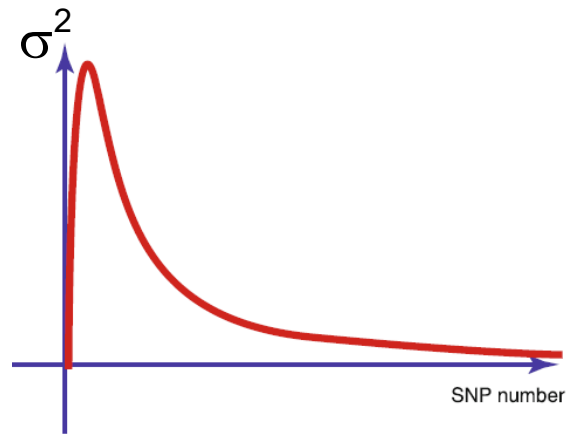
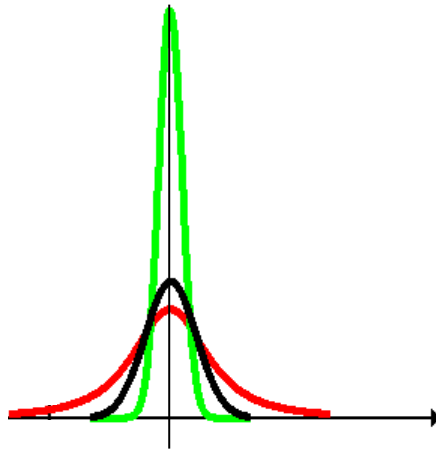
$$\begin{bmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z} \\ \mathbf{Z}'\mathbf{X} & \mathbf{Z}'\mathbf{Z} + I\frac{\sigma_e^2}{\sigma_u^2} \end{bmatrix} \begin{bmatrix} \mathbf{b} \\ \mathbf{u} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{Z}'\mathbf{y} \end{bmatrix}$$

$$\sigma_u^2 = \frac{\sigma_A^2}{50,000}$$

The Bayesian alphabet

PRIOR INFORMATION

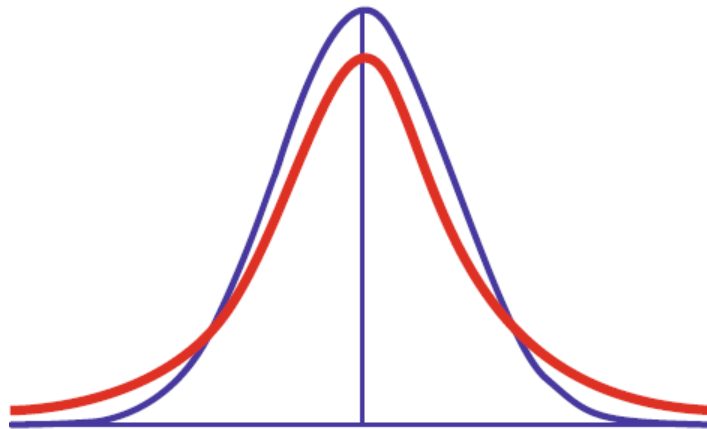
$$f(\mathbf{u}) \sim N(\mathbf{0}, \mathbf{D}) \rightarrow \left\{ \begin{array}{l} u_i \sim N(0, \sigma_i^2) \\ \sigma_i^2 \sim \text{IGamma}(v_e, S_e) \end{array} \right\} \rightarrow \text{Bayes A}$$



The Bayesian alphabet

PRIOR INFORMATION

$$f(\mathbf{u}) \sim t(\mathbf{0}, \mathbf{I}\sigma_u^2) \rightarrow u_i \sim t_v(0, \sigma_u^2) \rightarrow \text{Bayes A}$$



The Bayesian alphabet

PRIOR INFORMATION

$$f(\mathbf{u}) \sim N(\mathbf{0}, \mathbf{D}) \rightarrow \left\{ \begin{array}{c} u_i \sim N(0, \sigma_i^2) \\ \sigma_i^2 \sim \left\{ \begin{array}{cc} 0 & \xrightarrow{\text{PROB}} \pi \\ \text{IGamma}(v_e, S_e) & \xrightarrow{\text{PROB}} 1 - \pi \end{array} \right\} \end{array} \right\} \rightarrow \text{BayesB}$$

π is subjective prior information

π can be easily manipulated.

π should be around 0.95, but often it is 0.999 or more

The Bayesian alphabet

PRIOR INFORMATION

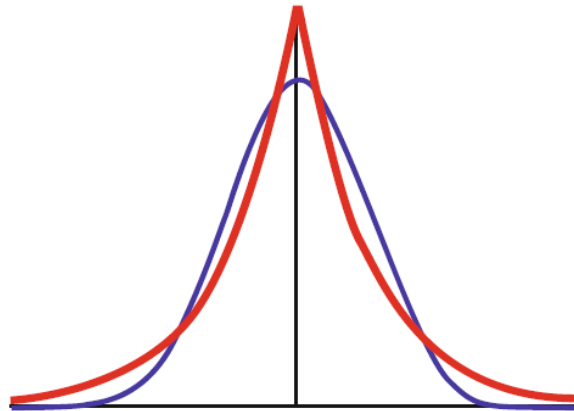
$$f(\mathbf{u}) \sim N(\mathbf{0}, I\sigma_u^2) \rightarrow \left\{ \begin{array}{c} u_i \sim N(0, \sigma_u^2) \\ \sigma_u^2 \sim \left\{ \begin{array}{cc} 0 & \xrightarrow{\text{PROB}} \pi \\ \text{Constant or IG} & \xrightarrow{\text{PROB}} 1 - \pi \end{array} \right\} \end{array} \right\} \rightarrow \text{Bayes C}$$

$$f(\mathbf{u}) \sim N(\mathbf{0}, I\sigma_u^2) \rightarrow \left\{ \begin{array}{c} u_i \sim N(0, \sigma_u^2) \\ \sigma_u^2 \sim \left\{ \begin{array}{cc} 0 & \xrightarrow{\text{PROB}} \pi \\ \text{constant or IG} & \xrightarrow{\text{PROB}} 1 - \pi \end{array} \right\} \\ \pi \sim \text{Beta}(\alpha, \beta) \propto x^{\alpha-1} (1-x)^{\beta-1} \end{array} \right\} \rightarrow \text{Bayes C}\pi$$

The Bayesian alphabet

PRIOR INFORMATION

$$f(\mathbf{u}) \rightarrow \left\{ \begin{array}{l} u_i | \lambda \sim \frac{\lambda}{2} \exp(\lambda |u_i|) \\ \lambda^2 \sim \text{Gamma}(\alpha, \beta) \end{array} \right\} \rightarrow \text{Bayes L}$$



The Bayesian alphabet

PRIOR INFORMATION

$$f(\mathbf{u}) \sim N(\mathbf{0}, \mathbf{D}) \rightarrow \left\{ \begin{array}{l} u_i \sim N(0, \sigma_i^2) \\ \sigma_i^2 = 0 \xrightarrow{\text{PROB}} \pi_1 \\ \sigma_i^2 = 0.0001 \sigma_A^2 \xrightarrow{\text{PROB}} \pi_2 \\ \sigma_i^2 = 0.001 \sigma_A^2 \xrightarrow{\text{PROB}} \pi_3 \\ \sigma_i^2 = 0.01 \sigma_A^2 \xrightarrow{\text{PROB}} \pi_4 \end{array} \right\} \rightarrow \text{BayesR}$$

