

GENETIC AND GENOMIC EVALUATION NOTES

CHAPTER 03

POLYGENIC-GENOMIC MODELS WITH INCOMPLETE GENOTYPE INFORMATION

Genotyping of all animals in a population may not be feasible. Two alternative approaches could be considered in such cases:

- 1) a two-step procedure: a) predict or impute missing genotypes using pedigree and available genotype information (Gengler et al., 2007; VanRaden, 2011; VanRaden et al., 2013), and b) compute genomic predictions of marker effects using mixed model or Bayesian procedures (Meuwissen et al., 2001; VanRaden, 2008) based on actual and predicted genotypes; and
- 2) a joint polygenic-genomic evaluation with mixed model procedures that use phenotypic, pedigree, and genotypic information (Aguilar et al., 2009; Legarra et al., 2009). Errors of prediction or imputation of the missing genotypes make the first alternative less desirable than the second one.

Polygenic-Genomic Models with Incomplete Genotype Information

- a. [Prediction of unknown genotypes](#)
- b. [Polygenic-Genomic Marker Model with actual and predicted genotypes](#)
- c. [Polygenic-Genomic Animal Model with actual and predicted genotypes](#)
- d. [Polygenic-Genomic Animal Model with phenotypic, pedigree, and genotypic information](#)

Polygenic-Genomic Models with **actual** and **predicted** genotypes

Step 1: Prediction of unknown genotypes

Gene Content Model

Description based on the procedure developed by Gengler et al. (2007) to compute gene contents in large populations.

Consider the following model in scalar form:

$$\text{gene content} = \text{mean} + \text{breed regression} + \text{gene content deviation} + \text{error}$$

$$E[\text{gene content}] = \text{mean} + \text{breed regression}$$

$$\text{var}(\text{gene content}) = \text{var}(\text{gene content deviation}) + \text{var}(\text{error})$$

where

$$\text{gene content deviation} = \text{gene content} - (\text{mean} + \text{breed regression})$$

$$\text{error} = \text{pedigree and genotyping errors (Gengler et al., 2007)}$$

In matrix notation, the gene content model is:

$$y_k = 1_k b + Z_k Q_k g + \begin{bmatrix} I_k & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} d_k \\ d_u \end{bmatrix} + \varepsilon_k$$

$$E[y_k] = 1_k b + Z_k Q_k g$$

$$\text{var} \begin{bmatrix} d_k \\ d_u \\ \varepsilon \end{bmatrix} = \begin{bmatrix} A_{kk} * \sigma_{gc}^2 & A_{ku} * \sigma_{gc}^2 & 0 \\ A_{uk} * \sigma_{gc}^2 & A_{uu} * \sigma_{gc}^2 & 0 \\ 0 & 0 & I_k * \sigma_{\varepsilon}^2 \end{bmatrix}$$

$$\text{var}(y) = \begin{bmatrix} I_k & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} A_{kk} * \sigma_{gc}^2 & A_{ku} * \sigma_{gc}^2 \\ A_{uk} * \sigma_{gc}^2 & A_{uu} * \sigma_{gc}^2 \end{bmatrix} \begin{bmatrix} I_k & 0 \\ 0 & 0 \end{bmatrix} + I_k * \sigma_{\varepsilon}^2$$

$$\text{var}(y) = A_{kk} * \sigma_{gc}^2 + I_k * \sigma_{\varepsilon}^2$$

where

y = vector of observed single marker locus gene contents [0, 1, 2] for genotyped animals,

b = unknown mean,

g = vector unknown breed effects,

d_k = vector of gene content deviations for genotyped animals,

d_u = vector of gene content deviations for non-genotyped animals,

1_k = $k \times 1$ vector of 1's relating observed single marker locus gene contents to the overall mean,

I_k = identity incidence matrix relating observed single marker locus gene contents to gene content deviations,

Z_k = $k \times \text{nbg}$ incidence matrix relating genotyped animals to breed groups, nbg = number of breed groups,

Q_k = $\text{nbg} \times \text{nbr}$ incidence matrix relating breed groups to expected breed fractions for genotyped animals,

$A = \begin{bmatrix} A_{kk} & A_{ku} \\ A_{uk} & A_{uu} \end{bmatrix}$ = additive genetic relationship matrix including all genotyped and non-genotyped animals,

σ_{gc}^2 = gene content variance, and

σ_{ϵ}^2 = error variance due to pedigree and genotyping errors.

The mixed model equations for the gene content model are:

$$\begin{bmatrix} 1'1 & 1'Z_k Q_k & 1'I_k & 0 \\ Q'_k Z'_k 1 & Q'_k Z'_k Z_k Q_k & Q'_k Z'_k I_k & 0 \\ I_k 1 & I_k Z_k Q_k & I_k + A^{kk} * \frac{\sigma_{\epsilon}^2}{\sigma_{gc}^2} & A^{ku} * \frac{\sigma_{\epsilon}^2}{\sigma_{gc}^2} \\ 0 & 0 & A^{uk} * \frac{\sigma_{\epsilon}^2}{\sigma_{gc}^2} & I_u + A^{uu} * \frac{\sigma_{\epsilon}^2}{\sigma_{gc}^2} \end{bmatrix} \begin{bmatrix} b \\ g \\ d_k \\ d_u \end{bmatrix} = \begin{bmatrix} 1'y \\ Q'_k Z'_k y \\ I_k y \\ 0 \end{bmatrix}$$

The value of the ratio $\frac{\sigma_{\epsilon}^2}{\sigma_{gc}^2}$ is defined to be a small number (e.g., 0.01; Gengler et al., 2007).

Example

Prediction of the genetic value of animals for weaning weight using the following phenotypic and genomic information for 60 marker loci on 12 calves.

Anim	Breed Compsn	loc 1 ¹	loc 2 ¹	loc 3 ¹	loc 4 ¹	...	loc 60 ¹	Sex	WWt (kg)	Sire	Breed Comp	Dam	Breed Compsn
1	A	2	1	1	2	...	2	M	289	0	A	0	A
2	B	0	1	2	0	...	1	F	245	0	B	0	B
3	½ A ½ B	1	2	0	0	...	2	F	256	0	A	2	B
4	½ A ½ B	1	2	0	0	...	1	F	261	1	A	0	B
5	½ A ½ B	1	2	0	1	...	1	M	292	1	A	2	B
6	¾ A ¼ B	2	1	2	2	...	1	M	286	1	A	3	½ A ½ B
7	¼ A ¾ B	- ²	-	-	-	...	-	M	272	0	B	3	½ A ½ B
8	¾ A ¼ B	-	-	-	-	...	-	F	264	1	A	4	½ A ½ B
9	5/8 A 3/8 B	-	-	-	-	...	-	F	270	5	½ A ½ B	8	¾ A ¼ B
10	½ A ½ B	-	-	-	-	...	-	M	278	5	½ A ½ B	3	½ A ½ B
11	3/8 A 5/8 B	-	-	-	-	...	-	F	259	6	¾ A ¼ B	0	B
12	3/8 A 5/8 B	-	-	-	-	...	-	M	280	6	¾ A ¼ B	2	B

¹Number of “second” alleles (VanRaden, 2008).

²Non-genotyped animals.

To predict for the gene content for the 60 marker loci in non-genotyped animals 7 to 12, a series of 60 mixed model equations need to be solved (i.e., one for each marker locus).

The mixed model for marker locus 1 is:

$$\begin{bmatrix} 2 \\ 0 \\ 1 \\ 1 \\ 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} [mean] + \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1/2 & 1/2 \\ 1/2 & 1/2 \\ 3/4 & 1/4 \end{bmatrix} \begin{bmatrix} breed_A \\ breed_B \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} genanim_1 \\ genanim_2 \\ genanim_3 \\ genanim_4 \\ genanim_5 \\ genanim_6 \\ nongenanim_7 \\ nongenanim_8 \\ nongenanim_9 \\ nongenanim_{10} \\ nongenanim_{11} \\ nongenanim_{12} \end{bmatrix} + \begin{bmatrix} error_{anim 1} \\ error_{anim 2} \\ error_{anim 3} \\ error_{anim 4} \\ error_{anim 5} \\ error_{anim 6} \end{bmatrix}$$

The inverse of the additive relationship matrix for the Example is:

ainv											
2.8333333	0.5	0.5	-0.166667	-1	-1	0	-1	0	0	0	0
0.5	2.3333333	-0.666667	0	-1	0.5	0	0	0	0	0	-1
0.5	-0.666667	2.6666667	0	0.5	-1	-0.666667	0	0	-1	0	0
-0.166667	0	0	1.8333333	0	0	0	-1	0	0	0	0
-1	-1	0.5	0	3	0	0	0.5	-1	-1	0	0
-1	0.5	-1	0	0	2.8333333	0	0	0	0	-0.666667	-1
0	0	-0.666667	0	0	0	1.3333333	0	0	0	0	0
-1	0	0	-1	0.5	0	0	2.5	-1	0	0	0
0	0	0	0	-1	0	0	-1	2	0	0	0
0	0	-1	0	-1	0	0	0	0	2	0	0
0	0	0	0	0	-0.666667	0	0	0	0	1.3333333	0
0	-1	0	0	0	-1	0	0	0	0	0	2

Assuming that $\frac{\sigma_e^2}{\sigma_{gc}^2} = 0.01$, the LHS matrix, vector of effects, and the RHS of the mixed model equations are:

lhsgc															Effect	rhs
6.000	0.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	Mean	7.000
0.000	2.313	0.938	1.000	0.000	0.500	0.500	0.500	0.750	0.000	0.000	0.000	0.000	0.000	0.000	Breed _A	5.000
0.000	0.938	1.813	0.000	1.000	0.500	0.500	0.500	0.250	0.000	0.000	0.000	0.000	0.000	0.000	Breed _B	2.000
1.000	1.000	0.000	0.028	0.005	0.005	0.000	-0.010	-0.010	0.000	-0.010	0.000	0.000	0.000	0.000	\hat{d}_{k1}	2.000
1.000	0.000	1.000	0.005	1.023	-0.007	0.000	-0.010	0.005	0.000	0.000	0.000	0.000	0.000	-0.010	\hat{d}_{k2}	0.000
1.000	0.500	0.500	0.005	-0.007	0.277	0.000	0.005	-0.010	-0.007	0.000	0.000	-0.010	0.000	0.000	\hat{d}_{k3}	1.000
1.000	0.500	0.500	0.000	0.000	0.000	-0.265	0.000	0.000	0.000	-0.010	0.000	0.000	0.000	0.000	\hat{d}_{k4}	1.000
1.000	0.500	0.500	-0.010	-0.010	0.005	0.000	0.280	0.000	0.000	0.005	-0.010	-0.010	0.000	0.000	\hat{d}_{k5}	1.000
1.000	0.750	0.250	-0.010	0.005	-0.010	0.000	0.000	0.090	0.000	0.000	0.000	0.000	-0.005	-0.010	\hat{d}_{k6}	2.000
0.000	0.000	0.000	0.000	0.000	-0.007	0.000	0.000	0.000	0.013	0.000	0.000	0.000	0.000	0.000	\hat{d}_{k7}	0.000
0.000	0.000	0.000	-0.010	0.000	0.000	-0.010	0.005	0.000	0.000	0.025	-0.010	0.000	0.000	0.000	\hat{d}_{k8}	0.000
0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.010	0.000	0.000	-0.010	0.020	0.000	0.000	0.000	\hat{d}_{k9}	0.000
0.000	0.000	0.000	0.000	0.000	-0.010	0.000	-0.010	0.000	0.000	0.000	0.000	0.020	0.000	0.000	\hat{d}_{k10}	0.000
0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.005	0.000	0.000	0.000	0.000	0.010	0.000	\hat{d}_{k11}	0.000
0.000	0.000	0.000	0.000	-0.010	0.000	0.000	0.000	-0.010	0.000	0.000	0.000	0.000	0.000	0.020	\hat{d}_{k12}	0.000

The vector of **solutions** and their **standard errors** are:

Effect	SOL	SESOL
Mean ^o	0.53	0.26
Breed _A ^o	1.57	0.20
Breed _B ^o	-0.70	.
\hat{d}_{k1}	-2.30	0.98
\hat{d}_{k2}	0.19	0.78
\hat{d}_{k3}	0.38	1.60
\hat{d}_{k4}	0.09	1.67

\widehat{d}_{k5}	0.05	1.65
\widehat{d}_{k6}	5.41	2.85
\widehat{d}_{u7}	0.19	8.70
\widehat{d}_{u8}	-1.10	7.17
\widehat{d}_{u9}	-0.53	7.97
\widehat{d}_{u10}	0.22	7.12
\widehat{d}_{u11}	2.71	10.10
\widehat{d}_{u12}	2.80	7.20

Predictions of gene contents for marker locus j in non-genotyped animals were computed as

follows:

$$\widehat{g}_{c_{ij}} = \text{mean}^\circ + \sum_{b=1}^{n_b} \text{Prob}_{uib} * \text{Breed}_b^\circ + \widehat{d}_{uij}$$

where

mean° = generalized least squares solution for the overall mean,

n_b = number of breeds,

Prob_{uib} = expected fraction of breed b in non-genotyped animal i,

Breed_b° = generalized least squares solution for breed b,

\widehat{d}_{uij} = BLUP of non-genotyped animal i for marker locus j.

The matrix of variances of errors of prediction (VEP) for the vector of predicted gene contents for non-genotyped animals is:

$$\text{VEP}(\widehat{d}_u) = \text{var}(\widehat{d}_u - d_u) = K_{du} * (\text{ginvlhs}) * K_{du}'$$

where

K_{du} = $n_u \times n_{eq}$ matrix of coefficients for effects in the model (one row for each animal), where n_u = number of non-genotyped animals, and n_{eq} = number of equations. Each row has $1 + n_b + 1$ elements, i.e., a 1 for the mean, the probability of breed b in animal i for $b = 1$ to n_b , and a 1 for the animal.

$ginvlhs$ = $n_{eq} \times n_{eq}$ generalized inverse matrix of the left hand side of the MME.

Thus, the gene contents and their standard errors of prediction for marker locus 1 in non-genotyped animals 7 to 12 are:

Anim	Mean ^o	Prob _A	Br _A ^o	Prob _B	Br _B ^o	\hat{d}_{ui}	$\hat{g}c_{ui}$	$SEP(\hat{g}c_{ui})$
7	0.53	1	1.57	0	-0.70	0.19	0.59	8.71
8	0.53	0	1.57	1	-0.70	-1.10	0.43	7.2
9	0.53	0.5	1.57	0.5	-0.70	-0.53	0.72	7.99
10	0.53	0.5	1.57	0.5	-0.70	0.22	1.18	7.16
11	0.53	0.5	1.57	0.5	-0.70	2.71	3.39	10.12
12	0.53	0.75	1.57	0.25	-0.70	2.80	3.48	7.24

Predictions of gene content were highly inaccurate and outside the [0 to 2] range for some non-genotyped animals in most marker loci. **This occurred because of the small number of genotyped animals and the incompleteness of the pedigree in the example.** One alternative before performing the genomic evaluation in step 2 would be to restrict the predicted values of all estimates of gene content to the [0 to 2] range, i.e., any estimate below zero would be set to zero,

and any estimate above 2 would be set to 2. Thus, the predicted gene contents for animals 11 and 12 in the example would be set to 2.0. Consequently, the restricted estimates of gene content for the 6 non-genotyped animals in the example for marker 1 would be [0.59, 0.43, 0.72, 1.18, 2.0, 2.0].

Step 2: Computation of polygenic-genomic predictions using phenotypes, actual genotypes, and predicted genotypes.

The data for the example with predicted genotypes for non-genotyped animals restricted to the range [0 to 2] is:

Anim	Breed Compsn	loc 1 ¹	loc 2 ¹	loc 3 ¹	loc 4 ¹	...	loc 60 ¹	Sex	WWt (kg)	Sire	Breed Comp	Dam	Breed Compsn
1	A	2	1	1	2	...	2	M	289	0	A	0	A
2	B	0	1	2	0	...	1	F	245	0	B	0	B
3	½ A ½ B	1	2	0	0	...	2	F	256	0	A	2	B
4	½ A ½ B	1	2	0	0	...	1	F	261	1	A	0	B
5	½ A ½ B	1	2	0	1	...	1	M	292	1	A	2	B
6	¾ A ¼ B	2	1	2	2	...	1	M	286	1	A	3	½ A ½ B
7	¼ A ¾ B	0.6 ²	2	0	1.2	...	2	M	272	0	B	3	½ A ½ B
8	¾ A ¼ B	0.4	2	0	0	...	2	F	264	1	A	4	½ A ½ B
9	5/8 A 3/8 B	0.7	2	0.4	0.6	...	2	F	270	5	½ A ½ B	8	¾ A ¼ B
10	½ A ½ B	1.2	2	0.5	2	...	2	M	278	5	½ A ½ B	3	½ A ½ B
11	3/8 A 5/8 B	2	0.5	2	1.6	...	0	F	259	6	¾ A ¼ B	0	B
12	3/8 A 5/8 B	2	0.1	2	1.6	...	0	M	280	6	¾ A ¼ B	2	B

¹Number of “second” alleles (VanRaden, 2008).

²Predicted genotypes for non-genotyped animals (Gengler et al., 2007).

Additive Genetic and Residual Variances

Variance	Additive Genetic	Residual
Intrabreed AA	36 kg ²	49 kg ²
Intrabreed BB	44 kg ²	16 kg ²
Interbreed AB	22 kg ²	25 kg ²

Two alternative models:a) **Polygenic-Genomic Marker Model**b) **Polygenic-Genomic Animal Model****Polygenic-Genomic Marker Model with phenotypes, actual genotypes, and predicted genotypes**

Wean wt = overall mean + breed regression + heterosis regression + sex + sum regression markers 1 to 60 + residual

E [Wean wt] = overall mean + breed regression + heterosis regression + sex

Var (Wean wt) = sum [var(ma j regression), j = 1 to 60] + var (residual)

cov(Wean wt_{ik}, Wean wt_{i'k'}) = sum [cov(ma j regression, ma j' regression), j, j' = 1 to 60] + cov (residual_{ik}, residual_{i'k'})

cov(Wean wt_{ik}, Wean wt_{i'k'}) = 0 + cov (residual_{ik}, residual_{i'k'})

cov(Wean wt_{ik}, Wean wt_{i'k'}) = cov (residual_{ik}, residual_{i'k'})

The polygenic-genomic marker model, in matrix notation, is:

$$y = Xb + Z_a Q_a g_a + Z_n Q_n g_n + Z_{ma} m_a + \varepsilon$$

$$E[y] = Xb + Z_a Q_a g_a + Z_n Q_n g_n$$

$$\text{var} \begin{bmatrix} m_a \\ \varepsilon \end{bmatrix} = \begin{bmatrix} G_{ma} & 0 \\ 0 & R \end{bmatrix}$$

$$\text{var}(y) = Z_{ma}G_{ma}Z_{ma}' + R$$

where

- y = vector of animal records,
- b = vector of unknown fixed effects (mean, sex of calf),
- g_a = vector of unknown additive group genetic effects (breed),
- g_n = vector of unknown nonadditive group genetic effects (heterosis at 1 locus),
- m_a = vector of unknown random marker locus additive genomic effects,
- ε = vector of unknown random residual effects
- X = known incidence matrix relating records to fixed effects in vector b,
- Z_a = known incidence matrix relating records to elements of vector a_a ,
- Z_n = known incidence matrix relating records to random polygenic nonadditive genetic effects in vector a_n (a_n is assumed to be part of the residual),
- Z_{ma} = known incidence matrix relating records to elements of vector m_a ; elements of Z_{ma} will be 0, 1, 2 depending on the number of “second” alleles in marker locus j , $j = 1, \dots, 60$. Values [0,1, 2] are called gene content. Other parameterizations are: 1) [-1, 0, 1] obtained by subtracting 1 from [0, 1, 2]; and 2) deviate [-1, 0, 1] from $2(p_j - 0.5)$, which yields [-2 p_j , 1 - 2 p_j , 2 - 2 p_j], where 0.5 is the allele frequency used to center matrix Z_{ma} . Other frequencies used to center matrix Z_{ma} are the base population mean computed using a linear regression on gene content (Gengler et al., 2007) and the current population mean (Aguilar et al., 2009).
- Q_a = known incidence matrix relating random polygenic additive genetic effects to

additive genetic groups in vector g_a ,

Q_n = known incidence matrix relating random polygenic nonadditive genetic effects to nonadditive genetic groups in vector g_n ,

G_{ma} = $nma \times nma$ diagonal matrix of marker locus additive genomic variances (assumed to be equal for all markers associated with a trait),

$$= I * \sigma_{ma}^2$$

$$= I * \left\{ \sigma_a^2 / 2 \sum_{j=1}^{j=nma} p_j (1 - p_j) \right\}, \text{ where } \sigma_a^2 = \text{polygenic additive genetic variance, } nma = \text{number of marker loci (in linkage disequilibrium with QTL), and } p_j = \text{frequency of "second" allele in marker } j,$$

G_a = matrix of polygenic additive genetic variances and covariances among animals (purebred and crossbred) in the multibreed population; covariances are due to additive relationships among animals.

R = matrix of residual variances for purebred and crossbred animals (sums of nonadditive genetic and environmental variances).

The mixed model with the information of the 12 calves can be written as follows:

$$\begin{bmatrix} 289 \\ 245 \\ 256 \\ 261 \\ 292 \\ 286 \\ 272 \\ 264 \\ 270 \\ 278 \\ 259 \\ 280 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 1/2 & 1/2 & 1 & 0 & 1 \\ 1 & 1/2 & 1/2 & 1 & 0 & 1 \\ 1 & 1/2 & 1/2 & 1 & 1 & 0 \\ 1 & 3/4 & 1/4 & 1/2 & 1 & 0 \\ 1 & 1/4 & 3/4 & 1/2 & 1 & 0 \\ 1 & 3/4 & 1/4 & 1/2 & 0 & 1 \\ 1 & 5/8 & 3/8 & 1/2 & 0 & 1 \\ 1 & 1/2 & 1/2 & 1/2 & 1 & 0 \\ 1 & 3/8 & 5/8 & 3/4 & 0 & 1 \\ 1 & 3/8 & 5/8 & 3/4 & 1 & 0 \end{bmatrix} \begin{bmatrix} \text{mean} \\ \text{breed}_A \\ \text{breed}_B \\ \text{heterosis}_{AB} \\ \text{sex}_M \\ \text{sex}_F \end{bmatrix} + \begin{bmatrix} 2 & 1 & 1 & 2 & \dots & 2 \\ 0 & 1 & 2 & 0 & \dots & 1 \\ 1 & 2 & 0 & 0 & \dots & 2 \\ 1 & 2 & 0 & 0 & \dots & 1 \\ 1 & 2 & 0 & 1 & \dots & 1 \\ 2 & 1 & 2 & 2 & \dots & 1 \\ 0.6 & 2 & 0 & 1.2 & \dots & 2 \\ 0.4 & 2 & 0 & 0 & \dots & 2 \\ 0.7 & 2 & 0.4 & 0.6 & \dots & 2 \\ 1.2 & 2 & 0.5 & 2 & \dots & 2 \\ 2 & 0.5 & 2 & 1.6 & \dots & 0 \\ 2 & 0.1 & 2 & 1.6 & \dots & 0 \end{bmatrix} \begin{bmatrix} ma_1 \\ ma_2 \\ ma_3 \\ ma_4 \\ \vdots \\ ma_{60} \end{bmatrix} + \begin{bmatrix} \text{residual}_{anim 1} \\ \text{residual}_{anim 2} \\ \text{residual}_{anim 3} \\ \text{residual}_{anim 4} \\ \text{residual}_{anim 5} \\ \text{residual}_{anim 6} \\ \text{residual}_{anim 7} \\ \text{residual}_{anim 8} \\ \text{residual}_{anim 9} \\ \text{residual}_{anim 10} \\ \text{residual}_{anim 11} \\ \text{residual}_{anim 12} \end{bmatrix}$$

invr = R ⁻¹											
0	0	0	0	0.0307692	0	0	0	0	0	0	0
0	0	0	0	0	0.0212766	0	0	0	0	0	0
0	0	0	0	0	0	0.0327869	0	0	0	0	0
0	0	0	0	0	0	0	0.0212766	0	0	0	0
0	0	0	0	0	0	0	0	0.021025	0	0	0
0	0	0	0	0	0	0	0	0	0.0222222	0	0
0	0	0	0	0	0	0	0	0	0	0.0302457	0
0	0	0	0	0	0	0	0	0	0	0	0.0302457

The diagonal matrix of marker locus additive genomic variances is equal to:

$$G_{ma} = I * \left\{ \sigma_a^2 / 2 \sum_{j=1}^{j=60} p_j (1 - p_j) \right\}$$

Where

$$\left\{ \sigma_a^2 / 2 \sum_{j=1}^{j=60} p_j (1 - p_j) \right\} = 36 / 20.974159 = 1.7163978$$

Matrix G_{ma}^{-1} is equal to:

$gmainv = G_{ma}^{-1}$									
0.5826155	0	0	0	0	0	...	0	0	0
0	0.5826155	0	0	0	0	...	0	0	0
0	0	0.5826155	0	0	0	...	0	0	0
0	0	0	0.5826155	0	0	...	0	0	0
0	0	0	0	0.5826155	0	...	0	0	0
0	0	0	0	0	0.5826155	...	0	0	0
...
0	0	0	0	0	0	...	0.5826155	0	0
0	0	0	0	0	0	...	0	0.5826155	0

$gmainv = G_{ma}^{-1}$									
0	0	0	0	0	0	...	0	0	0.5826155

The LHS matrix, the vector of effects, and the RHS vector of the mixed model equations are:

lhs																Effects	rhs
0.354	0.154	0.201	0.197	0.158	0.197	0.366	0.502	0.309	0.256	0.298	0.422	...	0.457	0.624	0.442	Mean	94.88
0.154	0.092	0.062	0.095	0.082	0.071	0.198	0.232	0.098	0.145	0.144	0.182	...	0.18	0.278	0.215	Breed _A	42.1
0.201	0.062	0.139	0.102	0.075	0.125	0.168	0.27	0.211	0.111	0.154	0.24	...	0.277	0.346	0.227	Breed _B	52.78
0.197	0.095	0.102	0.156	0.092	0.105	0.239	0.307	0.112	0.153	0.194	0.261	...	0.289	0.336	0.231	Heterosis _{AB}	53.36
0.158	0.082	0.075	0.092	0.158	0	0.22	0.216	0.123	0.187	0.204	0.188	...	0.235	0.261	0.203	Sex _M	44.53
0.197	0.071	0.125	0.105	0	0.197	0.146	0.286	0.185	0.069	0.094	0.234	...	0.222	0.362	0.239	Sex _F	50.35
0.366	0.198	0.168	0.239	0.22	0.146	1.141	0.445	0.368	0.46	0.426	0.522	...	0.515	0.67	0.387	ma ₁	100.49
0.502	0.232	0.27	0.307	0.216	0.286	0.445	1.454	0.225	0.223	0.388	0.514	...	0.626	0.835	0.76	ma ₂	134.74
0.309	0.098	0.211	0.112	0.123	0.185	0.368	0.225	1.180	0.368	0.259	0.472	...	0.41	0.618	0.208	ma ₃	81.3
0.256	0.145	0.111	0.153	0.187	0.069	0.46	0.223	0.368	1.032	0.339	0.402	...	0.385	0.51	0.196	ma ₄	71.16
0.298	0.144	0.154	0.194	0.204	0.094	0.426	0.388	0.259	0.339	0.987	0.366	...	0.478	0.5	0.348	ma ₅	81.72
0.422	0.182	0.24	0.261	0.188	0.234	0.522	0.514	0.472	0.402	0.366	1.217	...	0.561	0.763	0.405	ma ₆	113.08
...
0.457	0.18	0.277	0.289	0.235	0.222	0.515	0.626	0.41	0.385	0.478	0.561	...	1.317	0.804	0.496	ma ₅₈	123.22
0.624	0.278	0.346	0.336	0.261	0.362	0.67	0.835	0.618	0.51	0.5	0.763	...	0.804	1.776	0.715	ma ₅₉	167.2
0.442	0.215	0.227	0.231	0.203	0.239	0.387	0.76	0.208	0.196	0.348	0.405	...	0.496	0.715	1.322	ma ₆₀	118.74

The vector of **solutions** and their **standard errors** are:

Effect	SOL	SESOL	Effect	SOL	SESOL	Effect	SOL	SESOL
Mean ^o	133.91	6.1	\widehat{ma}_{17}	0	1.31	\widehat{ma}_{39}	0.01	1.25
Breed _A ^o	77.5	6.4	\widehat{ma}_{18}	-0.15	1.27	\widehat{ma}_{40}	-0.24	1.29
Breed _B ^o	56.41	6.54	\widehat{ma}_{19}	0.02	1.26	\widehat{ma}_{41}	0.3	1.24

Heterosis _{AB} ^o	7.4	6.98	\widehat{ma}_{20}	0.05	1.25	\widehat{ma}_{42}	0.18	1.28
Sex _M ^o	78.21	3.78	\widehat{ma}_{21}	0	1.31	\widehat{ma}_{43}	0.05	1.28
Sex _F ^o	55.7	3.75	\widehat{ma}_{22}	-0.02	1.27	\widehat{ma}_{44}	0.07	1.29
\widehat{ma}_1	-0.04	1.29	\widehat{ma}_{23}	-0.32	1.27	\widehat{ma}_{45}	0.01	1.28
\widehat{ma}_2	-0.03	1.28	\widehat{ma}_{24}	0.14	1.25	\widehat{ma}_{46}	-0.44	1.25
\widehat{ma}_3	0.03	1.26	\widehat{ma}_{25}	-0.14	1.28	\widehat{ma}_{47}	0.19	1.27
\widehat{ma}_4	0.19	1.27	\widehat{ma}_{26}	-0.23	1.27	\widehat{ma}_{48}	-0.05	1.29
\widehat{ma}_5	-0.08	1.27	\widehat{ma}_{27}	0.2	1.28	\widehat{ma}_{49}	-0.2	1.26
\widehat{ma}_6	-0.01	1.28	\widehat{ma}_{28}	-0.25	1.27	\widehat{ma}_{50}	0.07	1.28
\widehat{ma}_7	0.04	1.29	\widehat{ma}_{29}	-0.17	1.26	\widehat{ma}_{51}	-0.08	1.25
\widehat{ma}_8	-0.19	1.26	\widehat{ma}_{30}	-0.23	1.27	\widehat{ma}_{52}	-0.07	1.29
\widehat{ma}_9	-0.28	1.25	\widehat{ma}_{31}	0.08	1.28	\widehat{ma}_{53}	0.06	1.27
\widehat{ma}_{10}	-0.05	1.3	\widehat{ma}_{32}	-0.09	1.29	\widehat{ma}_{54}	-0.04	1.29
\widehat{ma}_{11}	0.14	1.29	\widehat{ma}_{33}	-0.07	1.28	\widehat{ma}_{55}	0	1.31
\widehat{ma}_{12}	0.34	1.26	\widehat{ma}_{34}	-0.02	1.27	\widehat{ma}_{56}	0.25	1.28
\widehat{ma}_{13}	-0.04	1.25	\widehat{ma}_{35}	0.02	1.27	\widehat{ma}_{57}	0	1.31
\widehat{ma}_{14}	0.17	1.28	\widehat{ma}_{36}	0.22	1.25	\widehat{ma}_{58}	0.21	1.27
\widehat{ma}_{15}	-0.31	1.27	\widehat{ma}_{37}	-0.08	1.29	\widehat{ma}_{59}	0.1	1.28
\widehat{ma}_{16}	0.03	1.27	\widehat{ma}_{38}	0.27	1.25	\widehat{ma}_{60}	-0.16	1.28

Prediction of polygenic-genomic values

Predictions of additive polygenic-genomic values (\hat{u}_{ai}) are computed as the sum of the breed effect, the sum of the weighted random effects of the 60 markers, and the animal genomic deviation from the breed effect, i.e.,

$$\hat{u}_{ai} = \text{Prob}_{A\text{anim } i} * (\text{Breed}_A^\circ - \text{Breed}_B^\circ) + \sum_{j=1}^{60} w_{\text{ma } j \text{ Anim } i} * \hat{m}_j$$

where

$\text{Prob}_{A\text{anim } i}$ = expected fraction of breed A in animal i,

Breed_A° = generalized least squares solution for breed A,

Breed_B° = generalized least squares solution for breed B,

$w_{\text{ma } j \text{ Anim } i}$ = number of “second” alleles in marker locus j of animal i, and

\hat{m}_j = BLUP of m_j .

The matrix of variances of errors of prediction (VEP) for the vector of predicted genetic values (\hat{u}_a) is:

$$\text{VEP}(\hat{u}_a) = \text{var}(\hat{u}_a - u_a) = \mathbf{K}_a * (\text{ginvlhs}) * \mathbf{K}_a'$$

where

\mathbf{K}_a = $n_{\text{anim}} \times n_{\text{eq}}$ matrix specifying multipliers for breed differences, **allele difference in marker locus j, j = 1, 2, 3, 4, ..., 60**, and random animal deviations, where n_{anim} = number of animals, and n_{eq} = number of equations, and

ginvlhs = $n_{\text{eq}} \times n_{\text{eq}}$ generalized inverse matrix of the left hand side of the MME.

The standard errors of prediction (SEP) of the \hat{u}_{ai} are computed as the square roots of the diagonal elements of the $VEP(\hat{u}_{ai})$ matrix.

Matrix K_a for the Example is:

ka															
0.00	1.00	-1.00	0.00	0.00	0.00	2.00	1.00	1.00	2.00	1.00	1.00	...	0.00	2.00	2.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	2.00	0.00	0.00	1.00	...	1.00	2.00	1.00
0.00	0.50	-0.50	0.00	0.00	0.00	1.00	2.00	0.00	0.00	1.00	1.00	...	1.00	1.00	2.00
0.00	0.50	-0.50	0.00	0.00	0.00	1.00	2.00	0.00	0.00	0.00	2.00	...	1.00	2.00	1.00
0.00	0.50	-0.50	0.00	0.00	0.00	1.00	2.00	0.00	1.00	1.00	1.00	...	2.00	2.00	1.00
0.00	0.75	-0.75	0.00	0.00	0.00	2.00	1.00	2.00	2.00	1.00	2.00	...	1.00	2.00	1.00
0.00	0.25	-0.25	0.00	0.00	0.00	0.58	2.00	0.00	0.00	1.18	0.96	...	1.45	0.40	2.00
0.00	0.75	-0.75	0.00	0.00	0.00	0.42	2.00	0.00	0.00	0.00	0.62	...	0.00	2.00	2.00
0.00	0.63	-0.63	0.00	0.00	0.00	0.72	2.00	0.00	0.40	0.64	0.24	...	1.77	2.00	2.00
0.00	0.50	-0.50	0.00	0.00	0.00	1.19	2.00	0.00	0.54	2.00	0.11	...	2.00	1.93	2.00
0.00	0.38	-0.38	0.00	0.00	0.00	2.00	0.51	2.00	2.00	1.65	2.00	...	2.00	2.00	0.00
0.00	0.38	-0.38	0.00	0.00	0.00	2.00	0.10	2.00	2.00	1.59	2.00	...	2.00	2.00	0.00

We can verify that predictions for markers 1 to 60 are not confounded by computing matrix $K_a * \text{ginvlhs} * \text{lhs}$. If multipliers for marker loci $j, j = 1, \dots, 60$ in matrix K_a are recreated exactly in matrix $(K_a * \text{ginvlhs} * \text{lhs})$, then they were predicted unbiasedly.

Matrix $K_a * \text{ginvlhs} * \text{lhs}$ is:

kagl = $K_a * \text{ginvlhs} * \text{lhs}$															
0.00	1.00	-1.00	0.00	0.00	0.00	2.00	1.00	1.00	2.00	1.00	1.00	...	0.00	2.00	2.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	2.00	0.00	0.00	1.00	...	1.00	2.00	1.00
0.00	0.50	-0.50	0.00	0.00	0.00	1.00	2.00	0.00	0.00	1.00	1.00	...	1.00	1.00	2.00
0.00	0.50	-0.50	0.00	0.00	0.00	1.00	2.00	0.00	0.00	0.00	2.00	...	1.00	2.00	1.00
0.00	0.50	-0.50	0.00	0.00	0.00	1.00	2.00	0.00	1.00	1.00	1.00	...	2.00	2.00	1.00

kagl = $K_a * \text{ginvlhs} * \text{lhs}$															
0.00	0.75	-0.75	0.00	0.00	0.00	2.00	1.00	2.00	2.00	1.00	2.00	...	1.00	2.00	1.00
0.00	0.25	-0.25	0.00	0.00	0.00	0.58	2.00	0.00	0.00	1.18	0.96	...	1.45	0.40	2.00
0.00	0.75	-0.75	0.00	0.00	0.00	0.42	2.00	0.00	0.00	0.00	0.62	...	0.00	2.00	2.00
0.00	0.63	-0.63	0.00	0.00	0.00	0.72	2.00	0.00	0.40	0.64	0.24	...	1.77	2.00	2.00
0.00	0.50	-0.50	0.00	0.00	0.00	1.19	2.00	0.00	0.54	2.00	0.11	...	2.00	1.93	2.00
0.00	0.38	-0.38	0.00	0.00	0.00	2.00	0.51	2.00	2.00	1.65	2.00	...	2.00	2.00	0.00
0.00	0.38	-0.38	0.00	0.00	0.00	2.00	0.10	2.00	2.00	1.59	2.00	...	2.00	2.00	0.00

Thus, genomic predictions for all animals were unbiased because $K_a * \text{ginvlhs} * \text{lhs} = K_a$.

The \hat{u}_{ai} and the SEP (\hat{u}_a) for the animals in the Example are:

Anim	Prob _A	Br _A ^o - Br _B ^o	W _{ma1}	W _{ma2}	W _{ma3}	...	W _{ma60}	\widehat{ma}_1	\widehat{ma}_2	\widehat{ma}_3	...	\widehat{ma}_{60}	\hat{u}_a	SEP (\hat{u}_a)
1	1	21.09	2	1	1	...	2	-0.04	-0.03	0.03	...	-0.16	20.58	14.8
2	0	21.09	0	1	2	...	1	-0.04	-0.03	0.03	...	-0.16	-0.92	13.25
3	0.5	21.09	1	2	0	...	2	-0.04	-0.03	0.03	...	-0.16	5.16	13.8
4	0.5	21.09	1	2	0	...	1	-0.04	-0.03	0.03	...	-0.16	8.6	13.65
5	0.5	21.09	1	2	0	...	1	-0.04	-0.03	0.03	...	-0.16	13.18	13.6
6	0.75	21.09	2	1	2	...	1	-0.04	-0.03	0.03	...	-0.16	14.27	15.13
7	0.25	21.09	0.6	2	0	...	2	-0.04	-0.03	0.03	...	-0.16	-0.78	12.92
8	0.75	21.09	0.4	2	0	...	2	-0.04	-0.03	0.03	...	-0.16	14.78	13.75
9	0.625	21.09	0.7	2	0.4	...	2	-0.04	-0.03	0.03	...	-0.16	15.97	13.49
10	0.5	21.09	1.2	2	0.5	...	2	-0.04	-0.03	0.03	...	-0.16	9.58	14.81
11	0.375	21.09	2	0.5	2	...	0	-0.04	-0.03	0.03	...	-0.16	6.23	13.33
12	0.375	21.09	2	0.1	2	...	0	-0.04	-0.03	0.03	...	-0.16	6.22	13.58

Mauricio A. Elzo, University of Florida, 2010, 2011, 2012, 2014.

[GEV03 - 20]

Polygenic-Genomic Animal Model with phenotypes, actual genotypes, and predicted genotypes

$$\text{Wean wt} = \text{overall mean} + \text{breed regression} + \text{heterosis regression} + \text{sex} + \text{animal genomic} + \text{residual}$$

$$E[\text{Wean wt}] = \text{overall mean} + \text{breed regression} + \text{heterosis regression} + \text{sex}$$

$$\text{Var}(\text{Wean wt}) = \text{var}(\text{animal genomic}) + \text{var}(\text{residual})$$

$$\text{cov}(\text{Wean wt}_{ik}, \text{Wean wt}_{i'k'}) = \text{cov}(\text{animal genomic } j, \text{animal genomic } j') + \text{cov}(\text{residual}_{ik}, \text{residual}_{i'k'})$$

The polygenic-genomic animal model, in matrix notation, is:

$$y = Xb + Z_a Q_a g_a + Z_n Q_n g_n + Z_a Z_{ma} m_a + \varepsilon$$

Letting $a_a = Z_{ma} m_a$, the animal genomic model becomes:

$$y = Xb + Z_a Q_a g_a + Z_n Q_n g_n + Z_a a_a + \varepsilon$$

$$E[y] = Xb + Z_a Q_a g_a + Z_n Q_n g_n$$

$$\text{var} \begin{bmatrix} a_a \\ \varepsilon \end{bmatrix} = \begin{bmatrix} Z_{ma} G_{ma} Z_{ma}' & 0 \\ 0 & R \end{bmatrix}$$

$$\text{var}(y) = Z_a Q_a G_{ma} Z_{ma}' + R$$

where

y = vector of animal records,

b = vector of unknown fixed effects (mean, sex of calf),

g_a = vector of unknown additive group genetic effects (breed),

g_n = vector of unknown nonadditive group genetic effects (heterosis at 1 locus),

a_a = vector of unknown random animal additive genomic effects,

ε = vector of unknown random residual effects

- X = known incidence matrix relating records to fixed effects in vector b ,
- Z_a = known incidence matrix relating records to elements of vector a_a ,
- Z_n = known incidence matrix relating records to random polygenic nonadditive genetic effects in vector a_n (a_n is assumed to be part of the residual),
- Q_a = known incidence matrix relating random polygenic additive genetic effects to additive genetic groups in vector g_a ,
- Q_n = known incidence matrix relating random polygenic nonadditive genetic effects to nonadditive genetic groups in vector g_n ,
- Z_{ma} = known incidence matrix relating records to elements of vector ma ; elements of Z_{ma} will be 0, 1, 2 depending on the number of “second” alleles in marker locus j , $j = 1, \dots, 60$. Values [0,1, 2] are called gene content. Other parameterizations are: 1) [-1, 0, 1] obtained by subtracting 1 from [0, 1, 2]; and 2) deviate [-1, 0, 1] from $2(p_j - 0.5)$, which yields $[-2p_j, 1 - 2p_j, 2 - 2p_j]$, where 0.5 is the allele frequency used to center matrix Z_{ma} . Other frequencies used to center matrix Z_{ma} are the base population mean (computed using a linear regression on gene content; Gengler et al., 2007 and the current population mean (Aguilar et al., 2009).
- G_{ma} = $nma \times nma$ diagonal matrix of marker locus additive genomic variances (assumed to be equal for all markers associated with a trait),
- = $I * \sigma_a^2$
- = $I * \{\sigma_a^2 / kma\}$, where $kma = 2 \sum_{j=1}^{nma} p_j (1 - p_j)$, nma = number of marker loci (in linkage disequilibrium with QTL), p_j = frequency of the “second” allele in marker j , and σ_a^2 = polygenic additive genetic variance,

R = matrix of residual variances for purebred and crossbred animals (sums of nonadditive genetic and environmental variances).

Substituting $(I * (\sigma_a^2/kma))$ for G_{ma} in $var \begin{bmatrix} u_a \\ \varepsilon \end{bmatrix}$ yields:

$$var \begin{bmatrix} u_a \\ \varepsilon \end{bmatrix} = \begin{bmatrix} Z_{ma}(I * (\sigma_a^2/kma))Z_{ma}' & 0 \\ 0 & R \end{bmatrix}$$

$$var \begin{bmatrix} u_a \\ \varepsilon \end{bmatrix} = \begin{bmatrix} Z_{ma}Z_{ma}' \left(\frac{\sigma_a^2}{kma}\right) & 0 \\ 0 & R \end{bmatrix}$$

$$var \begin{bmatrix} u_a \\ \varepsilon \end{bmatrix} = \begin{bmatrix} \left(\frac{Z_{ma}Z_{ma}'}{2 \sum_{j=1}^{j=nma} p_j(1-p_j)}\right) * \sigma_a^2 & 0 \\ 0 & R \end{bmatrix}$$

$$var \begin{bmatrix} u_a \\ \varepsilon \end{bmatrix} = \begin{bmatrix} A_{gg} * \sigma_a^2 & 0 \\ 0 & R \end{bmatrix}$$

$$var \begin{bmatrix} u_a \\ \varepsilon \end{bmatrix} = \begin{bmatrix} G_{gg} & 0 \\ 0 & R \end{bmatrix}$$

where

$G_{gg} = \left(\frac{Z_{ma}Z_{ma}'}{2 \sum_{j=1}^{j=nma} p_j(1-p_j)}\right) * \sigma_a^2$ is the genomic additive covariance matrix,

$A_{gg} = Z_{ma}Z_{ma}'/2 \sum_{j=1}^{j=nma} p_j(1 - p_j)$ is the genomic relationship matrix, and

$\sigma_a^2 =$ polygenic additive genetic variance.

The mixed model with the information of the 12 calves can be written as follows:

$$\begin{bmatrix} 289 \\ 245 \\ 256 \\ 261 \\ 292 \\ 286 \\ 272 \\ 264 \\ 270 \\ 278 \\ 259 \\ 280 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 1/2 & 1/2 & 1 & 0 & 1 \\ 1 & 1/2 & 1/2 & 1 & 0 & 1 \\ 1 & 1/2 & 1/2 & 1 & 1 & 0 \\ 1 & 3/4 & 1/4 & 1/2 & 1 & 0 \\ 1 & 1/4 & 3/4 & 1/2 & 1 & 0 \\ 1 & 3/4 & 1/4 & 1/2 & 0 & 1 \\ 1 & 5/8 & 3/8 & 1/2 & 0 & 1 \\ 1 & 1/2 & 1/2 & 1/2 & 1 & 0 \\ 1 & 3/8 & 5/8 & 3/4 & 0 & 1 \\ 1 & 3/8 & 5/8 & 3/4 & 1 & 0 \end{bmatrix} \begin{bmatrix} mean \\ breed_A \\ breed_B \\ heterosis_{AB} \\ sex_M \\ sex_F \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_{a1} \\ a_{a2} \\ a_{a3} \\ a_{a4} \\ a_{a5} \\ a_{a6} \\ a_{a7} \\ a_{a8} \\ a_{a9} \\ a_{a10} \\ a_{a11} \\ a_{a12} \end{bmatrix} + \begin{bmatrix} residual_{anim 1} \\ residual_{anim 2} \\ residual_{anim 3} \\ residual_{anim 4} \\ residual_{anim 5} \\ residual_{anim 6} \\ residual_{anim 7} \\ residual_{anim 8} \\ residual_{anim 9} \\ residual_{anim 10} \\ residual_{anim 11} \\ residual_{anim 12} \end{bmatrix}$$

The mixed model equations (MME) are:

$$\begin{bmatrix} X'R^{-1}X & X'R^{-1}Z_aQ_a & X'R^{-1}Z_nQ_n & X'R^{-1}Z_a \\ Q'_aZ'_aR^{-1}X & Q'_aZ'_aR^{-1}Z_aQ_a & Q'_aZ'_aR^{-1}Z_nQ_n & Q'_aZ'_aR^{-1}Z_a \\ Q'_nZ'_nR^{-1}X & Q'_nZ'_nR^{-1}Z_aQ_a & Q'_nZ'_nR^{-1}Z_nQ_n & Q'_nZ'_nR^{-1}Z_a \\ Z'_aR^{-1}X & Z'_aR^{-1}Z_aQ_a & Z'_aR^{-1}Z_nQ_n & Z'_aR^{-1}Z_a + G_{gg}^{-1} \end{bmatrix} \begin{bmatrix} b \\ g_a \\ g_n \\ a_a \end{bmatrix} = \begin{bmatrix} X'R^{-1}y \\ Q'_aZ'_aR^{-1}y \\ Q'_nZ'_nR^{-1}y \\ Z'_aR^{-1}y \end{bmatrix}$$

Matrix $[X \ Z_aQ_a \ Z_nQ_n \ Z_a]$ is:

$xf = [X \ Z_aQ_a \ Z_nQ_n \ Z_a]$															
1	1	0	0	1	0	1	0	0	0	0	0	0	0	0	0
1	0	1	0	0	1	0	1	0	0	0	0	0	0	0	0
1	0.5	0.5	1	0	1	0	0	1	0	0	0	0	0	0	0
1	0.5	0.5	1	0	1	0	0	0	1	0	0	0	0	0	0
1	0.5	0.5	1	1	0	0	0	0	0	1	0	0	0	0	0
1	0.75	0.25	0.5	1	0	0	0	0	0	1	0	0	0	0	0
1	0.25	0.75	0.5	1	0	0	0	0	0	0	1	0	0	0	0
1	0.75	0.25	0.5	0	1	0	0	0	0	0	0	1	0	0	0
1	0.625	0.375	0.5	0	1	0	0	0	0	0	0	0	1	0	0
1	0.5	0.5	0.5	1	0	0	0	0	0	0	0	0	0	1	0
1	0.375	0.625	0.75	0	1	0	0	0	0	0	0	0	0	0	1
1	0.375	0.625	0.75	1	0	0	0	0	0	0	0	0	0	0	1

The inverse of the matrix of residual effects, R^{-1} , is:

$invr = R^{-1}$										
0.0204082	0	0	0	0	0	0	0	0	0	0
0	0.0625	0	0	0	0	0	0	0	0	0
0	0	0.0307692	0	0	0	0	0	0	0	0
0	0	0	0.0307692	0	0	0	0	0	0	0
0	0	0	0	0.0307692	0	0	0	0	0	0

invr = R ⁻¹											
0	0	0	0	0	0.0212766	0	0	0	0	0	0
0	0	0	0	0	0	0.0327869	0	0	0	0	0
0	0	0	0	0	0	0	0.0212766	0	0	0	0
0	0	0	0	0	0	0	0	0.021025	0	0	0
0	0	0	0	0	0	0	0	0	0.0222222	0	0
0	0	0	0	0	0	0	0	0	0	0.0302457	0
0	0	0	0	0	0	0	0	0	0	0	0.0302457

Matrix $G_{gg} = \left(\frac{Z_{ma}Z'_{ma}}{2\sum_{j=1}^{nma} p_j(1-p_j)} \right) * \sigma_a^2 = A_{gg} * \sigma_a^2$, where σ_a^2 was assumed to be equal to 36.

Matrix Z_{ma} is equal to:

$zg = Z_{ma}$											
2	1	1	2	1	1	2	2	...	0	2	2
0	1	2	0	0	1	2	2	...	1	2	1
1	2	0	0	1	1	1	2	...	1	1	2
1	2	0	0	0	2	2	0	...	1	2	1
1	2	0	1	1	1	1	1	...	2	2	1
2	1	2	2	1	2	2	1	...	1	2	1
0.584991	2	0	0	1.180194	0.961235	0.508471	2	...	1.448531	0.400564	2
0.424377	2	0	0	0	0.624769	2	0.773394	...	0	2	2
0.722955	2	0	0.399815	0.64053	0.242628	1.425643	1.463978	...	1.773901	2	2
1.185669	2	0	0.539372	2	0.114667	0.649897	2	...	2	1.931438	2
2	0.514714	2	2	1.646977	2	2	0	...	2	2	0
2	0.099822	2	2	1.588911	2	2	0.405518	...	2	2	0

Matrix $Z_{ma}Z'_{ma}$ is equal to:

$zgtzg = Z_{ma}Z'_{ma}$											
98	73	76	65	84	79	68.743163	72.370693	86.576595	98.583695	67.314599	68.261997
73	115	79	80	86	84	78.916343	73.738808	83.891166	95.319959	80.832056	85.120853
76	79	98	72	79	79	98.442288	64.047467	74.497754	100.29407	73.365519	73.353516
65	80	72	92	73	76	70.157557	77.163616	72.00266	78.80684	77.044925	76.980411
84	86	79	73	108	80	73.864936	68.218303	99.890424	113.01213	71.820043	71.828885
79	84	79	76	80	105	75.341501	49.287978	65.251567	85.448518	107.3146	107.85648
68.743163	78.916343	98.442288	70.157557	73.864936	75.341501	103.44974	63.084326	69.950518	96.461978	72.528838	72.542226
72.370693	73.738808	64.047467	77.163616	68.218303	49.287978	63.084326	101.88041	86.836956	83.608555	40.522903	40.916832
86.576595	83.891166	74.497754	72.00266	99.890424	65.251567	69.950518	86.836956	111.06138	112.64821	52.738989	52.958523
98.583695	95.319959	100.29407	78.80684	113.01213	85.448518	96.461978	83.608555	112.64821	137.1486	75.966448	75.787714
67.314599	80.832056	73.365519	77.044925	71.820043	107.3146	72.528838	40.522903	52.738989	75.966448	125.88171	126.42116
68.261997	85.120853	73.353516	76.980411	71.828885	107.85648	72.542226	40.916832	52.958523	75.787714	126.42116	129.03988

and $2 \sum_{j=1}^{j=nma} p_j(1 - p_j) = 20.974159$.

Thus, matrix $A_{gg} = \left(\frac{Z_{ma}Z'_{ma}}{2 \sum_{j=1}^{j=nma} p_j(1 - p_j)} \right)$ is equal to:

$agg = A_{gg} = \left(\frac{Z_{ma}Z'_{ma}}{2 \sum_{j=1}^{j=nma} p_j(1 - p_j)} \right)$											
4.6724161	3.4804733	3.6235064	3.0990515	4.0049281	3.7665395	3.277517	3.4504693	4.1277743	4.7002454	3.2094063	3.2545761
3.4804733	5.4829373	3.7665395	3.8142173	4.1002836	4.0049281	3.762551	3.5156979	3.9997392	4.5446379	3.8538878	4.0583678
3.6235064	3.7665395	4.6724161	3.4327955	3.7665395	3.7665395	4.6935034	3.0536369	3.5518827	4.781792	3.4979004	3.4973281
3.0990515	3.8142173	3.4327955	4.3863499	3.4804733	3.6235064	3.344952	3.678985	3.4329224	3.7573301	3.673326	3.6702502
4.0049281	4.1002836	3.7665395	3.4804733	5.1491933	3.8142173	3.5217114	3.2524929	4.7625472	5.3881602	3.4242156	3.4246372
3.7665395	4.0049281	3.7665395	3.6235064	3.8142173	5.0061602	3.5921107	2.3499382	3.1110457	4.0739902	5.116515	5.1423506
3.277517	3.762551	4.6935034	3.344952	3.5217114	3.5921107	4.9322473	3.0077166	3.3350809	4.5990868	3.4580093	3.4586476
3.4504693	3.5156979	3.0536369	3.678985	3.2524929	2.3499382	3.0077166	4.8574253	4.1401877	3.9862649	1.9320394	1.9508211

$\text{agg} = A_{gg} = \left(\frac{Z_{ma}Z'_{ma}}{2 \sum_{j=1}^{j=nma} p_j(1-p_j)} \right)$											
4.1277743	3.9997392	3.5518827	3.4329224	4.7625472	3.1110457	3.3350809	4.1401877	5.2951531	5.3708095	2.5144745	2.5249414
4.7002454	4.5446379	4.781792	3.7573301	5.3881602	4.0739902	4.5990868	3.9862649	5.3708095	6.5389317	3.6219067	3.6133851
3.2094063	3.8538878	3.4979004	3.673326	3.4242156	5.116515	3.4580093	1.9320394	2.5144745	3.6219067	6.0017525	6.0274723
3.2545761	4.0583678	3.4973281	3.6702502	3.4246372	5.1423506	3.4586476	1.9508211	2.5249414	3.6133851	6.0274723	6.1523269

Matrix $G_{gg}^{-1} = A_{gg}^{-1} * \sigma_a^{-2}$, where A_{gg}^{-1} is equal to:

$\text{ginvagg} = A_{gg}^{-1}$											
3.4021279	0.8447791	-3.272922	1.8141808	-0.106482	-2.911899	2.7164288	-2.096063	-0.415663	-0.337269	2.7001554	-2.224733
0.8447791	1.2519103	0.299945	0.1482126	-0.154155	-1.338481	-0.291879	-0.611813	-0.186395	-0.274437	2.9102629	-2.582508
-3.272922	0.299945	17.380473	-3.897236	1.1399623	-1.933243	-12.61894	2.7631857	1.2680707	-3.423036	2.954552	-0.227036
1.8141808	0.1482126	-3.897236	3.2982777	-1.249157	-0.957112	2.6211581	-2.560128	0.1086624	1.2162572	-0.70732	-0.042081
-0.106482	-0.154155	1.1399623	-1.249157	3.5112633	-0.813722	-0.350338	1.4758335	-1.87485	-1.69949	0.3030485	0.1805007
-2.911899	-1.338481	-1.933243	-0.957112	-0.813722	6.8455845	0.7408813	1.7002717	-0.054334	1.6606085	-6.12881	2.9201964
2.7164288	-0.291879	-12.61894	2.6211581	-0.350338	0.7408813	10.25007	-2.23475	-0.36277	1.192772	-1.822217	0.12085
-2.096063	-0.611813	2.7631857	-2.560128	1.4758335	1.7002717	-2.23475	3.1793202	-1.29383	-0.276505	-1.044329	1.1909758
-0.415663	-0.186395	1.2680707	0.1086624	-1.87485	-0.054334	-0.36277	-1.29383	3.84929	-1.373299	0.2913733	0.2017412
-0.337269	-0.274437	-3.423036	1.2162572	-1.69949	1.6606085	1.192772	-0.276505	-1.373299	3.8867951	-2.362772	1.1504978
2.7001554	2.9102629	2.954552	-0.70732	0.3030485	-6.12881	-1.822217	-1.044329	0.2913733	-2.362772	20.008044	-16.63004
-2.224733	-2.582508	-0.227036	-0.042081	0.1805007	2.9201964	0.12085	1.1909758	0.2017412	1.1504978	-16.63004	15.74431

Thus, matrix $G_{gg}^{-1} = A_{gg}^{-1} * (36)^{-1}$, is equal to:

$g_{inv} = G_{gg}^{-1} = A_{gg}^{-1} * (36)^{-1}$											
0.0945036	0.0234661	-0.090915	0.0503939	-0.002958	-0.080886	0.0754564	-0.058224	-0.011546	-0.009369	0.0750043	-0.061798
0.0234661	0.0347753	0.0083318	0.004117	-0.004282	-0.03718	-0.008108	-0.016995	-0.005178	-0.007623	0.0808406	-0.071736
-0.090915	0.0083318	0.4827909	-0.108257	0.0316656	-0.053701	-0.350526	0.0767552	0.0352242	-0.095084	0.0820709	-0.006307
0.0503939	0.004117	-0.108257	0.0916188	-0.034699	-0.026586	0.0728099	-0.071115	0.0030184	0.0337849	-0.019648	-0.001169
-0.002958	-0.004282	0.0316656	-0.034699	0.0975351	-0.022603	-0.009732	0.0409954	-0.052079	-0.047208	0.008418	0.0050139
-0.080886	-0.03718	-0.053701	-0.026586	-0.022603	0.1901551	0.02058	0.0472298	-0.001509	0.046128	-0.170245	0.0811166
0.0754564	-0.008108	-0.350526	0.0728099	-0.009732	0.02058	0.2847242	-0.062076	-0.010077	0.0331326	-0.050617	0.0033569
-0.058224	-0.016995	0.0767552	-0.071115	0.0409954	0.0472298	-0.062076	0.0883144	-0.03594	-0.007681	-0.029009	0.0330827
-0.011546	-0.005178	0.0352242	0.0030184	-0.052079	-0.001509	-0.010077	-0.03594	0.1069247	-0.038147	0.0080937	0.0056039
-0.009369	-0.007623	-0.095084	0.0337849	-0.047208	0.046128	0.0331326	-0.007681	-0.038147	0.1079665	-0.065633	0.0319583
0.0750043	0.0808406	0.0820709	-0.019648	0.008418	-0.170245	-0.050617	-0.029009	0.0080937	-0.065633	0.555779	-0.461946
-0.061798	-0.071736	-0.006307	-0.001169	0.0050139	0.0811166	0.0033569	0.0330827	0.0056039	0.0319583	-0.461946	0.4373419

The LHS matrix, the vector of effects and RHS vector of the mixed model equations are:

lhs																	Effect	rhs	
0.354	0.154	0.201	0.197	0.158	0.197	0.020	0.063	0.031	0.031	0.031	0.021	0.033	0.021	0.021	0.022	0.030	0.030	Mean	94.88
0.154	0.092	0.062	0.095	0.082	0.071	0.020	0.000	0.015	0.015	0.015	0.016	0.008	0.016	0.013	0.011	0.011	0.011	Breed _A	42.10
0.201	0.062	0.139	0.102	0.075	0.125	0.000	0.063	0.015	0.015	0.015	0.005	0.025	0.005	0.008	0.011	0.019	0.019	Breed _B	52.78
0.197	0.095	0.102	0.156	0.092	0.105	0.000	0.000	0.031	0.031	0.031	0.011	0.016	0.011	0.011	0.011	0.023	0.023	Heterosis _{AB}	53.36
0.158	0.082	0.075	0.092	0.158	0.000	0.020	0.000	0.000	0.000	0.031	0.021	0.033	0.000	0.000	0.022	0.000	0.030	Sex _M	44.53
0.197	0.071	0.125	0.105	0.000	0.197	0.000	0.063	0.031	0.031	0.000	0.000	0.000	0.021	0.021	0.000	0.030	0.000	Sex _F	50.35
0.020	0.020	0.000	0.000	0.020	0.000	0.115	0.023	-0.091	0.050	-0.003	-0.081	0.075	-0.058	-0.012	-0.009	0.075	-0.062	a _{a1}	5.90
0.063	0.000	0.063	0.000	0.000	0.063	0.023	0.097	0.008	0.004	-0.004	-0.037	-0.008	-0.017	-0.005	-0.008	0.081	-0.072	a _{a2}	15.31
0.031	0.015	0.015	0.031	0.000	0.031	-0.091	0.008	0.514	-0.108	0.032	-0.054	-0.351	0.077	0.035	-0.095	0.082	-0.006	a _{a3}	7.88
0.031	0.015	0.015	0.031	0.000	0.031	0.050	0.004	-0.108	0.122	-0.035	-0.027	0.073	-0.071	0.003	0.034	-0.020	-0.001	a _{a4}	8.03
0.031	0.015	0.015	0.031	0.031	0.000	-0.003	-0.004	0.032	-0.035	0.128	-0.023	-0.010	0.041	-0.052	-0.047	0.008	0.005	a _{a5}	8.98
0.021	0.016	0.005	0.011	0.021	0.000	-0.081	-0.037	-0.054	-0.027	-0.023	0.211	0.021	0.047	-0.002	0.046	-0.170	0.081	a _{a6}	6.09
0.033	0.008	0.025	0.016	0.033	0.000	0.075	-0.008	-0.351	0.073	-0.010	0.021	0.318	-0.062	-0.010	0.033	-0.051	0.003	a _{a7}	8.92
0.021	0.016	0.005	0.011	0.000	0.021	-0.058	-0.017	0.077	-0.071	0.041	0.047	-0.062	0.110	-0.036	-0.008	-0.029	0.033	a _{a8}	5.62
0.021	0.013	0.008	0.011	0.000	0.021	-0.012	-0.005	0.035	0.003	-0.052	-0.002	-0.010	-0.036	0.128	-0.038	0.008	0.006	a _{a9}	5.68
0.022	0.011	0.011	0.011	0.022	0.000	-0.009	-0.008	-0.095	0.034	-0.047	0.046	0.033	-0.008	-0.038	0.130	-0.066	0.032	a _{a10}	6.18
0.030	0.011	0.019	0.023	0.000	0.030	0.075	0.081	0.082	-0.020	0.008	-0.170	-0.051	-0.029	0.008	-0.066	0.586	-0.462	a _{a11}	7.83
0.030	0.011	0.019	0.023	0.030	0.000	-0.062	-0.072	-0.006	-0.001	0.005	0.081	0.003	0.033	0.006	0.032	-0.462	0.468	a _{a12}	8.47

The vector of **solutions** and their **standard errors** are:

Effect	SOL	SESOL
Mean ^o	133.91	6.10
Breed _A ^o	77.50	6.40
Breed _B ^o	56.41	6.54
Heterosis _{AB} ^o	7.40	6.98
Sex _M ^o	78.21	3.78
Sex _F ^o	55.70	3.75
â_{a1}	-0.52	12.22
â_{a2}	-0.92	13.25
â_{a3}	-5.39	11.71
â_{a4}	-1.95	11.92
â_{a5}	2.64	11.78
â_{a6}	-1.55	11.87
â_{a7}	-6.05	11.92
â_{a8}	-1.04	11.62
â_{a9}	2.79	11.60
â_{a10}	-0.97	13.19
â_{a11}	-1.68	11.86
â_{a12}	-1.69	12.01

Prediction of polygenic-genomic values

Predictions of additive polygenic-genomic values (\hat{u}_{ai}) are computed as the sum of the breed effect and the animal genomic deviation from the breed effect, i.e.,

$$\hat{u}_{ai} = \text{Prob}_{A_{anim\ i}} * (\text{Breed}_A^\circ - \text{Breed}_B^\circ) + \hat{a}_{ai}$$

where

- $\text{Prob}_{A_{anim\ i}}$ = expected fraction of breed A in animal i,
- Breed_A° = generalized least squares solution for breed A,
- Breed_B° = generalized least squares solution for breed B,
- \hat{a}_{ai} = Genomic BLUP (GBLUP) of a_{ai} .

The matrix of variances of errors of prediction (VEP) for the vector of predicted genetic values (\hat{u}_a) is:

$$\text{VEP}(\hat{u}_a) = \text{var}(\hat{u}_a - u_a) = \mathbf{K}_a * (\text{ginvlhs}) * \mathbf{K}_a'$$

where

- \mathbf{K}_a = $n_{anim} \times n_{eq}$ matrix specifying multipliers for breed differences and and **random animal additive genomic deviations**, where n_{anim} = number of animals, and n_{eq} = number of equations, and
- ginvlhs** = $n_{eq} \times n_{eq}$ generalized inverse matrix of the left hand side of the MME.

The standard errors of prediction (SEP) of the \hat{u}_{ai} are computed as the square roots of the diagonal elements of the $\text{VEP}(\hat{u}_{ai})$ matrix.

Matrix \mathbf{K}_a for the Example is:

ka																	
0.00	1.00	-1.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.50	-0.50	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.50	-0.50	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.50	-0.50	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.75	-0.75	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.25	-0.25	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00
0.00	0.75	-0.75	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00
0.00	0.63	-0.63	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00
0.00	0.50	-0.50	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00
0.00	0.38	-0.38	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00
0.00	0.38	-0.38	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00

The \hat{u}_{ai} and the SEP (\hat{u}_a) for the animals in the Example are:

Anim	Prob _A	Br _A ^o - Br _B ^o	\hat{u}_a	SEP (\hat{u}_a)
1	1	21.09	20.58	14.80
2	0	21.09	-0.92	13.25
3	0.5	21.09	5.16	13.80
4	0.5	21.09	8.60	13.65
5	0.5	21.09	13.18	13.60
6	0.75	21.09	14.27	15.13

7	0.25	21.09	-0.78	12.92
8	0.75	21.09	14.78	13.75
9	0.625	21.09	15.97	13.49
10	0.5	21.09	9.58	14.81
11	0.375	21.09	6.23	13.33
12	0.375	21.09	6.22	13.58

Computation of marker BLUP from animal genomic BLUP

Marker BLUP can be predicted from the GBLUP of a_a by conditioning m_a on \hat{a}_a and assuming normality. Firstly, under normality,

$$m_a | a_a = Z_{ma} m_a \sim MVN \{E[m_a | a_a = Z_{ma} m_a], \text{var}(m_a | a_a = Z_{ma} m_a)\}$$

where,

$$E[m_a | a_a = Z_{ma} m_a] = E[m_a] + \text{cov}(m_a, m_a' Z_{ma}') \text{var}(Z_{ma} m_a)^{-1} (Z_{ma} m_a - E[Z_{ma} m_a])$$

$$\text{var}(m_a | a_a = Z_{ma} m_a) = \text{var}(m_a) - \text{cov}(m_a, m_a' Z_{ma}') \text{var}(Z_{ma} m_a)^{-1} \text{cov}(Z_{ma} m_a, m_a')$$

$$E[m_a] = 0$$

$$E[a_a = Z_{ma} m_a] = 0$$

$$\text{var}(m_a) = (I * (\sigma_a^2 / kma))$$

$$\text{cov}(m_a, m_a' Z_{ma}') = (I * (\sigma_a^2 / kma)) Z_{ma}'$$

$$\text{var}(Z_{ma} m_a)^{-1} = (Z_{ma} (I * (\sigma_a^2 / kma)) Z_{ma}')^{-1}$$

$$\text{cov}(Z_{ma} m_a, m_a') = Z_{ma} (I * (\sigma_a^2 / kma))$$

and,

$$kma = 2 \sum_{j=1}^{j=nma} p_j(1 - p_j)$$

Thus,

$$E[m_a|a_a] = (I * (\sigma_a^2/kma))Z_{ma}'(Z_{ma}(I * (\sigma_a^2/kma))Z_{ma}')^{-1}a_a$$

and,

$$\begin{aligned} \text{var}(m_a|a_a) &= (I * (\sigma_a^2/kma)) \\ &\quad - (I * (\sigma_a^2/kma))Z_{ma}'(Z_{ma}(I * (\sigma_a^2/kma))Z_{ma}')^{-1}Z_{ma}(I * (\sigma_a^2/kma)) \end{aligned}$$

But $E[a_a|\hat{a}_a] = \hat{a}_a$, the MLE of a_a and the BLUP of a_a under normality. By the invariance property of MLE, the MLE of m_a (and the BLUP of m_a under normality) can be predicted as

$\hat{m}_a = K\hat{a}_a$, where

$$\begin{aligned} K &= \text{cov}(m_a, m_a'Z_{ma}')\text{var}(Z_{ma}m_a)^{-1} \\ K &= (I * (\sigma_a^2/kma))Z_{ma}'(Z_{ma}(I * (\sigma_a^2/kma))Z_{ma}')^{-1} \end{aligned}$$

Thus,

$$\hat{m}_a = (I * (\sigma_a^2/kma))Z_{ma}'(Z_{ma}(I * (\sigma_a^2/kma))Z_{ma}')^{-1}\hat{a}_a$$

In other words, by the invariance property of MLE,

$$\hat{m}_a = E[m_a|\hat{a}_a] = (I * (\sigma_a^2/kma))Z_{ma}'(Z_{ma}(I * (\sigma_a^2/kma))Z_{ma}')^{-1}\hat{a}_a$$

where $\hat{a}_a = \text{BLUP of } a_a = Z_{ma}m_a$ obtained by solving the MME for the animal genomic model.

The error variance of prediction of \hat{m}_a , i.e., $\text{var}(\hat{m}_a - m_a)$, can be computed as follows:

$$\text{var}(m_a) = E[\text{var}(m_a|\hat{a}_a)] + \text{var}(E[m_a|\hat{a}_a])$$

or,

$$\text{var}(m_a) = \text{var}(\hat{m}_a - m_a) + \text{var}(\hat{m}_a)$$

Thus,

$$\text{var}(\hat{m}_a - m_a) = \text{var}(m_a) - \text{var}(\hat{m}_a)$$

where

$$\text{var}(m_a) = (I * (\sigma_a^2/kma))$$

$$\text{var}(\hat{m}_a) = \text{cov}(m_a, m_a' Z_{ma}') \text{var}(Z_{ma} m_a)^{-1} \text{var}(\hat{a}_a) \text{var}(Z_{ma} m_a)^{-1} \text{cov}(Z_{ma} m_a, m_a')$$

$$\begin{aligned} \text{var}(\hat{m}_a) = & (I \\ & * (\sigma_a^2/kma)) Z_{ma}' (Z_{ma} (I \\ & * (\sigma_a^2/kma)) Z_{ma}')^{-1} \text{var}(\hat{a}_a) (Z_{ma} (I * (\sigma_a^2/kma)) Z_{ma}')^{-1} Z_{ma} (I \\ & * (\sigma_a^2/kma)) \end{aligned}$$

where,

$$\text{var}(\hat{a}_a) = \text{var}(a_a) - B^{22}$$

$$\text{var}(\hat{a}_a) = \text{var}(Z_{ma} m_a) - B^{22}$$

$$\text{var}(\hat{a}_a) = Z_{ma} (I * (\sigma_a^2/kma)) Z_{ma}' - B^{22}$$

and,

B^{22} = submatrix of the inverse of the LHS of the MME for the animal genomic model of error corresponding to animal genomic effects a_a .

The standard errors of prediction (**SEP**) of the marker BLUP are computed as the square root of the diagonal elements of $\text{var}(\hat{m}_a)$, the matrix of error variances of prediction.

The MLE of m_a for the Example is $\hat{m}_a = K \hat{a}_a$, where $K = \text{cov}(m_a, m_a' Z_{ma}') \text{var}(Z_{ma} m_a)^{-1}$.

The $n_{ma} \times n_{anim}$ $\text{cov}(m_a, m_a' Z_{ma}') = (I * (\sigma_a^2/kma)) Z_{ma}'$ is equal to:

covtzgg = cov(m _a , m _a 'Z _{ma})											
3.433	0.000	1.716	1.716	1.716	3.433	1.004	0.728	1.241	2.035	3.433	3.433
1.716	1.716	3.433	3.433	3.433	1.716	3.433	3.433	3.433	3.433	0.883	0.171
1.716	3.433	0.000	0.000	0.000	3.433	0	0	0	0	3.433	3.433
3.433	0.000	0.000	0.000	1.716	3.433	0	0	0.686	0.926	3.433	3.433
1.716	0.000	1.716	0.000	1.716	1.716	2.026	0	1.099	3.433	2.827	2.727
1.716	1.716	1.716	3.433	1.716	3.433	1.65	1.072	0.416	0.197	3.433	3.433
3.433	3.433	1.716	3.433	1.716	3.433	0.873	3.433	2.447	1.115	3.433	3.433
3.433	3.433	3.433	0.000	1.716	1.716	3.433	1.327	2.513	3.433	0	0.696
...
0.000	1.716	1.716	1.716	3.433	1.716	2.486	0	3.045	3.433	3.433	3.433
3.433	3.433	1.716	3.433	3.433	3.433	0.688	3.433	3.433	3.315	3.433	3.433
3.433	1.716	3.433	1.716	1.716	1.716	3.433	3.433	3.433	3.433	0	0

The $\text{nanim} \times \text{nanim } \text{var}(Z_{ma}m_a)^{-1} = (Z_{ma}(I * (\sigma_a^2/kma))Z_{ma}')^{-1} = G_{gg}^{-1}$ is equal to:

ginv _{ggg} = var(Z _{ma} m _a) ⁻¹ = G _{gg} ⁻¹											
0.0945036	0.0234661	-0.090915	0.0503939	-0.002958	-0.080886	0.0754564	-0.058224	-0.011546	-0.009369	0.0750043	-0.061798
0.0234661	0.0347753	0.0083318	0.004117	-0.004282	-0.03718	-0.008108	-0.016995	-0.005178	-0.007623	0.0808406	-0.071736
-0.090915	0.0083318	0.4827909	-0.108257	0.0316656	-0.053701	-0.350526	0.0767552	0.0352242	-0.095084	0.0820709	-0.006307
0.0503939	0.004117	-0.108257	0.0916188	-0.034699	-0.026586	0.0728099	-0.071115	0.0030184	0.0337849	-0.019648	-0.001169
-0.002958	-0.004282	0.0316656	-0.034699	0.0975351	-0.022603	-0.009732	0.0409954	-0.052079	-0.047208	0.008418	0.0050139
-0.080886	-0.03718	-0.053701	-0.026586	-0.022603	0.1901551	0.02058	0.0472298	-0.001509	0.046128	-0.170245	0.0811166
0.0754564	-0.008108	-0.350526	0.0728099	-0.009732	0.02058	0.2847242	-0.062076	-0.010077	0.0331326	-0.050617	0.0033569
-0.058224	-0.016995	0.0767552	-0.071115	0.0409954	0.0472298	-0.062076	0.0883144	-0.03594	-0.007681	-0.029009	0.0330827
-0.011546	-0.005178	0.0352242	0.0030184	-0.052079	-0.001509	-0.010077	-0.03594	0.1069247	-0.038147	0.0080937	0.0056039
-0.009369	-0.007623	-0.095084	0.0337849	-0.047208	0.046128	0.0331326	-0.007681	-0.038147	0.1079665	-0.065633	0.0319583
0.0750043	0.0808406	0.0820709	-0.019648	0.008418	-0.170245	-0.050617	-0.029009	0.0080937	-0.065633	0.555779	-0.461946
-0.061798	-0.071736	-0.006307	-0.001169	0.0050139	0.0811166	0.0033569	0.0330827	0.0056039	0.0319583	-0.461946	0.4373419

Thus, the matrix $K = cov(m_a, m_a'Z_{ma}')var(Z_{ma}m_a)^{-1}$ is equal to:

covgzg = K = cov(m _a , m _a 'Z _{ma} ')var(Z _{ma} m _a) ⁻¹											
0.030	-0.036	-0.010	0.007	-0.008	0.034	-0.03	-0.002	-0.003	0.024	-0.079	0.077
-0.023	-0.026	0.027	0.030	0.030	-0.017	-0.037	0.032	0.007	-0.048	0.159	-0.178
0.022	0.049	-0.034	-0.039	-0.040	0.048	0.01	0.018	0.004	0.001	0.144	-0.158
0.038	-0.020	-0.040	-0.021	0.002	0.051	0.174	0.015	-0.049	0.003	-0.046	0.024
-0.000	-0.020	0.016	-0.022	0.015	0.018	0.092	0	-0.047	0.116	-0.016	0.04
-0.007	-0.018	-0.013	0.033	-0.005	0.028	0.061	-0.019	-0.024	-0.063	-0.031	0.008
0.037	0.016	-0.028	0.024	-0.034	0.008	-0.181	0.028	0.053	-0.114	0.163	-0.105
0.036	0.040	0.034	-0.049	-0.029	-0.017	0.032	-0.033	0.031	0.014	-0.2	0.151
...
-0.055	-0.013	0.008	-0.001	0.057	0.012	0.14	-0.074	0.092	0.02	-0.022	0.071
0.019	0.005	-0.033	0.019	0.005	0.008	-0.188	0.058	-0.013	-0.001	0.051	-0.021
0.034	-0.002	0.030	-0.002	-0.023	-0.019	0.029	0.004	0.064	0.012	-0.104	0.043

The $n_{anim} \times 1$ vector \hat{a}_a is equal to:

gebv = \hat{a}_a
-0.52
-0.92
-5.39
-1.95
2.64
-1.55
-6.05
-1.04
2.79
-0.97
-1.68

$gebv = \hat{a}_a$
-1.69

The variance of prediction errors of \hat{m}_a is $var(\hat{m}_a - m_a) = var(m_a) - var(\hat{m}_a)$, where

$$var(m_a) = (I * (\sigma_a^2/kma)) = G_{ma}, \quad var(\hat{m}_a) = Kvar(\hat{a}_a)K', \quad var(\hat{a}_a) = var(Z_{ma}m_a) -$$

$$B^{22}, \quad \text{and } var(Z_{ma}m_a) = Z_{ma}(I * (\sigma_a^2/kma))Z'_{ma} = G_{gg}. \quad \text{Thus,}$$

$$var(\hat{m}_a - m_a) = G_{ma} - K(G_{gg} - B^{22})K'$$

Matrix G_{ma} is equal to:

$covg = var(m_a) = (I * (\sigma_a^2/kma)) = G_{ma}$									
1.7163978	0	0	0	0	0	...	0	0	0
0	1.7163978	0	0	0	0	...	0	0	0
0	0	1.7163978	0	0	0	...	0	0	0
0	0	0	1.7163978	0	0	...	0	0	0
0	0	0	0	1.7163978	0	...	0	0	0
0	0	0	0	0	1.7163978	...	0	0	0
...
0	0	0	0	0	0	...	1.7163978	0	0
0	0	0	0	0	0	...	0	1.7163978	0
0	0	0	0	0	0	...	0	0	1.7163978

Matrix G_{gg} is equal to:

$covagg = var(Z_{ma}m_a) = Z_{ma}(I * (\sigma_a^2/kma))Z'_{ma} = G_{gg}$											
168.21	125.30	130.45	111.57	144.18	135.60	117.99	124.22	148.60	169.21	115.54	117.16
125.30	197.39	135.60	137.31	147.61	144.18	135.45	126.57	143.99	163.61	138.74	146.10

$covagg = var(Z_{ma}m_a) = Z_{ma}(I * (\sigma_a^2/kma))Z'_{ma} = G_{gg}$											
130.45	135.60	168.21	123.58	135.60	135.60	168.97	109.93	127.87	172.14	125.92	125.90
111.57	137.31	123.58	157.91	125.30	130.45	120.42	132.44	123.59	135.26	132.24	132.13
144.18	147.61	135.60	125.30	185.37	137.31	126.78	117.09	171.45	193.97	123.27	123.29
135.60	144.18	135.60	130.45	137.31	180.22	129.32	84.598	112.00	146.66	184.19	185.12
117.99	135.45	168.97	120.42	126.78	129.32	177.56	108.28	120.06	165.57	124.49	124.51
124.22	126.57	109.93	132.44	117.09	84.598	108.28	174.87	149.05	143.51	69.553	70.230
148.60	143.99	127.87	123.59	171.45	112.00	120.06	149.05	190.63	193.35	90.521	90.898
169.21	163.61	172.14	135.26	193.97	146.66	165.57	143.51	193.35	235.40	130.39	130.08
115.54	138.74	125.92	132.24	123.27	184.19	124.49	69.553	90.521	130.39	216.06	216.99
117.16	146.10	125.90	132.13	123.29	185.12	124.51	70.230	90.898	130.08	216.99	221.48

Matrix B^{22} (quadrant 22 of the inverse of the LHS) is equal to:

$gevp = B^{22}$											
149.23	113.40	117.70	111.25	117.75	127.30	109.11	119.95	124.01	137.80	113.32	114.97
113.40	175.64	131.79	128.83	126.79	131.91	136.01	120.40	126.16	145.56	131.59	137.70
117.70	131.79	137.03	124.91	124.93	128.99	136.57	110.19	118.85	144.77	126.79	128.05
111.25	128.83	124.91	142.00	122.73	125.74	123.20	121.42	119.87	134.65	126.41	126.37
117.75	126.79	124.93	122.73	138.69	127.94	123.00	106.70	126.01	146.72	125.68	126.12
127.30	131.91	128.99	125.74	127.94	140.89	126.00	110.91	121.74	142.14	133.47	133.89
109.11	136.01	136.57	123.20	123.00	126.00	142.17	108.82	116.51	142.76	127.74	129.34
119.95	120.40	110.19	121.42	106.70	110.91	108.82	135.00	119.93	126.33	107.54	108.60
124.01	126.16	118.85	119.87	126.01	121.74	116.51	119.93	134.45	142.76	117.38	118.38
137.80	145.56	144.77	134.65	146.72	142.14	142.76	126.33	142.76	174.04	140.81	141.77
113.32	131.59	126.79	126.41	125.68	133.47	127.74	107.54	117.38	140.81	140.77	140.97
114.97	137.70	128.05	126.37	126.12	133.89	129.34	108.60	118.38	141.77	140.97	144.34

The error variance of prediction matrix for \hat{m}_a , $var(\hat{m}_a - m_a) = G_{ma} - K(G_{gg} - B^{22})K'$, is equal to:

$evpm = var(\hat{m}_a - m_a) = G_{ma} - K(G_{gg} - B^{22})K'$									
1.66	0.04	-0.04	-0.06	-0.05	-0.02	...	-0.03	-0.01	0.03
0.04	1.64	0.07	0.07	0.02	0.02	...	0.01	0.03	-0.06
-0.04	0.07	1.6	-0.08	-0.01	-0.05	...	0	-0.04	0.07
-0.06	0.07	-0.08	1.62	-0.05	-0.03	...	-0.04	-0.03	0.07
-0.05	0.02	-0.01	-0.05	1.61	0.04	...	-0.07	0	0
-0.02	0.02	-0.05	-0.03	0.04	1.64	...	0.03	-0.01	0.04
...
-0.03	0.01	0	-0.04	-0.07	0.03	...	1.62	-0.02	0.01
-0.01	0.03	-0.04	-0.03	0	-0.01	...	-0.02	1.65	0.05
0.03	-0.06	0.07	0.07	0	0.04	...	0.01	0.05	1.64

The vector of **marker BLUP** and their **standard errors of prediction** computed using the above expressions are:

Marker	SOL	SESOL	Marker	SOL	SESOL	Marker	SOL	SESOL
\hat{m}_1	-0.04	1.29	\hat{m}_{21}	0.00	1.31	\hat{m}_{41}	0.30	1.24
\hat{m}_2	-0.03	1.28	\hat{m}_{22}	-0.02	1.27	\hat{m}_{42}	0.18	1.28
\hat{m}_3	0.03	1.26	\hat{m}_{23}	-0.32	1.27	\hat{m}_{43}	0.05	1.28
\hat{m}_4	0.19	1.27	\hat{m}_{24}	0.14	1.25	\hat{m}_{44}	0.07	1.29
\hat{m}_5	-0.08	1.27	\hat{m}_{25}	-0.14	1.28	\hat{m}_{45}	0.01	1.28
\hat{m}_6	-0.01	1.28	\hat{m}_{26}	-0.23	1.27	\hat{m}_{46}	-0.44	1.25
\hat{m}_7	0.04	1.29	\hat{m}_{27}	0.20	1.28	\hat{m}_{47}	0.19	1.27
\hat{m}_8	-0.19	1.26	\hat{m}_{28}	-0.25	1.27	\hat{m}_{48}	-0.05	1.29

$\widehat{m}a_9$	-0.28	1.25	$\widehat{m}a_{29}$	-0.17	1.26	$\widehat{m}a_{49}$	-0.20	1.26
$\widehat{m}a_{10}$	-0.05	1.30	$\widehat{m}a_{30}$	-0.23	1.27	$\widehat{m}a_{50}$	0.07	1.28
$\widehat{m}a_{11}$	0.14	1.29	$\widehat{m}a_{31}$	0.08	1.28	$\widehat{m}a_{51}$	-0.08	1.25
$\widehat{m}a_{12}$	0.34	1.26	$\widehat{m}a_{32}$	-0.09	1.29	$\widehat{m}a_{52}$	-0.07	1.29
$\widehat{m}a_{13}$	-0.04	1.25	$\widehat{m}a_{33}$	-0.07	1.28	$\widehat{m}a_{53}$	0.06	1.27
$\widehat{m}a_{14}$	0.17	1.28	$\widehat{m}a_{34}$	-0.02	1.27	$\widehat{m}a_{54}$	-0.04	1.29
$\widehat{m}a_{15}$	-0.31	1.27	$\widehat{m}a_{35}$	0.02	1.27	$\widehat{m}a_{55}$	0.00	1.31
$\widehat{m}a_{16}$	0.03	1.27	$\widehat{m}a_{36}$	0.22	1.25	$\widehat{m}a_{56}$	0.25	1.28
$\widehat{m}a_{17}$	0.00	1.31	$\widehat{m}a_{37}$	-0.08	1.29	$\widehat{m}a_{57}$	0.00	1.31
$\widehat{m}a_{18}$	-0.15	1.27	$\widehat{m}a_{38}$	0.27	1.25	$\widehat{m}a_{58}$	0.21	1.27
$\widehat{m}a_{19}$	0.02	1.26	$\widehat{m}a_{39}$	0.01	1.25	$\widehat{m}a_{59}$	0.10	1.28
$\widehat{m}a_{20}$	0.05	1.25	$\widehat{m}a_{40}$	-0.24	1.29	$\widehat{m}a_{60}$	-0.16	1.28

Polygenic-Genomic Animal Model with phenotypic, pedigree, and genotypic information**Theoretical Development**

Consider the following **genetic model**:

$$y = Xb + Z_a Q_a g_a + Z_n Q_n g_n + Z_a a_a + \varepsilon$$

$$E[y] = Xb + Z_a Q_a g_a + Z_n Q_n g_n$$

$$\text{var} \begin{bmatrix} a_a \\ \varepsilon \end{bmatrix} = \begin{bmatrix} G_a & 0 \\ 0 & R \end{bmatrix}$$

$$\text{var}(y) = Z_a G_a Z_a' + R$$

where

y = vector of phenotypic animal records,

b = vector of unknown fixed effects (mean, sex of calf),

g_a = vector of unknown additive group genetic effects (breed),

g_n = vector of unknown nonadditive group genetic effects (heterosis at 1 locus),

a_a = vector of unknown random additive genetic effects,

ε = vector of unknown random residual effects

X = known incidence matrix relating records to fixed effects in vector b ,

Z_a = known incidence matrix relating records to elements of vector a_a ,

Z_n = known incidence matrix relating records to random nonadditive genetic effects in vector a_n (a_n is assumed to be part of the residual),

Q_a = known incidence matrix relating random additive genetic effects to additive genetic groups in vector g_a ,

Q_n = known incidence matrix relating random nonadditive genetic effects to nonadditive genetic groups in vector g_n ,

G_a = matrix of additive genetic variances and covariances among animals (purebred and crossbred) in the multibreed population; covariances are due to additive relationships among animals.

R = matrix of residual variances for purebred and crossbred animals (sums of nonadditive genetic and environmental variances).

Polygenic-Genomic Model

Assume that only a fraction of animals in the population are genotyped. Thus, a polygenic-genomic model that uses phenotypic, pedigree, and genotypic information can be written as follows:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} [b] + \begin{bmatrix} Z_{a1} Q_{a1} \\ Z_{a2} Q_{a2} \end{bmatrix} [g_a] + \begin{bmatrix} Z_{n1} Q_{n1} \\ Z_{n2} Q_{n2} \end{bmatrix} [g_n] + \begin{bmatrix} Z_{a1} & 0 \\ 0 & Z_{a2} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix}$$

$$E \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} [b] + \begin{bmatrix} Z_{a1} Q_{a1} \\ Z_{a2} Q_{a2} \end{bmatrix} [g_a] + \begin{bmatrix} Z_{n1} Q_{n1} \\ Z_{n2} Q_{n2} \end{bmatrix} [g_n]$$

$$\text{var} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \text{var} \left(\begin{bmatrix} Z_{a1} & 0 \\ 0 & Z_{a2} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix} \right)$$

$$\text{var} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} Z_{a1} & 0 \\ 0 & Z_{a2} \end{bmatrix} \text{var} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \begin{bmatrix} Z_{a1}' & 0 \\ 0 & Z_{a2}' \end{bmatrix} + \begin{bmatrix} R_1 & 0 \\ 0 & R_2 \end{bmatrix}$$

$$\text{var} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} Z_{a1} & 0 \\ 0 & Z_{a2} \end{bmatrix} \begin{bmatrix} G_{ag11} & G_{ag12} \\ G_{ag21} & G_{ag22} \end{bmatrix} \begin{bmatrix} Z_{a1}' & 0 \\ 0 & Z_{a2}' \end{bmatrix} + \begin{bmatrix} R_1 & 0 \\ 0 & R_2 \end{bmatrix}$$

where subscript 1 = non-genotyped animals and subscript 2 = genotyped animals, and

G_{ag11} = additive polygenic-genomic variance matrix for non-genotyped animals,

G_{ag12} = additive polygenic-genomic covariance matrix between non-genotyped and genotyped animals,

$$G_{ag21} = G_{ag12}'$$

$$G_{ag22} = \text{additive polygenic-genomic variance for genotyped animals.}$$

An explicit expression for the polygenic-genomic covariance matrix $\begin{bmatrix} G_{ag11} & G_{ag12} \\ G_{ag21} & G_{ag22} \end{bmatrix}$ can be

derived by conditioning the polygenic-genomic values of the non-genotyped animals (a_1) on the vector of genetic values of genotyped animals (a_2). Thus,

1) The polygenic-genomic variance for non-genotyped animals a_1 is:

$$G_{ag11} = E[\text{var}(a_1|a_2)] + \text{var}(E[a_1|a_2])$$

The conditional distribution of $a_1|a_2$ is:

$$p(a_1|a_2) = \text{MVN}(G_{12}G_{22}^{-1}a_2, G_{11} - G_{12}G_{22}^{-1}G_{21})$$

where

$$E[a_1|a_2] = G_{12}G_{22}^{-1}a_2$$

and

$$\text{var}(a_1|a_2) = G_{11} - G_{12}G_{22}^{-1}G_{21}$$

Thus,

$$G_{ag11} = E[G_{11} - G_{12}G_{22}^{-1}G_{21}] + \text{var}(G_{12}G_{22}^{-1}a_2)$$

But, in genomic terms, $a_2 = Z_{ma2} ma_2$, where ma_2 = vector of marker locus additive genomic effects, and Z_{ma2} = known incidence matrix relating records to elements of vector ma . Elements of Z_{ma2} will be 0, 1, 2 depending on the number of “second” alleles in marker locus j , $j = 1, \dots, 60$. Values [0, 1, 2] are called gene content. Other parameterizations are: 1) [-1, 0, 1] obtained by subtracting 1 from [0, 1, 2]; and 2) deviate [-1, 0, 1] from $2(p_j - 0.5)$, which yields $[-2p_j, 1 - 2p_j, 2 - 2p_j]$, where 0.5 is the allele frequency used to center matrix Z_{ma} . Other frequencies used to center matrix Z_{ma} are

the base population mean computed using a linear regression on gene content (Gengler et al., 2007) and the current population mean (Aguilar et al., 2009).

Thus, substituting $Z_{ma} a_2$ for a_2 yields:

$$G_{ag11} = E[G_{11} - G_{12}G_{22}^{-1}G_{21}] + var(G_{12}G_{22}^{-1}Z_{ma}a_2)$$

$$G_{ag11} = E[G_{11} - G_{12}G_{22}^{-1}G_{21}] + cov(G_{12}G_{22}^{-1}Z_{ma}a_2, a_2'Z_{ma}'G_{22}^{-1}G_{21})$$

$$G_{ag11} = G_{11} - G_{12}G_{22}^{-1}G_{21} + G_{12}G_{22}^{-1}Z_{ma}(I * (\sigma_a^2/kma))Z_{ma}'G_{22}^{-1}G_{21}$$

$$G_{ag11} = G_{11} - G_{12}G_{22}^{-1}G_{21} + G_{12}G_{22}^{-1}Z_{ma}Z_{ma}'(\sigma_a^2/kma)G_{22}^{-1}G_{21}$$

where σ_a^2 = polygenic additive genetic variance, $kma = 2 \sum_{j=1}^{j=nma} p_j(1 - p_j)$, nma = number of marker loci, and p_j = frequency of “second” allele in marker locus j .

Let the genomic variance for genotyped animals a_2 be:

$$var(a_2) = Z_{ma}Z_{ma}' \left(\frac{\sigma_a^2}{kma} \right)$$

$$var(a_2) = \left(\frac{Z_{ma}Z_{ma}'}{2 \sum_{j=1}^{j=nma} p_j(1 - p_j)} \right) * \sigma_a^2$$

$$var(a_2) = A_{gg} * \sigma_a^2$$

$$var(a_2) = G_{gg}$$

where

$A_{gg} = Z_{ma}Z_{ma}' / 2 \sum_{j=1}^{j=nma} p_j(1 - p_j)$ is the genomic relationship matrix, and

$\sigma_a^2 =$ polygenic additive genetic variance.

Thus,

$$G_{ag11} = G_{11} - G_{12}G_{22}^{-1}G_{21} + G_{12}G_{22}^{-1}G_{gg}G_{22}^{-1}G_{21}$$

$$G_{ag11} = G_{11} + G_{12}G_{22}^{-1}G_{gg}G_{22}^{-1}G_{21} - G_{12}G_{22}^{-1}G_{22}G_{22}^{-1}G_{21}$$

$$G_{ag11} = G_{11} + G_{12}G_{22}^{-1}(G_{gg} - G_{22})G_{22}^{-1}G_{21}$$

2) The polygenic-genomic covariance between non-genotyped a_1 and genotyped animals a_2 is:

$$G_{ag12} = E[\text{cov}(a_1|a_2, a_2|a_2)] + \text{cov}(E[a_1|a_2], E[a_2|a_2])$$

$$G_{ag12} = E[\text{cov}(a_1, a_2|a_2)] + \text{cov}(E[a_1|a_2], E[a_2|a_2])$$

$$G_{ag12} = 0 + \text{cov}(G_{12}G_{22}^{-1}Z_{ma}ma_2, ma_2'Z_{ma}')$$

$$G_{ag12} = G_{12}G_{22}^{-1}Z_{ma}(I * (\sigma_a^2/kma))Z_{ma}'$$

$$G_{ag12} = G_{12}G_{22}^{-1}Z_{ma}Z_{ma}'(\sigma_a^2/kma)$$

$$G_{ag12} = G_{12}G_{22}^{-1}G_{gg}$$

Thus, the polygenic-genomic covariance matrix for non-genotyped and genotyped animals is:

$$\begin{bmatrix} G_{ag11} & G_{ag12} \\ G_{ag21} & G_{ag22} \end{bmatrix} = \begin{bmatrix} G_{11} + G_{12}G_{22}^{-1}(G_{gg} - G_{22})G_{22}^{-1}G_{21} & G_{12}G_{22}^{-1}G_{gg} \\ G_{gg}G_{22}^{-1}G_{21} & G_{gg} \end{bmatrix}$$

Adding and subtracting terms yields (Legarra et al., 2009):

$$\begin{bmatrix} G_{ag11} & G_{ag12} \\ G_{ag21} & G_{ag22} \end{bmatrix} = \begin{bmatrix} G_{11} + G_{12}G_{22}^{-1}(G_{gg} - G_{22})G_{22}^{-1}G_{21} & G_{12}G_{22}^{-1}G_{gg} - G_{12}G_{22}^{-1}G_{22} + G_{12} \\ G_{gg}G_{22}^{-1}G_{21} - G_{22}G_{22}^{-1}G_{21} + G_{21} & G_{gg} - G_{22} + G_{22} \end{bmatrix}$$

$$\begin{bmatrix} G_{ag11} & G_{ag12} \\ G_{ag21} & G_{ag22} \end{bmatrix} = \begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix} + \begin{bmatrix} G_{12}G_{22}^{-1}(G_{gg} - G_{22})G_{22}^{-1}G_{21} & G_{12}G_{22}^{-1}(G_{gg} - G_{22}) \\ (G_{gg} - G_{22})G_{22}^{-1}G_{21} & G_{gg} - G_{22} \end{bmatrix}$$

where

$$\begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix} = \text{polygenic additive genetic covariance matrix, and}$$

$$\begin{bmatrix} G_{12}G_{22}^{-1}(G_{gg} - G_{22})G_{22}^{-1}G_{21} & G_{12}G_{22}^{-1}(G_{gg} - G_{22}) \\ (G_{gg} - G_{22})G_{22}^{-1}G_{21} & G_{gg} - G_{22} \end{bmatrix} = \text{difference between the polygenic-genomic}$$

covariance matrix and the polygenic additive genetic covariance matrix.

Inverse of the polygenic-genomic covariance matrix

The inverse of the polygenic-genomic covariance matrix can be obtained directly by expressing the joint distribution of the polygenic-genomic values of all animals as the product of the conditional distribution of the polygenic-genomic values of all animals given the polygenic-genomic values of genotyped animals times the distribution of the polygenic-genomic values of genotyped animals (Aguilar et al., 2010), i.e.,

$$f(a_1, a_2) = f(a_1, a_2 | a_2) f(a_2)$$

$$f(a_1, a_2) = f(a_1 | a_2) f(a_2)$$

$$f(a_1, a_2) \propto \exp \left[-\frac{1}{2} (a_1 - G_{12} G_{22}^{-1} a_2)' G^{11} (a_1 - G_{12} G_{22}^{-1} a_2) \right] \exp \left[-\frac{1}{2} a_2' G^{gg} a_2 \right]$$

$$f(a_1, a_2) \propto \exp \left[-\frac{1}{2} \begin{bmatrix} a_1' & a_2' \end{bmatrix} \begin{bmatrix} G^{11} & -G^{11} G_{12} G_{22}^{-1} \\ -G_{22}^{-1} G_{21} G^{11} & G^{gg} + G_{22}^{-1} G_{21} G^{11} G_{12} G_{22}^{-1} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \right]$$

But,

$$G^{22} = G_{22}^{-1} + G_{22}^{-1} G_{21} G^{11} G_{12} G_{22}^{-1}$$

Proof:

$$\begin{bmatrix} G^{11} & G^{12} \\ G^{21} & G^{22} \end{bmatrix} \begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$$

where

$$G^{11} G_{11} + G^{12} G_{21} = I$$

$$G^{11} G_{12} + G^{12} G_{22} = 0$$

$$G^{21} G_{11} + G^{22} G_{21} = 0$$

$$G^{21} G_{12} + G^{22} G_{22} = I$$

Also,

$$\begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix} \begin{bmatrix} G^{11} & G^{12} \\ G^{21} & G^{22} \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$$

where,

$$G_{11}G^{11} + G_{12}G^{21} = I$$

$$G_{11}G^{12} + G_{12}G^{22} = 0$$

$$G_{21}G^{11} + G_{22}G^{21} = 0$$

$$G_{21}G^{12} + G_{22}G^{22} = I$$

From the [fourth equation of the first set](#):

$$G^{21}G_{12} + G^{22}G_{22} = I$$

$$G^{22}G_{22} = I - G^{21}G_{12}$$

$$G^{22} = (I - G^{21}G_{12})G_{22}^{-1}$$

From the [third equation of the second set](#):

$$G_{21}G^{11} + G_{22}G^{21} = 0$$

$$G_{22}G^{21} = -G_{21}G^{11}$$

$$G^{21} = -G_{22}^{-1}G_{21}G^{11}$$

Substituting $-G_{22}^{-1}G_{21}G^{11}$ for G^{21} in $G^{22} = (I - G^{21}G_{12})G_{22}^{-1}$ yields:

$$G^{22} = (I - (-G_{22}^{-1}G_{21}G^{11})G_{12})G_{22}^{-1}$$

$$G^{22} = (I + G_{22}^{-1}G_{21}G^{11}G_{12})G_{22}^{-1}$$

$$G^{22} = G_{22}^{-1} + G_{22}^{-1}G_{21}G^{11}G_{12}G_{22}^{-1}$$

[Adding and subtracting \$G_{22}^{-1}\$ in \$f\(a_1, a_2\)\$ yields:](#)

$$f(a_1, a_2) \propto \exp \left[-\frac{1}{2} \begin{bmatrix} a_1' & a_2' \end{bmatrix} \begin{bmatrix} G^{11} & -G^{11}G_{12}G_{22}^{-1} \\ -G_{22}^{-1}G_{21}G^{11} & G^{22} - G_{22}^{-1} + G_{22}^{-1} + G_{22}^{-1}G_{21}G^{11}G_{12}G_{22}^{-1} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \right]$$

$$f(a_1, a_2) \propto \exp \left[-\frac{1}{2} [a_1' \quad a_2'] \begin{bmatrix} G^{11} & G^{12} \\ G^{21} & G^{22} + G^{gg} - G_{22}^{-1} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \right]$$

$$f(a_1, a_2) \propto \exp \left[-\frac{1}{2} [a_1' \quad a_2'] \left[\begin{bmatrix} G^{11} & G^{12} \\ G^{21} & G^{22} \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & G^{gg} - G_{22}^{-1} \end{bmatrix} \right] \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \right]$$

Thus, the inverse of the polygenic-genomic covariance matrix is:

$$\begin{bmatrix} G_{ag11} & G_{ag12} \\ G_{ag21} & G_{ag22} \end{bmatrix}^{-1} = \begin{bmatrix} G^{11} & G^{12} \\ G^{21} & G^{22} + G^{gg} - G_{22}^{-1} \end{bmatrix}$$

$$\begin{bmatrix} G_{ag11} & G_{ag12} \\ G_{ag21} & G_{ag22} \end{bmatrix}^{-1} = \begin{bmatrix} G^{11} & G^{12} \\ G^{21} & G^{22} \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & G^{gg} - G_{22}^{-1} \end{bmatrix}$$

where

$\begin{bmatrix} G^{11} & G^{12} \\ G^{21} & G^{22} \end{bmatrix}$ = inverse of the polygenic additive genetic covariance matrix for all animals (i.e., non-genotyped and genotyped),

G^{gg} = inverse of the marker locus additive genomic covariance matrix for genotyped animals, and

G_{22}^{-1} = inverse of the polygenic additive genetic covariance matrix for genotyped animals.

Because the additive genomic covariance matrix G_{gg} is frequently singular or close to singular,

VanRaden (2008) suggested using a **weighted sum of G_{gg} and G_{22}** , i.e.,

$$G_{wgg} = wG_{gg} + (1 - w)G_{22}$$

where values of w could be between 0.95 and 0.98.

Thus, the **approximate inverse of the polygenic-genomic additive covariance matrix** is:

$$\begin{bmatrix} G_{ag11} & G_{ag12} \\ G_{ag21} & G_{ag22} \end{bmatrix}^{-1} = \begin{bmatrix} G^{11} & G^{12} \\ G^{21} & G^{22} \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & (wG_{gg} + (1 - w)G_{22})^{-1} - G_{22}^{-1} \end{bmatrix}$$

or

$$\begin{bmatrix} G_{ag11} & G_{ag12} \\ G_{ag21} & G_{ag22} \end{bmatrix}^{-1} = \begin{bmatrix} G^{11} & G^{12} \\ G^{21} & G^{22} \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & G^{wgg} - G_{22}^{-1} \end{bmatrix}$$

where

$$G^{wgg} = (wG_{gg} + (1-w)G_{22})^{-1}$$

Mixed Model Equations

The mixed model equations for the polygenic-genomic model are:

$$\begin{bmatrix} X'R^{-1}X & X'R^{-1}Z_aQ_a & X'R^{-1}Z_nQ_n & X'R^{-1}Z_{a1} & X'R^{-1}Z_{a2} \\ Q'_aZ'_aR^{-1}X & Q'_aZ'_aR^{-1}Z_aQ_a & Q'_aZ'_aR^{-1}Z_nQ_n & Q'_aZ'_aR^{-1}Z_{a1} & Q'_aZ'_aR^{-1}Z_{a2} \\ Q'_nZ'_nR^{-1}X & Q'_nZ'_nR^{-1}Z_aQ_a & Q'_nZ'_nR^{-1}Z_nQ_n & Q'_nZ'_nR^{-1}Z_{a1} & Q'_nZ'_nR^{-1}Z_{a2} \\ Z'_{a1}R_1^{-1}X_1 & Z'_{a1}R_1^{-1}Z_{a1}Q_{a1} & Z'_{a1}R_1^{-1}Z_{n1}Q_{n1} & Z'_{a1}R_1^{-1}Z_{a1} + G^{11} & G^{12} \\ Z'_{a2}R_2^{-1}X_2 & Z'_{a2}R_2^{-1}Z_{a2}Q_{a2} & Z'_{a2}R_2^{-1}Z_{n2}Q_{n2} & G^{21} & Z'_{a2}R_2^{-1}Z_{a2} + G^{22} + G^{gg} - G_{22}^{-1} \end{bmatrix} \begin{bmatrix} b \\ g_a \\ g_n \\ a_1 \\ a_2 \end{bmatrix}$$

$$= \begin{bmatrix} X'R^{-1}y \\ Q'_aZ'_aR^{-1}y \\ Q'_nZ'_nR^{-1}y \\ Z'_{a1}R_1^{-1}y_1 \\ Z'_{a2}R_2^{-1}y_2 \end{bmatrix}$$

where $y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$, $X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$, $Z_aQ_a = \begin{bmatrix} Z_{a1}Q_{a1} \\ Z_{a2}Q_{a2} \end{bmatrix}$, $Z_nQ_n = \begin{bmatrix} Z_{n1}Q_{n1} \\ Z_{n2}Q_{n2} \end{bmatrix}$, and $R^{-1} = \begin{bmatrix} R_1^{-1} & 0 \\ 0 & R_2^{-1} \end{bmatrix}$.

Remark: In the Example: Genotyped animals are 1 to 6 (a₂) and non-genotyped animals are 7 to 12 (a₁).

Matrix $[X \ Z_aQ_a \ Z_nQ_n \ Z_{a2} \ Z_{a1}]$ in the Example is equal to:

The polygenic additive genetic covariance matrix, i.e., matrix $\begin{bmatrix} G_{22} & G_{21} \\ G_{12} & G_{11} \end{bmatrix}$ is equal to:

$ga = \begin{bmatrix} G_{22} & G_{21} \\ G_{12} & G_{11} \end{bmatrix}$											
36	0	0	18	18	18	0	27	22.5	9	9	9
0	44	22	0	22	11	11	0	11	22	5.5	27.5
0	22	40	0	11	20	20	0	5.5	25.5	10	21
18	0	0	40	9	9	0	29	19	4.5	4.5	4.5
18	22	11	9	40	14.5	5.5	13.5	26.75	25.5	7.25	18.25
18	11	20	9	14.5	43.5	10	13.5	14	17.25	21.75	27.25
0	11	20	0	5.5	10	47.5	0	2.75	12.75	5	10.5
27	0	0	29	13.5	13.5	0	52.5	33	6.75	6.75	6.75
22.5	11	5.5	19	26.75	14	2.75	33	55.375	16.125	7	12.5
9	22	25.5	4.5	25.5	17.25	12.75	6.75	16.125	56.5	8.625	19.625
9	5.5	10	4.5	7.25	21.75	5	6.75	7	8.625	45.125	13.625
9	27.5	21	4.5	18.25	27.25	10.5	6.75	12.5	19.625	13.625	50.625

The genomic variance for genotyped animals a_2 , i.e., $var(a_2) = Z_{ma}Z'_{ma} \left(\frac{\sigma_a^2}{kma} \right) = G_{gg}$, where

$$kma = 2 \sum_{j=1}^{j=nma} p_j(1 - p_j).$$

Matrix $\begin{bmatrix} Z_{ma} \\ \mathbf{0} \end{bmatrix}$ is equal to:

$zg = \begin{bmatrix} Z_{ma} \\ \mathbf{0} \end{bmatrix}$											
2	1	1	2	1	1	2	2	...	0	2	2
0	1	2	0	0	1	2	2	...	1	2	1
1	2	0	0	1	1	1	2	...	1	1	2
1	2	0	0	0	2	2	0	...	1	2	1
1	2	0	1	1	1	1	1	...	2	2	1
2	1	2	2	1	2	2	1	...	1	2	1

0	0	0	0	0	0	0	0	...	0	0	0
0	0	0	0	0	0	0	0	...	0	0	0
0	0	0	0	0	0	0	0	...	0	0	0
0	0	0	0	0	0	0	0	...	0	0	0
0	0	0	0	0	0	0	0	...	0	0	0
0	0	0	0	0	0	0	0	...	0	0	0

Matrix $\begin{bmatrix} Z_{ma}Z'_{ma} & 0 \\ 0 & 0 \end{bmatrix}$ is equal to:

$zgtzg = \begin{bmatrix} Z_{ma}Z'_{ma} & 0 \\ 0 & 0 \end{bmatrix}$											
98	73	76	65	84	79	0	0	0	0	0	0
73	115	79	80	86	84	0	0	0	0	0	0
76	79	98	72	79	79	0	0	0	0	0	0
65	80	72	92	73	76	0	0	0	0	0	0
84	86	79	73	108	80	0	0	0	0	0	0
79	84	79	76	80	105	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0

The additive genetic variance $\sigma_a^2 = 36$ and $kma = 2 \sum_{j=1}^{j=nma} p_j(1 - p_j) = 20.974159$.

Thus, matrix $\begin{bmatrix} G_{gg} & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} Z_{ma}Z'_{ma} \left(\frac{\sigma_a^2}{kma}\right) & 0 \\ 0 & 0 \end{bmatrix}$ is equal to:

$ga22 = \begin{bmatrix} G_{22} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$											
0	0	0	0	0	0	0	0	0	0	0	0

Weighing factor w for $G_{wgg} = wG_{gg} + (1 - w)G_{22}$ was assumed to be 1 for the Example, consequently $G_{wgg} = G_{gg}$.

Thus, the **inverse of the polygenic-genomic covariance matrix** for the Example was computed as

follows: $\begin{bmatrix} G_{ag22} & G_{ag21} \\ G_{ag12} & G_{ag11} \end{bmatrix}^{-1} = \begin{bmatrix} G^{22} & G^{21} \\ G^{12} & G^{11} \end{bmatrix} + \begin{bmatrix} G^{gg} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} + \begin{bmatrix} -G_{22}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$, where:

1) The inverse of matrix $\begin{bmatrix} G_{22} & G_{21} \\ G_{12} & G_{11} \end{bmatrix}$, i.e., matrix $\begin{bmatrix} G^{22} & G^{21} \\ G^{12} & G^{11} \end{bmatrix}$ is equal to:

$gainv = \begin{bmatrix} G^{22} & G^{21} \\ G^{12} & G^{11} \end{bmatrix}$											
0.069	0.013	0.010	-0.006	-0.025	-0.020	0.000	-0.020	0.000	0.000	0.000	0.000
0.013	0.055	-0.017	0.000	-0.025	0.011	0.000	0.000	0.000	0.000	0.000	-0.022
0.010	-0.017	0.059	0.000	0.008	-0.020	-0.013	0.000	0.000	-0.016	0.000	0.000
-0.006	0.000	0.000	0.042	0.000	0.000	0.000	-0.020	0.000	0.000	0.000	0.000
-0.025	-0.025	0.008	0.000	0.068	0.000	0.000	0.010	-0.020	-0.016	0.000	0.000
-0.020	0.011	-0.020	0.000	0.000	0.059	0.000	0.000	0.000	0.000	-0.015	-0.022
0.000	0.000	-0.013	0.000	0.000	0.000	0.027	0.000	0.000	0.000	0.000	0.000
-0.020	0.000	0.000	-0.020	0.010	0.000	0.000	0.051	-0.020	0.000	0.000	0.000
0.000	0.000	0.000	0.000	-0.020	0.000	0.000	-0.020	0.039	0.000	0.000	0.000
0.000	0.000	-0.016	0.000	-0.016	0.000	0.000	0.000	0.000	0.032	0.000	0.000
0.000	0.000	0.000	0.000	0.000	-0.015	0.000	0.000	0.000	0.000	0.029	0.000
0.000	-0.022	0.000	0.000	0.000	-0.022	0.000	0.000	0.000	0.000	0.000	0.043

$ginvga22 = \begin{bmatrix} G_{z22}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$											
0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Thus, the inverse of the polygenic-genomic covariance matrix, i.e., matrix $\begin{bmatrix} G_{ag22} & G_{ag21} \\ G_{ag12} & G_{ag11} \end{bmatrix}^{-1}$ is:

$ginvgag = \begin{bmatrix} G_{ag22} & G_{ag21} \\ G_{ag12} & G_{ag11} \end{bmatrix}^{-1}$											
0.034	0.002	-0.006	0.011	-0.011	-0.007	0.000	-0.020	0.000	0.000	0.000	0.000
0.002	0.029	-0.003	-0.006	-0.006	0.006	0.000	0.000	0.000	0.000	0.000	-0.022
-0.006	-0.003	0.037	-0.005	0.005	-0.004	-0.013	0.000	0.000	-0.016	0.000	0.000
0.011	-0.006	-0.005	0.032	-0.003	-0.006	0.000	-0.020	0.000	0.000	0.000	0.000
-0.011	-0.006	0.005	-0.003	0.041	-0.000	0.000	0.010	-0.020	-0.016	0.000	0.000
-0.007	0.006	-0.004	-0.006	-0.000	0.040	0.000	0.000	0.000	0.000	-0.015	-0.022
0.000	0.000	-0.013	0.000	0.000	0.000	0.027	0.000	0.000	0.000	0.000	0.000
-0.020	0.000	0.000	-0.020	0.010	0.000	0.000	0.051	-0.020	0.000	0.000	0.000
0.000	0.000	0.000	0.000	-0.020	0.000	0.000	-0.020	0.039	0.000	0.000	0.000
0.000	0.000	-0.016	0.000	-0.016	0.000	0.000	0.000	0.000	0.032	0.000	0.000
0.000	0.000	0.000	0.000	0.000	-0.015	0.000	0.000	0.000	0.000	0.029	0.000
0.000	-0.022	0.000	0.000	0.000	-0.022	0.000	0.000	0.000	0.000	0.000	0.043

The LHS matrix, the vector of effects, and the RHS vector of the mixed model equations for the Example are:

lhs																	Effect	rhs	
0.354	0.154	0.201	0.197	0.158	0.197	0.020	-0.063	0.031	0.031	0.031	0.021	0.033	0.021	0.021	0.022	0.030	0.030	Mean	94.88
0.154	0.092	0.062	0.095	0.082	0.071	0.020	0.000	0.015	0.015	0.015	0.016	0.008	0.016	0.013	0.011	0.011	0.011	Breed _A	42.10
0.201	0.062	0.139	0.102	0.075	0.125	0.000	-0.063	0.015	0.015	0.015	0.005	0.025	0.005	0.008	0.011	0.019	0.019	Breed _B	52.78
0.197	0.095	0.102	0.156	0.092	0.105	0.000	0.000	0.031	0.031	0.031	0.011	0.016	0.011	0.011	0.011	0.023	0.023	Heterosis _{AB}	53.36
0.158	0.082	0.075	0.092	0.158	0.000	0.020	0.000	0.000	0.000	0.031	0.021	0.033	0.000	0.000	0.022	0.000	0.030	Sex _M	44.53
0.197	0.071	0.125	0.105	0.000	0.197	0.000	0.063	0.031	0.031	0.000	0.000	0.000	0.021	0.021	0.000	0.030	0.000	Sex _F	50.35
0.020	0.020	0.000	0.000	0.020	0.000	0.054	0.002	-0.006	0.011	-0.011	-0.007	0.000	-0.020	0.000	0.000	0.000	0.000	a ₁	5.90
0.063	0.000	0.063	0.000	0.000	0.063	0.002	0.091	-0.003	-0.006	-0.006	0.006	0.000	0.000	0.000	0.000	0.000	-0.022	a ₂	15.31
0.031	0.015	0.015	0.031	0.000	0.031	-0.006	-0.003	0.068	-0.005	0.005	-0.004	-0.013	0.000	0.000	-0.016	0.000	0.000	a ₃	7.88
0.031	0.015	0.015	0.031	0.000	0.031	0.011	-0.006	-0.005	0.063	-0.003	-0.006	0.000	-0.020	0.000	0.000	0.000	0.000	a ₄	8.03
0.031	0.015	0.015	0.031	0.031	0.000	-0.011	-0.006	0.005	-0.003	0.071	0	0.000	0.010	-0.020	-0.016	0.000	0.000	a ₅	8.98
0.021	0.016	0.005	0.011	0.021	0.000	-0.007	0.006	-0.004	-0.006	0	0.061	0.000	0.000	0.000	0.000	-0.015	-0.022	a ₆	6.09
0.033	0.008	0.025	0.016	0.033	0.000	0.000	0.000	-0.013	0.000	0.000	0.000	0.059	0.000	0.000	0.000	0.000	0.000	a ₇	8.92
0.021	0.016	0.005	0.011	0.000	0.021	-0.020	0.000	0.000	-0.020	0.010	0.000	0.000	0.072	-0.020	0.000	0.000	0.000	a ₈	5.62
0.021	0.013	0.008	0.011	0.000	0.021	0.000	0.000	0.000	0.000	-0.020	0.000	0.000	-0.020	0.060	0.000	0.000	0.000	a ₉	5.68
0.022	0.011	0.011	0.011	0.022	0.000	0.000	0.000	-0.016	0.000	-0.016	0.000	0.000	0.000	0.000	0.054	0.000	0.000	a ₁₀	6.18
0.030	0.011	0.019	0.023	0.000	0.030	0.000	0.000	0.000	0.000	0.000	-0.015	0.000	0.000	0.000	0.000	0.059	0.000	a ₁₁	7.83
0.030	0.011	0.019	0.023	0.030	0.000	0.000	-0.022	0.000	0.000	0.000	-0.022	0.000	0.000	0.000	0.000	0.000	0.073	a ₁₂	8.47

Solutions to the mixed model equations were computed as the product of the generalized inverse of the LHS times the RHS. Standard errors of solutions were computed as square roots of the diagonal elements of the LHS. The vector of solutions and their standard errors for the example are:

Effect	SOL	SESOL
Mean ^o	132.89	4.78
Breed _A ^o	77.37	6.96
Breed _B ^o	55.52	5.83
Heterosis _{AB} ^o	7.3	6.56
Sex _M ^o	77.61	3.37
Sex _F ^o	55.28	3.51
\hat{a}_1	1.49	10.91
\hat{a}_2	1.21	9.65
\hat{a}_3	-3.54	9.87
\hat{a}_4	-0.28	9.51
\hat{a}_5	5.03	9.61
\hat{a}_6	0.18	10.49
\hat{a}_7	-2.52	7.93
\hat{a}_8	1.14	10.33
\hat{a}_9	5.15	10.08
\hat{a}_{10}	-0.62	9.88
\hat{a}_{11}	0.88	7.87
\hat{a}_{12}	0.53	9.50

Prediction of polygenic-genomic values

Predictions of additive polygenic-genomic values (\hat{u}_{ai}) are computed as the sum of the breed effect and the animal polygenic-genomic deviation from the breed effect, i.e.,

$$\hat{u}_{ai} = \text{Prob}_{A_{\text{anim } i}} * (\text{Breed}_A^\circ - \text{Breed}_B^\circ) + \hat{a}_i$$

where

$\text{Prob}_{A_{\text{anim } i}}$ = expected fraction of breed A in animal i,

Breed_A° = generalized least squares solution for breed A,

Breed_B° = generalized least squares solution for breed B, and

\hat{a}_i = BLUP of a_i .

The matrix of variances of errors of prediction (VEP) for the vector of predicted genetic values (\hat{u}_a) is:

$$\text{VEP}(\hat{u}_a) = \text{var}(\hat{u}_a - u_a) = \mathbf{K}_a * (\text{ginvlhs}) * \mathbf{K}_a'$$

where

\mathbf{K}_a = $n_{\text{anim}} \times n_{\text{eq}}$ matrix specifying multipliers for breed differences and random animal deviations, where n_{anim} = number of animals, and n_{eq} = number of equations, and

ginvlhs = $n_{\text{eq}} \times n_{\text{eq}}$ generalized inverse matrix of the left hand side of the MME.

The standard errors of prediction (SEP) of \hat{u}_{ai} are computed as the square roots of the diagonal elements of the $\text{VEP}(\hat{u}_{ai})$ matrix.

Matrix \mathbf{K}_a for the Example is:

ka																
0.0	1.00	-1.00	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.00	0.00	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.50	-0.50	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.50	-0.50	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.50	-0.50	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0
0.0	0.75	-0.75	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0
0.0	0.25	-0.25	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0
0.0	0.75	-0.75	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0
0.0	0.625	-0.625	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0
0.0	0.50	-0.50	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0
0.0	0.375	-0.375	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0
0.0	0.375	-0.375	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0

The \hat{u}_{ai} and the SEP (\hat{u}_a) for the animals in Example are:

Animal	Prob _{Aanim}	Breed _A ^o - Breed _B ^o	\hat{a}_{add}	\hat{u}_a	SEP (\hat{u}_a)
1	1	21.85	1.49	23.35	11.73
2	0	21.85	1.21	1.21	9.65
3	0.5	21.85	-3.54	7.38	10.45
4	0.5	21.85	-0.28	10.65	10.08
5	0.5	21.85	5.03	15.96	10.36
6	0.75	21.85	0.18	16.57	11.65
7	0.25	21.85	-2.52	2.94	8.62

8	0.75	21.85	1.14	17.54	10.33
9	0.625	21.85	5.15	18.81	10.24
10	0.5	21.85	-0.62	10.31	10.64
11	0.375	21.85	0.88	9.08	8.38
12	0.375	21.85	0.53	8.73	10.41

Summary of predictions by polygenic-genomic models

Animal	Polygenic		Marker Polygenic-Genomic with Actual + Predicted Genotypes		Animal Polygenic-Genomic with Actual + Predicted Genotypes		Animal Polygenic-Genomic with Phenotypes Pedigree Genotypes	
	\hat{u}_a	SEP (\hat{u}_a)	\hat{u}_a	SEP (\hat{u}_a)	\hat{u}_a	SEP (\hat{u}_a)	\hat{u}_a	SEP (\hat{u}_a)
1	22.40	9.76	20.58	14.8	20.58	14.8	23.35	11.73
2	0.63	5.73	-0.92	13.25	-0.92	13.25	1.21	9.65
3	7.12	7.93	5.16	13.8	5.16	13.8	7.38	10.45
4	10.59	6.52	8.6	13.65	8.6	13.65	10.65	10.08
5	14.04	7.82	13.18	13.6	13.18	13.6	15.96	10.36
6	15.01	9.10	14.27	15.13	14.27	15.13	16.57	11.65
7	2.20	7.10	-0.78	12.92	-0.78	12.92	2.94	8.62
8	17.12	7.72	14.78	13.75	14.78	13.75	17.54	10.33
9	17.89	7.75	15.97	13.49	15.97	13.49	18.81	10.24
10	9.16	8.49	9.58	14.81	9.58	14.81	10.31	10.64
11	8.45	6.59	6.23	13.33	6.23	13.33	9.08	8.38
12	7.59	7.86	6.22	13.58	6.22	13.58	8.73	10.41

References

- Aguilar, I., I. Misztal, D. L. Johnson, A. Legarra, S. Tsuruta, and T. J. Lawlor. 2009. A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluations of Holstein final score. *J. Dairy Sci.* 93:743-752.
- Gengler, N., P. Mayeres, and M. Szydlowski. 2007. A simple method to approximate gene content in large pedigree populations: application to the myostatin gene in dual purpose Belgian Blue cattle. *Animal* 1:21-28.
- Gianola, D., Campos, G. de los, Hill, W. G., Manfredi, E., Fernando, R., 2009. Additive genetic variability and the Bayesian alphabet. *Genetics* 183, 347-363.
- Legarra, A., I. Aguilar, and I. Misztal. 2009. A relationship matrix including full pedigree and genomic information. *J. Dairy Sci.* 92:4656-4663.
- Meuwissen, T. H. E., B. J. Hayes, and M. E. Goddard. 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157:1819-1829.
- Nkrumah, J. D., C. Li, J. Yu, C. Hansen, D. H. Keisler, and S. S. Moore. 2005. Polymorphisms in the bovine leptin promoter associated with serum leptin concentration, growth, feed intake, feeding behavior, and measures of carcass merit. *J. Anim. Sci.* 83:20-28.
- VanRaden, P. M. 2008. Efficient methods to compute genomic predictions. *J. Dairy Sci.* 91:4414-4423.
- VanRaden, P. M. 2011. Findhap.f90. Available at: <http://aipl.arsusda.gov/software/findhap>.
- VanRaden, P. M., Null, D. J., Sargolzaei, M., Wiggans, G. R., Tooker, M. E., Cole, J. B., Sonstegard, T. S., Connor, E. E., Winters, M., van Kaam, J. B. C. H. M., Valentini, A., Van Doormaal, B. J., Faust, M. A., Doak, G. A., 2013. Genomic imputation and evaluation using high-density Holstein genotypes. *J. Dairy Sci.* 96, 668-678.