IT-Solutions for
Animal Production

A SNP MACE model for international genomic evaluation:

- technical challenges and possible solutions

Zengting Liu and Mike E. Goddard

vit, IT solutions for animal production, Germany

University of Melbourne, Melbourne 3083, VIC, Australia

Overview

Current status of (inter)national conventional and genomic evaluations

- A SNP MACE model for international evaluation
- Solving algorithms for the SNP MACE model
- Approximation of prediction error (co)variances
- Further development and extension
  - Different SNP sets across countries

    Countries use MACE info as phenotypes

International evaluation based on SNP effects

- Concept of a SNP based international evaluation (Goddard, 2011)
- A SNP-Focus Model replacing the Animal Model (Schaeffer, 2014)
- Interbull estimation of SNP effects (Goddard, 2016)
- A SNP MACE model proposed (Goddard, 2017)
  - Interbull Technical Workshop, Slovenia
- Interbull project on the SNP MACE model

A SNP MACE model

- A SNP genomic model for Multiple Across Country Evaluation

$$\mathbf{g}_i^N \quad ==\gg \quad \mathbf{g}_i \qquad [1]$$

for country i (i = 1, …, c)

- A SNP BLUP model for national genomic evaluation

$$\mathbf{y}_i = \mu_i \mathbf{1} + \mathbf{Z}_i \mathbf{g}_i^N + \mathbf{e}_i \qquad [2]$$

where $\mathbf{y}_i$ is phenotype after absorbing all other effects (including the residual polygenic effect)

$$var(\mathbf{e}_i) = \mathbf{R}_i^{-1} = diag\{n_{ik}\sigma_{e_i}^{-2}\}$$

$$var(\mathbf{g}_i) = \mathbf{B}_i \sigma_i^2 \quad \text{with DGV variance } \sigma_i^2$$

$$\mathbf{B}_i = \frac{1}{\sum_i 2p_{ii}(1-p_{ii})}\mathbf{I} = \theta_i \mathbf{I}$$

SNP genetic (co)variances between countries (I)

- A (co)variance matrix for countries ($i = 1, \ldots, c$) **for a single SNP marker**

$$\mathbf{G}_{cou} = \begin{bmatrix} \sigma_1^2\theta_1 & r_{12}\sigma_1\sigma_2\sqrt{\theta_1\theta_2} & \cdots & r_{1c}\sigma_1\sigma_c\sqrt{\theta_1\theta_c} \\ & \sigma_2^2\theta_2 & \cdots & r_{2c}\sigma_2\sigma_c\sqrt{\theta_2\theta_c} \\ & & \ddots & \vdots \\ & symm. & & \sigma_c^2\theta_c \end{bmatrix} = \begin{bmatrix} g_{11} & g_{12} & \cdots & g_{1c} \\ & g_{22} & \cdots & g_{2c} \\ & & \ddots & \vdots \\ & symm. & & g_{cc} \end{bmatrix}$$

- Its inverse $\mathbf{G}_{cou}^{-1} = \begin{bmatrix} g^{11} & g^{12} & \cdots & g^{1c} \\ & g^{22} & \cdots & g^{2c} \\ & & \ddots & \vdots \\ & symm. & & g^{cc} \end{bmatrix}$

SNP genetic (co)variances between countries (II)

- Genetic (co)variance matrix for **ALL** SNP effects **(ordered by countries)**

$$\text{var}\begin{bmatrix}\mathbf{g_1}\\ \mathbf{g_2}\\ \vdots\\ \mathbf{g_c}\end{bmatrix} = \mathbf{G} = \mathbf{G_{cou}} \otimes \mathbf{I} = \begin{bmatrix} g_{11}\mathbf{I} & g_{12}\mathbf{I} & \cdots & g_{1c}\mathbf{I} \\ & g_{22}\mathbf{I} & \cdots & g_{2c}\mathbf{I} \\ & & \ddots & \vdots \\ symm. & & & g_{cc}\mathbf{I} \end{bmatrix}$$

- Its inverse $\mathbf{G^{-1}} = \mathbf{G_{cou}^{-1}} \otimes \mathbf{I} = \begin{bmatrix} g^{11}\mathbf{I} & g^{12}\mathbf{I} & \cdots & g^{1c}\mathbf{I} \\ & g^{22}\mathbf{I} & \cdots & g^{2c}\mathbf{I} \\ & & \ddots & \vdots \\ symm. & & & g^{cc}\mathbf{I} \end{bmatrix}$
- Inter-SNP genetic correlations: within or between countries are all 0
- Intra-SNP genetic correlations between countries to be estimated
  - Set to country correlations as in current MACE

Mixed model equations of the SNP MACE model

$$\begin{bmatrix} \ddots & \cdots & \cdots & \\ & \left[\begin{bmatrix} \mathbf{1'Z_i'R_i^{-1}Z_i1} & \mathbf{1'Z_i'R_i^{-1}Z_i} \\ \mathbf{Z_i'R_i^{-1}Z_i1} & \mathbf{Z_i'R_i^{-1}Z_i} \end{bmatrix} + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & g^{ii}\mathbf{I} \end{bmatrix}\right] & \left[\mathbf{\Psi}_{ii^+} + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & g^{ii^+}\mathbf{I} \end{bmatrix}\right] & \\ & \ddots & \left[\begin{bmatrix} \mathbf{1'Z_{i^+}'R_{i^+}^{-1}Z_{i^+}1} & \mathbf{1'Z_{i^+}'R_{i^+}^{-1}Z_{i^+}} \\ \mathbf{Z_{i^+}'R_{i^+}^{-1}Z_{i^+}1} & \mathbf{Z_{i^+}'R_{i^+}^{-1}Z_{i^+}} \end{bmatrix} + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & g^{i^+i^+}\mathbf{I} \end{bmatrix}\right] & \ddots \end{bmatrix}$$

$$\mathrm{x}\begin{bmatrix} \vdots \\ \hat{\mu}_i \\ \hat{\mathbf{g}}_i \\ \vdots \\ \hat{\mu}_{i^+} \\ \hat{\mathbf{g}}_{i^+} \\ \vdots \end{bmatrix} = \begin{bmatrix} \vdots \\ \mathbf{1'Z_i'R_i^{-1}y_i} \\ \mathbf{Z_i'R_i^{-1}y_i} \\ \vdots \\ \mathbf{1'Z_{i^+}'R_{i^+}^{-1}y_{i^+}} \\ \mathbf{Z_{i^+}'R_{i^+}^{-1}y_{i^+}} \\ \vdots \end{bmatrix}$$

Zero residual covariances between countries $\mathbf{\Psi}_{ii^+} = \mathbf{0}$,
if the countries do not use MACE EBV for national genomic evaluation

Solving the mixed model equations

- MME of the SNP MACE model have special structures:
  - Data contribution by country, zero residual covariances (off-diagonals)
  - SNP genetic contribution: only diagonal and sub-diagonals ≠ 0
    - Block-diagonal matrix in the SNP-major order
- Identical processes for every country or every SNP → parallel computing
- PCG algorithm using multiple cores

$$\mathbf{C}\,\mathbf{v} = \{\mathbf{Z}_i'\mathbf{R}_i^{-1}\mathbf{Z}_i\}\,\mathbf{v} + \{\mathbf{G}_{cou}^{-1}\}\mathbf{v}$$

for country i — parallelised by countries

for every SNP marker j — parallelised by SNP markers

- Conditioner may be the inverted diagonal block for country i

$$\mathbf{M}_i = (\mathbf{Z}_i'\mathbf{R}_i^{-1}\mathbf{Z}_i + \mathbf{G}^{ii})^{-1}$$

the matrix $\mathbf{M}_i$ is also used in reliability calculation.

Calculation of prediction error (co)variances

- Countries need to calculate reliabilities of DGV (sum of all SNP effects)
- Not only reliabilities of MACE SNP effect estimates but also (co)reliabilities between the SNP effect estimates
  - the whole PEC block of 50k x 50k
- Absorbing all the other countries into own SNP effects

$$\mathbf{C}_i = \left(\mathbf{Z}_i'\mathbf{R}_i^{-1}\mathbf{Z}_i + \mathbf{G}^{ii}\right) - \sum_{j\neq i}\mathbf{G}^{ij}\left(\mathbf{Z}_j'\mathbf{R}_j^{-1}\mathbf{Z}_j + \mathbf{G}^{jj}\right)^{-1}\mathbf{G}^{ji}$$
$$= \left(\mathbf{Z}_i'\mathbf{R}_i^{-1}\mathbf{Z}_i + \mathrm{g}^{ii}\,\mathbf{I}\right) - \sum_{j\neq i}\mathrm{g}^{ij}\,\mathbf{I}\left(\mathbf{Z}_j'\mathbf{R}_j^{-1}\mathbf{Z}_j + \mathbf{G}^{jj}\right)^{-1}\mathrm{g}^{ji}\,\mathbf{I}$$
$$= \left(\mathbf{Z}_i'\mathbf{R}_i^{-1}\mathbf{Z}_i + \mathrm{g}^{ii}\,\mathbf{I}\right) - \left(\mathrm{g}^{ij}\right)^2\sum_{j\neq i}\mathbf{M}_j$$

- Invert the own block matrix $\mathbf{C}_i^{-1}$
- Provide the PEC matrix to countries $\mathbf{C}_i^{-1}$

Three methods for handling different sets of SNP markers

- Method 1: conversion of country SNP effects to a common set
- Method 2: conversion of SNP effects for GBLUP models
- Method 3: direct modelling heterogeneous sets of SNP markers

Method 1: Conversion of country SNP effects to a common set of SNP markers

- SNP effects of national set of SNP markers for *i*-th country: $\mathbf{g}_c^N$
- SNP effects of a common set of SNP markers: $\mathbf{g}_c^N$
- Define DGV of all reference animals with own set of SNP markers:

$$\mathbf{u}_i = \mathbf{Z}_i \mathbf{g}_i^N$$

- A SNP BLUP model is fitted to model the DGV of reference animals:

$$\mathbf{u}_i = \mathbf{Z}_i^c \mathbf{g}_c^N + \xi$$

$$(\mathbf{Z}_i^{c\prime} \mathbf{R}_i^{-1} \mathbf{Z}_i^c + \sigma_i^{-2} \mathbf{B}_c^{-1}) \mathbf{g}_c^N = \mathbf{Z}_i^{c\prime} \mathbf{R}_i^{-1} \mathbf{u}_i$$

$$\mathbf{g}_c^N = (\mathbf{Z}_i^{c\prime} \mathbf{R}_i^{-1} \mathbf{Z}_i^c + \sigma_i^{-2} \mathbf{B}_c^{-1})^{-1} \mathbf{Z}_i^{c\prime} \mathbf{R}_i^{-1} (\mathbf{Z}_i \mathbf{g}_i^N)$$

- Additional data needed for the conversion

$$\mathbf{Z}_i' \mathbf{R}_i^{-1} \mathbf{Z}_i \qquad \mathbf{Z}_i^{c\prime} \mathbf{R}_i^{-1} \mathbf{Z}_i^c$$

  in addition to $\mathbf{Z}_i' \mathbf{R}_i^{-1} \mathbf{Z}_i$

- Back conversion of MACE SNP effect estimates to the own SNP set

Method 2: Conversion of country SNP effects for GBLUP models

- Country uses a GBLUP model with its own SNP set
- Assumption: equal GEBV for reference animals with both SNP sets
- For all reference animals: GEBV vector $\mathbf{u}_i^*$
- Genomic relationship matrix for all reference animals is invertible: $\mathbf{G}_{rel}^{-1}$
- Estimate SNP effects of the common set

$$\mathbf{g}_i^c = (1-k)\mathbf{B}_c\mathbf{Z}_i^{c\prime}\mathbf{G}_{rel}^{-1}\mathbf{u}_i^*$$

- Equal genomic relationship matrices

$$(1-k)\mathbf{Z}_i^{c\prime}\mathbf{B}_c\mathbf{Z}_i^c + k\mathbf{A}_i = (1-k)\mathbf{Z}_i{}'\mathbf{B}_i\mathbf{Z}_i + k\mathbf{A}_i$$

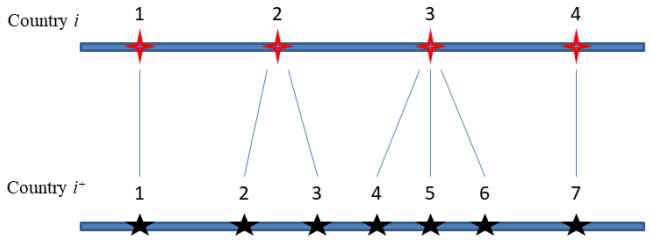
Summary

- The SNP MACE model is an efficient tool for utilizing phenotype info of foreign reference animals
  - Particularly useful for new traits with large-scale genotyped cows
- No requirement for direct access to original national genotype and phenotype data
  - Keep the current infra-structure of national evaluation systems
- Parallel computing for efficiently solving the SNP MACE equations
  - No more pedigree relationship matrix, difficult to be parallelized
- Direct modelling different sets of SNP reduces the need for conversion to a common set of SNP markers
- A gain in accuracy of prediction is expected, especially for novel traits

IT-Solutions for Animal Production

Method 3: Direct modelling of SNP effects of different marker sets

- Cross reference two SNP sets for a country pair (same allele coding)



$$m_i = 4 \quad m_{i^+} = 7$$

- SNP effect covariance matrix

$$\sigma_{i,i^+}\mathbf{B}_{i,i^+} = \sigma_{i,i^+}\sqrt{\theta_i\theta_{i^+}}\mathbf{E}_{i,i^+}$$

$$\mathbf{E}_{i,i^+} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

- Same modelling of **SNP/DGV variances** as in national genomic evaluations
- Correct **covariance of DGV** for any pair of countries
- But **covariances of SNP** are only correct for countries with fewer markers

Considering the SNP array differences

- Modelling SNP effect covariances between countries
  - Country A with more SNPs with unequal SNP covariance: less optimal
  - Country B with fewer SNPs with equal SNP covariance: exact modelling
  - But **DGV covariance** is correct for both countries
- For every one of all country pairs:
  - Set up a SNP cross-reference table
  - Determine the SNP covariances for each country pair
  - In case of a change in SNP arrays in one country, re-do the SNP cross-referencing with all the other countries
- Advantages of the procedure Method 3:
  - Countries do not have to be forced to use the same SNP arrays
  - SNP effects conversion to the common SNP set is not needed
- Disadvantages of Method 3:
  - More work of the SNP MACE, particularly when countries change their SNP sets
  - Approximate inverse of **G** matrix

  Residual covariances to be modelled as the procedure for genetic covariances

Introduction

- Interbull MACE / GMACE evaluation for bulls / genomic bulls
  - Based on national conventional / genomic evaluation
- National genomic evaluation uses MACE EBV of foreign reference bulls
  - Significant increase in accuracy of genomic prediction
  - Fear of domination of foreign reference bulls on own SNP effects
  - Negative impact of genomic pre-selection on conventional EBV of bulls
    - Single-step national evaluation beneficial
- LD info of foreign reference cows NOT used in own SNP effect estimation
  - More countries add cows into national reference population
  - No MACE for cows, exchanging genotype of millions of cows infeasible
- Novel traits have relatively small national reference population
  - MACE bull evaluation perhaps not ready yet
  - Expected to have the largest gain in accuracy of prediction

An example with 2 countries and 3 SNP markers (1)

- Data contribution: least squares parts

$$\begin{bmatrix} \varphi_{\mu_1} & \varphi_{\mu_1 s_1} & \varphi_{\mu_1 s_2} & \varphi_{\mu_1 s_3} & 0 & 0 & 0 & 0 \\ \varphi_{\mu_1 s_1} & \varphi_{1\,s_{11}} & \varphi_{1\,s_{12}} & \varphi_{1\,s_{13}} & 0 & & & \\ \varphi_{\mu_1 s_2} & \varphi_{1\,s_{21}} & \varphi_{1\,s_{22}} & \varphi_{1\,s_{23}} & 0 & & & \\ \varphi_{\mu_1 s_3} & \varphi_{1\,s_{31}} & \varphi_{1\,s_{32}} & \varphi_{1\,s_{33}} & 0 & & & \\ 0 & 0 & 0 & 0 & \varphi_{\mu_2} & \varphi_{\mu_2 s_1} & \varphi_{\mu_2 s_2} & \varphi_{\mu_2 s_3} \\ 0 & & & & \varphi_{\mu_2 s_1} & \varphi_{2\,s_{11}} & \varphi_{2\,s_{12}} & \varphi_{2\,s_{13}} \\ 0 & & & & \varphi_{\mu_2 s_2} & \varphi_{2\,s_{21}} & \varphi_{2\,s_{22}} & \varphi_{2\,s_{23}} \\ 0 & & & & \varphi_{\mu_2 s_3} & \varphi_{2\,s_{31}} & \varphi_{2\,s_{32}} & \varphi_{2\,s_{33}} \end{bmatrix} \begin{bmatrix} \hat{\mu}_1 \\ \hat{g}_{1s_1} \\ \hat{g}_{1s_2} \\ \hat{g}_{1s_3} \\ \hat{\mu}_2 \\ \hat{g}_{2s_1} \\ \hat{g}_{2s_2} \\ \hat{g}_{2s_3} \end{bmatrix} = \begin{bmatrix} \Delta_{\mu_1} \\ \Delta_{1s_1} \\ \Delta_{1s_2} \\ \Delta_{1s_3} \\ \Delta_{\mu_2} \\ \Delta_{2s_1} \\ \Delta_{2s_2} \\ \Delta_{2s_3} \end{bmatrix}$$

- Only the products of matrices or vectors are available, not the matrices or vectors themselves

$$\mathbf{Z_i}'\mathbf{R}_i^{-1}\mathbf{Z_i} \rightarrow \{\varphi_{1\,s_{11}}\}$$

$$\mathbf{Z_i}'\mathbf{R}_i^{-1}\mathbf{y_i} \rightarrow \{\Delta_{2\,s_3}\}$$

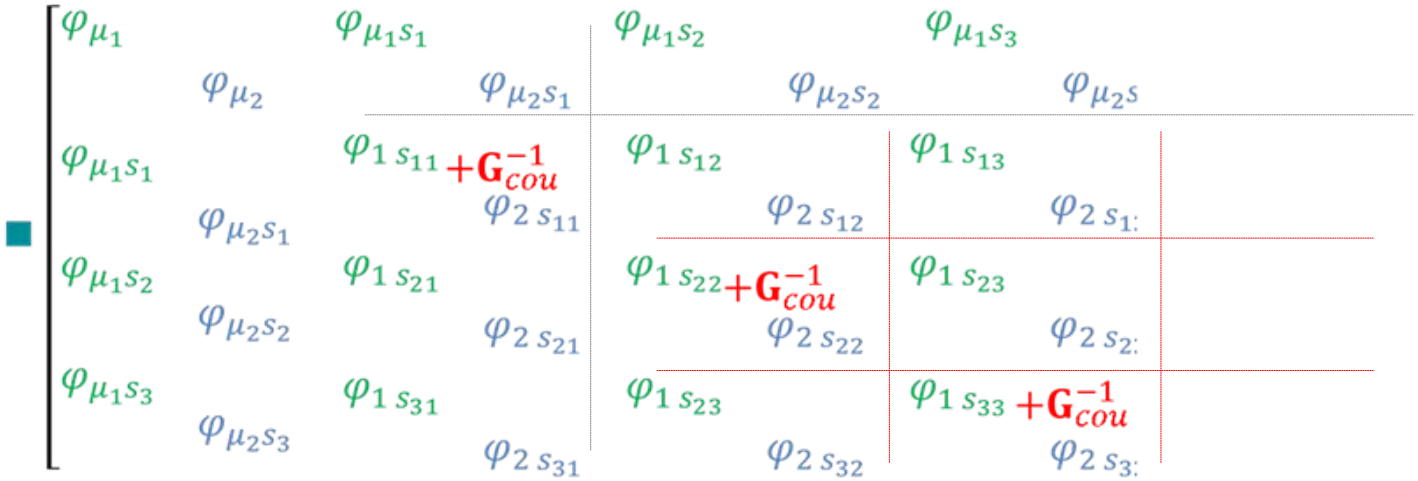
An example with 2 countries and 3 SNP markers (2)

- Data contribution: least squares parts + SNP genetic parts

$$\begin{bmatrix} \varphi_{\mu_1} & \varphi_{\mu_1 s_1} & \varphi_{\mu_1 s_2} & \varphi_{\mu_1 s_3} & 0 & 0 & 0 & 0 \\ \varphi_{\mu_1 s_1} & \varphi_{1\,s_{11}} & \varphi_{1\,s_{12}} & \varphi_{1\,s_{13}} & 0 & & & \\ \varphi_{\mu_1 s_2} & \varphi_{1\,s_{21}} & \varphi_{1\,s_{22}} & \varphi_{1\,s_{23}} & 0 & & & \\ \varphi_{\mu_1 s_3} & \varphi_{1\,s_{31}} & \varphi_{1\,s_{32}} & \varphi_{1\,s_{33}} & 0 & & & \\ 0 & 0 & 0 & 0 & \varphi_{\mu_2} & \varphi_{\mu_2 s_1} & \varphi_{\mu_2 s_2} & \varphi_{\mu_2 s_3} \\ 0 & & & & \varphi_{\mu_2 s_1} & \varphi_{2\,s_{11}} & \varphi_{2\,s_{12}} & \varphi_{2\,s_{13}} \\ 0 & & & & \varphi_{\mu_2 s_2} & \varphi_{2\,s_{21}} & \varphi_{2\,s_{22}} & \varphi_{2\,s_{23}} \\ 0 & & & & \varphi_{\mu_2 s_3} & \varphi_{2\,s_{31}} & \varphi_{2\,s_{32}} & \varphi_{2\,s_{33}} \end{bmatrix} + g^{11} \mathrm{I} + g^{12} \mathrm{I} + g^{21} \mathrm{I} + g^{22} \mathrm{I} \begin{bmatrix} \hat{\mu}_1 \\ \hat{g}_{1s_1} \\ \hat{g}_{1s_2} \\ \hat{g}_{1s_3} \\ \hat{\mu}_2 \\ \hat{g}_{2s_1} \\ \hat{g}_{2s_2} \\ \hat{g}_{2s_3} \end{bmatrix} = \begin{bmatrix} \Delta_{\mu_1} \\ \Delta_{1s_1} \\ \Delta_{1s_2} \\ \Delta_{1s_3} \\ \Delta_{\mu_2} \\ \Delta_{2s_1} \\ \Delta_{2s_2} \\ \Delta_{2s_3} \end{bmatrix}$$

- Order: SNP markers within country (in country-major order)

An example with 2 countries and 3 SNP markers (3)



- Order: countries within SNP marker (in SNP-major order)

$$\mathbf{G}_{cou}^{-1} = \begin{bmatrix} \sigma_1^2\theta_1 & \sigma_{12}\sqrt{\theta_1\theta_2} \\ \sigma_{12}\sqrt{\theta_1\theta_2} & \sigma_2^2\theta_2 \end{bmatrix}^{-1}$$

Why not fitting the residual polygenic effect?

- Needs to know the identifications of reference animals
- Needs to directly access genotypes and phenotypes of reference animals
  - Keep the infra-structure of current national evaluation systems
- In future, millions of cows will be added to ref. pop. worldwide
  - Exchange of genotypes of millions of reference cows may be infeasible
  - Estimating RPG of the millions of cows for all countries is challenging

A SNP MACE model: SNP effect covariances

Country $i = 1, \ldots, c$; SNP marker $j = 1, \ldots, k$; reference animal $k = 1, \ldots, n$

- Countries may have different sets of SNP markers

$$\text{var}\begin{bmatrix}\mathbf{g}_1\\ \mathbf{g}_2\\ \vdots\\ \mathbf{g}_c\end{bmatrix} = \begin{bmatrix}\sigma_1^2\mathbf{B}_1 & \sigma_{12}\mathbf{B}_{12} & \cdots & \sigma_{1c}\mathbf{B}_{1c}\\ & \sigma_2^2\mathbf{B}_2 & \cdots & \sigma_{2c}\mathbf{B}_{2c}\\ & & \ddots & \vdots\\ symm. & & & \sigma_c^2\mathbf{B}_c\end{bmatrix} = \mathbf{G}$$

DGV variance of country i, σ_i^2

DGV covariance between countries i and i^+. σ_{i,i^+}

$$\mathbf{B}_{i,i^+} = \frac{1}{\sqrt{\sum_j 2p_{ij}(1-p_{ij})}\sqrt{\sum_j 2p_{i^+j}(1-p_{i^+j})}}\mathbf{I} = \sqrt{\theta\theta}\,\mathbf{I}$$ for same SNP set

$\mathbf{B}_{i,i^+}$ is not a squared matrix for two different SNP sets $\sqrt{\theta_{ii}\theta_{i^+i^+}}\mathbf{I}_{i,i^+}$

National data for the SNP MACE evaluation: replacing the deregression step of national bull EBV in MACE

- $$\begin{bmatrix} \mathbf{1}'\mathbf{Z}_i'\mathbf{R}_i^{-1}\mathbf{Z}_i\mathbf{1} & \mathbf{1}'\mathbf{R}_i^{-1}\mathbf{Z}_i \\ \mathbf{Z}_i'\mathbf{R}_i^{-1}\mathbf{1} & \mathbf{Z}_i'\mathbf{R}_i^{-1}\mathbf{Z}_i + \sigma_i^{-2}\mathbf{B}_i^{-1} \end{bmatrix} \begin{bmatrix} \hat{\mu}_i \\ \hat{\mathbf{g}}_i^N \end{bmatrix} = \begin{bmatrix} \mathbf{1}'\mathbf{Z}_i'\mathbf{R}_i^{-1}\mathbf{y}_i \\ \mathbf{Z}_i'\mathbf{R}_i^{-1}\mathbf{y}_i \end{bmatrix}$$

- Re-written as:

$$(\boldsymbol{\Psi}_i + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \sigma_i^{-2}\mathbf{B}_i^{-1} \end{bmatrix}) \begin{bmatrix} \hat{\mu}_i \\ \hat{\mathbf{g}}_i^N \end{bmatrix} = \boldsymbol{\Delta}_i$$

- Least-square part of the LHS of MME:

$$\boldsymbol{\Psi}_i = \begin{bmatrix} \mathbf{1}'\mathbf{Z}_i'\mathbf{R}_i^{-1}\mathbf{Z}_i\mathbf{1} & \mathbf{1}'\mathbf{R}_i^{-1}\mathbf{Z}_i \\ \mathbf{Z}_i'\mathbf{R}_i^{-1}\mathbf{1} & \mathbf{Z}_i'\mathbf{R}_i^{-1}\mathbf{Z}_i \end{bmatrix}$$

- Right-hand-side of the MME:

$$\boldsymbol{\Delta}_i = \begin{bmatrix} \mathbf{1}'\mathbf{Z}_i'\mathbf{R}_i^{-1}\mathbf{y}_i \\ \mathbf{Z}_i'\mathbf{R}_i^{-1}\mathbf{y}_i \end{bmatrix}$$

Mixed model equations of the SNP MACE model

$$\begin{bmatrix} \ddots & & \cdots & & \cdots \\ & \begin{bmatrix} \mathbf{1}'\mathbf{Z}_i'\mathbf{R}_i^{-1}\mathbf{Z}_i\mathbf{1} & \mathbf{1}'\mathbf{R}_i^{-1}\mathbf{Z}_i \\ \mathbf{Z}_i'\mathbf{R}_i^{-1}\mathbf{1} & \mathbf{Z}_i'\mathbf{R}_i^{-1}\mathbf{Z}_i \end{bmatrix} + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}^{ii} \end{bmatrix} & \cdots & \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Psi}_{ii^+} \end{bmatrix} + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}^{ii^+} \end{bmatrix} & \vdots \\ \vdots & \cdots & \begin{bmatrix} \mathbf{1}'\mathbf{Z}_{i^+}'\mathbf{R}_{i^+}^{-1}\mathbf{Z}_{i^+}\mathbf{1} & \mathbf{1}'\mathbf{R}_{i^+}^{-1}\mathbf{Z}_{i^+} \\ \mathbf{Z}_{i^+}'\mathbf{R}_{i^+}^{-1}\mathbf{1} & \mathbf{Z}_{i^+}'\mathbf{R}_{i^+}^{-1}\mathbf{Z}_{i^+} \end{bmatrix} + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}^{i^+i^+} \end{bmatrix} & \vdots \\ & & & & \ddots \end{bmatrix}$$

$$\mathrm{X}\begin{bmatrix} \cdots \\ \begin{bmatrix} \hat{\mu}_i \\ \hat{\mathbf{g}}_i \end{bmatrix} \\ \vdots \\ \begin{bmatrix} \hat{\mu}_{i^+} \\ \hat{\mathbf{g}}_{i^+} \end{bmatrix} \\ \cdots \end{bmatrix} = \begin{bmatrix} \cdots \\ \begin{bmatrix} \mathbf{1}'\mathbf{Z}_i'\mathbf{R}_i^{-1}\mathbf{y}_i \\ \mathbf{Z}_i'\mathbf{R}_i^{-1}\mathbf{y}_i \end{bmatrix} \\ \vdots \\ \begin{bmatrix} \mathbf{1}'\mathbf{Z}_{i^+}'\mathbf{R}_{i^+}^{-1}\mathbf{y}_{i^+} \\ \mathbf{Z}_{i^+}'\mathbf{R}_{i^+}^{-1}\mathbf{y}_{i^+} \end{bmatrix} \\ \cdots \end{bmatrix}$$

Residual covariance: $\boldsymbol{\Psi}_{ii^+} = (\mathbf{Z}_i'\mathbf{R}_i^{-\frac{1}{2}})(\mathbf{R}_{i^+}^{-\frac{1}{2}}\mathbf{Z}_{i^+})$

$\boldsymbol{\Psi}_{ii^+} = 0$ when countries use only national phenotypes for SNP effect estimation

Mixed model equations of the SNP MACE model (mu)

$$\begin{bmatrix} \ddots & & \cdots & & \cdots & \\ & \begin{bmatrix} 1'R_i^{-1}1 & 1'R_i^{-1}Z_i \\ Z_i'R_i^{-1}1 & Z_i'R_i^{-1}Z_i \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & G^{ii} \end{bmatrix} & \cdots & \begin{bmatrix} 0 & 0 \\ 0 & \Psi_{ii'} \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & G^{ii'} \end{bmatrix} & \vdots & \\ & & \ddots & \vdots & \vdots & \\ & & & \begin{bmatrix} 1'R_{i'}^{-1}1 & 1'R_{i'}^{-1}Z_{i'} \\ Z_{i'}'R_{i'}^{-1}1 & Z_{i'}'R_{i'}^{-1}Z_{i'} \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & G^{i'i'} \end{bmatrix} & \vdots & \\ & & & & \ddots & \end{bmatrix}$$

$$X \begin{bmatrix} \cdots \\ \begin{bmatrix} \hat{\mu}_i \\ \hat{\mathbf{g}}_i \end{bmatrix} \\ \vdots \\ \begin{bmatrix} \hat{\mu}_{i+} \\ \hat{\mathbf{g}}_{i+} \end{bmatrix} \\ \cdots \end{bmatrix} = \begin{bmatrix} \cdots \\ \begin{bmatrix} 1'R_i^{-1}y_i \\ Z_i'R_i^{-1}y_i \end{bmatrix} \\ \vdots \\ \begin{bmatrix} 1'R_{i+}^{-1}y_{i+} \\ Z_{i+}'R_{i+}^{-1}y_{i+} \end{bmatrix} \\ \cdots \end{bmatrix}$$

Method 2: Conversion of country SNP effects using the genomic relationship matrix: DGV

- Country uses a SNP BLUP model with its own SNP set
- Assumption: equal DGV for reference animals with both SNP sets
- DGV genomic relationship matrix for all reference animals is invertible:

$$\mathbf{G}_{rel\_c}^{-1} = (\mathbf{Z}_i^c{}' \mathbf{B}_c \mathbf{Z}_i^c)^{-1} = (\mathbf{Z}_i{}' \mathbf{B}_i \mathbf{Z}_i)^{-1} = \mathbf{G}_{rel\_i}^{-1}$$

- SNP marker cross-referencing (7 own SNPs, 4 common SNPs)

$$\mathbf{B}_{c,i} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

- SNP effects converted to the common set:

$$\mathbf{g}_i^c = \mathbf{B}_{c,i} \mathbf{Z}_i{}' \mathbf{G}_{rel\_i}^{-1} (\mathbf{Z}_i \mathbf{g}_i)$$

- Back conversion of MACE SNP effects to own SNP set:

$$\mathbf{g}_i = \mathbf{B}_{c,i}{}' \mathbf{Z}_c{}' \mathbf{G}_{rel\_c}^{-1} (\mathbf{Z}_c \mathbf{g}_i^c)$$