

Supplementary material: Genomic selection for genotype performance and stability using information on multiple traits and multiple environments

Jon Bancic · Ben Ovenden · Gregor Gorjanc · Daniel J. Tolhurst

Received: 07 Jul 2022/ Accepted: 9 Jan 2023

Key message The inclusion of multiple traits and multiple environments within a partially separable factor analytic approach for genomic selection provides breeders with an informative framework to utilise genotype by environment by trait interaction for efficient selection

D.J. Tolhurst ✉

The Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, Easter Bush, Midlothian United Kingdom
E-mail: D.J.Tolhurst@sms.ed.ac.uk

Sparse implementation of the SFA-LMM

This section extends the sparse implementation of the average information (AI) algorithm for the partially separable factor analytic linear mixed model (SFA-LMM). [Thompson et al. \(2003\)](#) introduced the sparse implementation for factor analytic linear mixed models as a more efficient approach to estimate the key variance parameters, that is the loadings, score variances and specific variances. This approach also provides a natural way to handle cases where some specific variances are zero. When all specific variances are zero, this leads to a (fully) reduced rank factor analytic linear mixed model ([Kirkpatrick and Meyer, 2004](#)).

Preliminaries

The SFA-LMM for \mathbf{y} , the ns -vector of phenotypic data on v genotypes, p environments and s traits, is:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\tau} + \mathbf{Z}_{\Lambda_{te}}\mathbf{f}_{te} + (\mathbf{I}_s \otimes \mathbf{Z})\boldsymbol{\delta}_{te} + \boldsymbol{\varepsilon}, \quad (1)$$

where $\mathbf{Z}_{\Lambda_{te}} = \boldsymbol{\Lambda}_t \otimes \mathbf{Z}[\boldsymbol{\Lambda}_e \otimes \mathbf{I}_v]$ is the $n \times vk_t k_e$ separable factor analytic design matrix, and $\boldsymbol{\varepsilon} = (\mathbf{I}_s \otimes \mathbf{Z})\mathbf{u}_n + (\mathbf{I}_s \otimes \mathbf{Z}_p)\mathbf{u}_p + \mathbf{e}$, with $\text{var}(\boldsymbol{\varepsilon}) = \mathbf{R}_\varepsilon$. The SFA-LMM in Equation 1 can be extended by reordering and partitioning the regression residuals as $\boldsymbol{\delta}_{te} = (\boldsymbol{\delta}_{te_1}^\top, \boldsymbol{\delta}_{te_2}^\top)^\top$, where $\boldsymbol{\delta}_{te_1}$ is a $vp_1 s_1$ -vector with no zero elements and $\boldsymbol{\delta}_{te_2}$ is a $v(p_1 s_2 + p_2 s)$ -vector with all zero elements, such that $p = p_1 + p_2$ and $s = s_1 + s_2$, with $p_1 \geq 1$ due to the constraints required during estimation (see Section X of the manuscript). Two simpler models can also be obtained:

1. When $p_1 = p$ and $s_1 = s$, no specific variances are zero. This model is the conventional SFA-LMM.
2. When $s_1 = 0$, all specific variances are effectively zero. This model is referred to as the reduced rank SFA-LMM.

The model considered below allows some specific variances to be non-zero and some to be zero. The SFA-LMM can therefore be written as:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\tau} + \mathbf{Z}_{\Lambda_{te}}\mathbf{f}_{te} + \mathbf{Z}_1\boldsymbol{\delta}_{te_1} + \mathbf{Z}_2\boldsymbol{\delta}_{te_2} + \boldsymbol{\varepsilon}, \quad (2)$$

where $[\mathbf{Z}_1 \mathbf{Z}_2]$ is partitioned conformably with $\boldsymbol{\delta}_{te}$.

It is assumed that $\boldsymbol{\delta}_{te_2} = \mathbf{0}$, and that:

$$\begin{bmatrix} \mathbf{f}_{te} \\ \boldsymbol{\delta}_{te_1} \end{bmatrix} \sim N \left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \mathbf{D}_t \otimes \mathbf{D}_e & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Psi}_{t_1} \otimes \boldsymbol{\Psi}_{e_1} \end{bmatrix} \otimes \mathbf{G}_g \right).$$

In this model, the genotype scores and regression residuals are assumed to be independent. [Thompson et al. \(2003\)](#) also present an equivalent formulation where the random effects are assumed to be dependent. This formulation will be extended for the SFA-LMM in a subsequent paper.

The mixed model equations for the SFA-LMM in Equation 2 are given by:

$$\begin{bmatrix} \mathbf{X}^\top \mathbf{R}_\varepsilon^{-1} \mathbf{X} & & \mathbf{X}^\top \mathbf{R}_\varepsilon^{-1} \mathbf{Z}_{\Lambda_{te}} & & \mathbf{X}^\top \mathbf{R}_\varepsilon^{-1} \mathbf{Z}_1 \\ \mathbf{Z}_{\Lambda_{te}}^\top \mathbf{R}_\varepsilon^{-1} \mathbf{X} & \mathbf{Z}_{\Lambda_{te}}^\top \mathbf{R}_\varepsilon^{-1} \mathbf{Z}_{\Lambda_{te}} + \mathbf{D}_t^{-1} \otimes \mathbf{D}_e^{-1} \otimes \mathbf{G}_g^{-1} & & & \mathbf{Z}_{\Lambda_{te}}^\top \mathbf{R}_\varepsilon^{-1} \mathbf{Z}_1 \\ \mathbf{Z}_1^\top \mathbf{R}_\varepsilon^{-1} \mathbf{X} & & \mathbf{Z}_1^\top \mathbf{R}_\varepsilon^{-1} \mathbf{Z}_{\Lambda_{te}} & & \mathbf{Z}_1^\top \mathbf{R}_\varepsilon^{-1} \mathbf{Z}_1 + \boldsymbol{\Psi}_{t_1}^{-1} \otimes \boldsymbol{\Psi}_{e_1}^{-1} \otimes \mathbf{G}_g^{-1} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\tau}} \\ \hat{\mathbf{f}}_{te} \\ \hat{\boldsymbol{\delta}}_{te_1} \end{bmatrix} = \begin{bmatrix} \mathbf{X}^\top \mathbf{R}_\varepsilon^{-1} \mathbf{y} \\ \mathbf{Z}_{\Lambda_{te}}^\top \mathbf{R}_\varepsilon^{-1} \mathbf{y} \\ \mathbf{Z}_1^\top \mathbf{R}_\varepsilon^{-1} \mathbf{y} \end{bmatrix}. \quad (3)$$

Prediction of genotype scores and regression residuals

Following [Smith et al. \(2019\)](#), Equations 2 and 3 can be written as:

$$\mathbf{y} = \mathbf{W}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad \text{and} \quad \mathbf{C}\tilde{\boldsymbol{\beta}} = \mathbf{W}^\top \mathbf{R}_\varepsilon^{-1} \mathbf{y}, \quad (4)$$

where $\mathbf{W} = [\mathbf{X} \mathbf{Z}_{\Lambda_{te}} \mathbf{Z}_1]$ and $\mathbf{C} = \mathbf{W}^\top \mathbf{R}_\varepsilon^{-1} \mathbf{W} + \mathbf{G}_c^{-1}$, with:

$$\tilde{\boldsymbol{\beta}} = \begin{pmatrix} \hat{\boldsymbol{\tau}} \\ \hat{\mathbf{f}}_{te} \\ \hat{\boldsymbol{\delta}}_{te_1} \end{pmatrix} \quad \text{and} \quad \mathbf{G}_c^{-1} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_t^{-1} \otimes \mathbf{D}_e^{-1} \otimes \mathbf{G}_g^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \boldsymbol{\Psi}_{t_1}^{-1} \otimes \boldsymbol{\Psi}_{e_1}^{-1} \otimes \mathbf{G}_g^{-1} \end{bmatrix}.$$

The BLUPs of the key random effects are obtained via absorption of \mathbf{C} onto $\mathbf{y}^\top \mathbf{R}_\varepsilon^{-1} \mathbf{y}$, which gives:

$$\tilde{\mathbf{f}}_{te} = [\mathbf{D}_t \otimes \mathbf{D}_e \otimes \mathbf{G}_g] \mathbf{Z}_{\Lambda_{te}}^\top \mathbf{P} \mathbf{y} \quad \text{and} \quad \tilde{\boldsymbol{\delta}}_{te_1} = [\boldsymbol{\Psi}_{t_1} \otimes \boldsymbol{\Psi}_{e_1} \otimes \mathbf{G}_g] \mathbf{Z}_1^\top \mathbf{P} \mathbf{y},$$

where $\mathbf{P} = \mathbf{R}_\varepsilon^{-1} - \mathbf{R}_\varepsilon^{-1} \mathbf{W} \mathbf{C}^{-1} \mathbf{W}^\top \mathbf{R}_\varepsilon^{-1}$ is the $ns \times ns$ residual sum of squares matrix. Absorption of \mathbf{C} also produces the prediction error variance matrices, with:

$$\text{var}(\tilde{\mathbf{f}}_{te} - \mathbf{f}_{te}) = \mathbf{C}^{\tilde{\mathbf{f}}_{te}} \tilde{\mathbf{f}}_{te} \quad \text{and} \quad \text{var}(\tilde{\boldsymbol{\delta}}_{te_1} - \boldsymbol{\delta}_{te_1}) = \mathbf{C}^{\tilde{\boldsymbol{\delta}}_{te_1}} \tilde{\boldsymbol{\delta}}_{te_1},$$

where $\mathbf{C}^{\tilde{\mathbf{f}}_{te}} = [\mathbf{D}_t \otimes \mathbf{D}_e \otimes \mathbf{G}_g] (\mathbf{I}_{vk_t k_e} - \mathbf{Z}_{\Lambda_{te}}^\top \mathbf{P} \mathbf{Z}_{\Lambda_{te}} [\mathbf{D}_t \otimes \mathbf{D}_e \otimes \mathbf{G}_g])$ and $\mathbf{C}^{\tilde{\boldsymbol{\delta}}_{te_1}} = [\boldsymbol{\Psi}_{t_1} \otimes \boldsymbol{\Psi}_{e_1} \otimes \mathbf{G}_g] (\mathbf{I}_{vps} - \mathbf{Z}_1^\top \mathbf{P} \mathbf{Z}_1 [\boldsymbol{\Psi}_{t_1} \otimes \boldsymbol{\Psi}_{e_1} \otimes \mathbf{G}_g])$. These matrices are equivalent to the diagonal blocks in \mathbf{C}^{-1} corresponding to $\tilde{\mathbf{f}}_{te}$ and $\tilde{\boldsymbol{\delta}}_{te_1}$, respectively.

Lastly, note that the components in the equations above assume the variance parameters are known. The variance parameters are unknown, however, so they must be estimated. The resulting predictions are therefore referred to as EBLUPs.

Estimation of loadings and specific variances

The REML estimates of the key variance parameters are obtained by maximising the residual log-likelihood, which is given by:

$$\begin{aligned} l(\mathbf{y}_2) &= -\frac{1}{2} (\log |\mathbf{H}| + \log |\mathbf{X}^\top \mathbf{H}^{-1} \mathbf{X}| + \mathbf{y}^\top \mathbf{P} \mathbf{y}) \\ &= -\frac{1}{2} (\log |\mathbf{H}| + \log |\mathbf{X}^\top \mathbf{H}^{-1} \mathbf{X}| + \mathbf{y}_r^\top \mathbf{H}^{-1} \mathbf{y}_r), \end{aligned} \quad (5)$$

where $\mathbf{y}_2 = \mathbf{L}_2^\top \mathbf{y}$ such that $\mathbf{L}_2^\top \mathbf{X} = \mathbf{0}$ and $\mathbf{y}_r = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\tau}}$, with $\text{var}(\mathbf{y}_r) = \mathbf{H}$ ([Verbyla, 1990](#)). In particular, the REML estimates are obtained by solving a set of (score) equations, which are given by:

$$s(\boldsymbol{\kappa}) = \frac{\partial l(\mathbf{y}_2)}{\partial \boldsymbol{\kappa}^\top} = \mathbf{0}, \quad (6)$$

where $\boldsymbol{\kappa}$ is the b -vector of variance parameters in the IFA-LMM. The score equations were originally solved using numerical approaches based on the observed or

expected information. The AI algorithm is based on the average information, and computes updates as:

$$\boldsymbol{\kappa}^{(m+1)} = \boldsymbol{\kappa}^{(m)} + \left[\mathcal{I}_{\mathbf{a}}^{(m)} \right]^{-1} \mathbf{s}(\boldsymbol{\kappa}^{(m)}), \quad (7)$$

where $\boldsymbol{\kappa}^{(m)}$ is the b -vector of variance parameters, $\mathcal{I}_{\mathbf{a}}^{(m)}$ is the AI matrix and $\mathbf{s}(\boldsymbol{\kappa}^{(m)})$ is the score equations for the m^{th} iteration. Note that the superscript “ m ” is removed in the following for brevity. The score equation for κ_i is given by (Gilmour et al., 1995):

$$s(\kappa_i) = -\frac{1}{2} \left[\text{tr}(\mathbf{P}\dot{\mathbf{H}}_i) - \mathbf{y}^\top \mathbf{P}\mathbf{q}_i \right], \quad (8)$$

where $\dot{\mathbf{H}}_i = \frac{\partial \mathbf{H}}{\partial \kappa_i}$ and \mathbf{q}_i is the working variate for the i^{th} variance parameter, which is given by:

$$\mathbf{q}_i = \dot{\mathbf{H}}_i \mathbf{P} \mathbf{y}. \quad (9)$$

Next, let the vectors of key variance parameters in the SFA-LMM be denoted by:

$$\boldsymbol{\lambda}_{\mathbf{a}} = \text{vect}(\boldsymbol{\Lambda}_{\mathbf{a}}), \quad \mathbf{d}_{\mathbf{a}} = \text{diag}(\mathbf{D}_{\mathbf{a}}) \text{ and } \boldsymbol{\psi}_{\mathbf{a}_1} = \text{diag}(\boldsymbol{\Psi}_{\mathbf{a}_1}),$$

where $a = t$ and e for traits and environments.

The working variates for the key variance parameters are given by:

$$\mathbf{q}_i = \begin{cases} \left[\begin{array}{l} (\boldsymbol{\Lambda}_t \mathbf{D}_t \boldsymbol{\Lambda}_t^\top + \dot{\boldsymbol{\Lambda}}_t \mathbf{D}_t \boldsymbol{\Lambda}_t^\top) \otimes \mathbf{Z}(\boldsymbol{\Lambda}_e \mathbf{D}_e \boldsymbol{\Lambda}_e^\top \otimes \mathbf{G}_g) \mathbf{Z}^\top \\ \boldsymbol{\Lambda}_t \mathbf{D}_t \boldsymbol{\Lambda}_t^\top \otimes \mathbf{Z}[(\boldsymbol{\Lambda}_e \mathbf{D}_e \boldsymbol{\Lambda}_e^\top + \dot{\boldsymbol{\Lambda}}_e \mathbf{D}_e \boldsymbol{\Lambda}_e^\top) \otimes \mathbf{G}_g] \mathbf{Z}^\top \end{array} \right] \mathbf{P} \mathbf{y} & \kappa_i \in \boldsymbol{\lambda}_t \\ \left[\begin{array}{l} \boldsymbol{\Lambda}_{\Lambda_{te}} \left[\begin{array}{l} \dot{\mathbf{D}}_t \otimes \mathbf{D}_e \otimes \mathbf{G}_g \\ \mathbf{D}_t \otimes \dot{\mathbf{D}}_e \otimes \mathbf{G}_g \end{array} \right] \mathbf{Z}_{\Lambda_{te}}^\top \\ \mathbf{Z}_1 \left[\begin{array}{l} \dot{\boldsymbol{\Psi}}_{t_1} \otimes \boldsymbol{\Psi}_{e_1} \otimes \mathbf{G}_g \\ \boldsymbol{\Psi}_{t_1} \otimes \dot{\boldsymbol{\Psi}}_{e_1} \otimes \mathbf{G}_g \end{array} \right] \mathbf{Z}_1^\top \end{array} \right] \mathbf{P} \mathbf{y} & \kappa_i \in \mathbf{d}_t \\ & \kappa_i \in \mathbf{d}_e \\ \mathbf{Z}_1 \left[\begin{array}{l} \dot{\boldsymbol{\Psi}}_{t_1} \otimes \boldsymbol{\Psi}_{e_1} \otimes \mathbf{G}_g \\ \boldsymbol{\Psi}_{t_1} \otimes \dot{\boldsymbol{\Psi}}_{e_1} \otimes \mathbf{G}_g \end{array} \right] \mathbf{Z}_1^\top \mathbf{P} \mathbf{y} & \kappa_i \in \boldsymbol{\psi}_{t_1} \\ \mathbf{Z}_1 \left[\begin{array}{l} \boldsymbol{\Psi}_{t_1} \otimes \dot{\boldsymbol{\Psi}}_{e_1} \otimes \mathbf{G}_g \\ \dot{\boldsymbol{\Psi}}_{t_1} \otimes \boldsymbol{\Psi}_{e_1} \otimes \mathbf{G}_g \end{array} \right] \mathbf{Z}_1^\top \mathbf{P} \mathbf{y} & \kappa_i \in \boldsymbol{\psi}_{e_1} \end{cases}$$

where $\dot{\boldsymbol{\Lambda}}_{\mathbf{a}_i} = \frac{\partial \boldsymbol{\Lambda}_{\mathbf{a}}}{\partial \lambda_{a_i}}$, $\dot{\mathbf{D}}_{\mathbf{a}_i} = \frac{\partial \mathbf{D}_{\mathbf{a}}}{\partial d_{a_i}}$ and $\dot{\boldsymbol{\Psi}}_{\mathbf{a}_1} = \frac{\partial \boldsymbol{\Psi}_{\mathbf{a}_1}}{\partial \psi_{a_1}}$, with $a = t$ or e . The working variates for the specific variances can be simplified since the regression residuals are independent across trait by environment combinations. It therefore follows that:

$$\mathbf{q}_i = \begin{cases} \mathbf{Z}_{1..i} \left[\boldsymbol{\Psi}_{e_1} \otimes \mathbf{G}_g \right] \mathbf{Z}_{1..i}^\top \mathbf{P}_i \mathbf{y} & \kappa_i \in \boldsymbol{\psi}_{t_1} \\ \mathbf{Z}_{1..i} \left[\boldsymbol{\Psi}_{t_1} \otimes \mathbf{G}_g \right] \mathbf{Z}_{1..i}^\top \mathbf{P}_i \mathbf{y} & \kappa_i \in \boldsymbol{\psi}_{e_1} \end{cases}$$

where $\mathbf{Z}_{1..i}$ is the $n_1 \times vp_1$ design matrix for the i^{th} trait with $n_1 = \sum_{i=1}^{p_1} n_i$ or $\mathbf{Z}_{1..i}$ is the $n_i s_1 \times vs_1$ design matrix for the i^{th} environment, \mathbf{P}_i are the corresponding n_1 or $n_i s_1$ rows in \mathbf{P} and $\mathbf{Z}_{1..i}$ are the corresponding vp_1 or vs_1 columns in \mathbf{Z}_1 .

The trace term in Equation 8 is then given by:

$$\text{tr}(\mathbf{P}\dot{\mathbf{H}}_i) = \begin{cases} 2 \text{tr} \left[\left(\dot{\boldsymbol{\Lambda}}_t^\top \otimes [\boldsymbol{\Lambda}_e^\top \otimes \mathbf{I}_v] \mathbf{Z}^\top \right) \mathbf{R}_e^{-1} \mathbf{W} \mathbf{C}^{\cdot \tilde{f}_{te}} \right] & \kappa_i \in \boldsymbol{\lambda}_t \\ 2 \text{tr} \left[\left(\dot{\boldsymbol{\Lambda}}_t^\top \otimes [\dot{\boldsymbol{\Lambda}}_e^\top \otimes \mathbf{I}_v] \mathbf{Z}^\top \right) \mathbf{R}_e^{-1} \mathbf{W} \mathbf{C}^{\cdot \tilde{f}_{te}} \right] & \kappa_i \in \boldsymbol{\lambda}_e \\ \text{tr} \left[\left(\dot{\mathbf{D}}_t \mathbf{D}_t^{-1} \otimes \mathbf{D}_e \otimes \mathbf{I}_v \right) \mathbf{Z}_{\Lambda_{te}}^\top \mathbf{R}_e^{-1} \mathbf{W} \mathbf{C}^{\cdot \tilde{f}_{te}} \right] & \kappa_i \in \mathbf{d}_t \\ \text{tr} \left[\left(\mathbf{D}_t \otimes \dot{\mathbf{D}}_e \mathbf{D}_e^{-1} \otimes \mathbf{I}_v \right) \mathbf{Z}_{\Lambda_{te}}^\top \mathbf{R}_e^{-1} \mathbf{W} \mathbf{C}^{\cdot \tilde{f}_{te}} \right] & \kappa_i \in \mathbf{d}_e \\ \text{tr} \left[\left(\dot{\boldsymbol{\Psi}}_{t_1} \boldsymbol{\Psi}_{t_1}^{-1} \otimes \boldsymbol{\Psi}_{e_1} \otimes \mathbf{I}_v \right) \mathbf{Z}_1^\top \mathbf{R}_e^{-1} \mathbf{W} \mathbf{C}^{\cdot \tilde{\delta}_{te_1}} \right] & \kappa_i \in \boldsymbol{\psi}_{t_1} \\ \text{tr} \left[\left(\boldsymbol{\Psi}_{t_1} \otimes \dot{\boldsymbol{\Psi}}_{e_1} \boldsymbol{\Psi}_{e_1}^{-1} \otimes \mathbf{I}_v \right) \mathbf{Z}_1^\top \mathbf{R}_e^{-1} \mathbf{W} \mathbf{C}^{\cdot \tilde{\delta}_{te_1}} \right] & \kappa_i \in \boldsymbol{\psi}_{e_1} \end{cases}$$

where $\mathbf{C}^{\cdot \tilde{f}_{te}}$ is the column in the inverse coefficient matrix corresponding to \tilde{f}_{te} and $\mathbf{C}^{\cdot \tilde{\delta}_{te_1}}$ is the column corresponding to $\tilde{\delta}_{te_1}$.

The trace terms for the specific variances can be further simplified as:

$$\text{tr}(\mathbf{P}\dot{\mathbf{H}}_i) = \begin{cases} \frac{1}{\psi_{t_1}} \text{tr} \left[\left(\boldsymbol{\Psi}_{e_1} \otimes \mathbf{I}_v \right) \mathbf{Z}_{1..i}^\top \mathbf{R}_{\boldsymbol{\epsilon}_i}^{-1} \mathbf{W}_i \cdot \mathbf{C}^{\cdot \tilde{\delta}_{te_1}} \right] & \kappa_i \in \boldsymbol{\psi}_{t_1} \\ \frac{1}{\psi_{e_1}} \text{tr} \left[\left(\boldsymbol{\Psi}_{t_1} \otimes \mathbf{I}_v \right) \mathbf{Z}_{1..i}^\top \mathbf{R}_{\boldsymbol{\epsilon}_i}^{-1} \mathbf{W}_i \cdot \mathbf{C}^{\cdot \tilde{\delta}_{te_1}} \right] & \kappa_i \in \boldsymbol{\psi}_{e_1} \end{cases}$$

where $\mathbf{R}_{\boldsymbol{\epsilon}_i}$ is the i^{th} diagonal block in $\mathbf{R}_{\boldsymbol{\epsilon}}$, \mathbf{W}_i are the n_1 or $n_i s_1$ rows in \mathbf{W} corresponding to the i^{th} trait or environment and $\mathbf{C}^{\cdot \tilde{\delta}_{te_1}}$ are the vp_1 or vs_1 columns in \mathbf{C} corresponding to $\tilde{\delta}_{te_1}$. The trace terms avoid working with the dense genomic relationship matrix. When \mathbf{G}_g is not prohibitively large, the trace term for the specific variances can also be computed as:

$$\text{tr}(\mathbf{P}\dot{\mathbf{H}}_i) = \begin{cases} \frac{v}{\psi_{t_1}} \text{tr}(\boldsymbol{\Psi}_{e_1}) - \text{tr} \left[\left(\boldsymbol{\Psi}_{e_1} \otimes \mathbf{G}_g \right) \mathbf{C}^{\boldsymbol{\eta}_i \boldsymbol{\eta}_i} \right] & \kappa_i \in \boldsymbol{\psi}_{t_1} \\ \frac{v}{\psi_{e_1}} \text{tr}(\boldsymbol{\Psi}_{t_1}) - \text{tr} \left[\left(\boldsymbol{\Psi}_{t_1} \otimes \mathbf{G}_g \right) \mathbf{C}^{\boldsymbol{\eta}_i \boldsymbol{\eta}_i} \right] & \kappa_i \in \boldsymbol{\psi}_{e_1} \end{cases}$$

where $\mathbf{C}^{\boldsymbol{\eta}_i \boldsymbol{\eta}_i}$ is the prediction error variance matrix of $\boldsymbol{\eta}_i = \mathbf{Z}_{1..i}^\top \mathbf{P}_i \mathbf{y}$.

Lastly, the AI matrix in Equation 7 is given by:

$$\mathcal{I}_{\mathbf{a}} = \frac{1}{2} \mathbf{Q}^\top \mathbf{P} \mathbf{Q}, \quad (10)$$

where $\mathbf{Q} = [\mathbf{q}_1 \ \mathbf{q}_2 \ \dots \ \mathbf{q}_b]$ is the $ns \times b$ matrix of working variates across all b variance parameters. The AI matrix is obtained via absorption of \mathbf{C} onto $\mathbf{Q}^\top \mathbf{R}_{\boldsymbol{\epsilon}}^{-1} \mathbf{Q}$ (Smith, 1999).

References

- Gilmour AR, Thompson R, Cullis BR (1995) Average information reml: An efficient algorithm for variance parameter estimation in linear mixed models. *Biometrics* 51:1440–1450, URL <http://doi.org/10.2307/2533274>
- Kirkpatrick M, Meyer K (2004) Direct estimation of genetic principal components: Simplified analysis of complex phenotypes. *Genetics* 168:2295–2306, URL <http://doi.org/10.1534/genetics.104.029181>
- Smith AB (1999) Multiplicative mixed models for the analysis of multi-environment trial data. PhD thesis, University of Adelaide, URL <http://hdl.handle.net/2440/19539>
- Smith AB, Borg LM, Gogel BJ, Cullis BR (2019) Estimation of factor analytic mixed models for the analysis of multi-treatment multi-environment trial data. *International Biometric Society* 24:573–588, URL <http://doi.org/10.1007/s13253-019-00362-6>
- Thompson R, Cullis BR, Smith AB, Gilmour AR (2003) A sparse implementation of the average information algorithm for factor analytic and reduced rank variance models. *Australian and New Zealand Journal of Statistics* 45:445–459, URL <http://doi.org/10.1111/1467-842X.00297>
- Verbyla AP (1990) A Conditional Derivation of Residual Maximum Likelihood. *Australian Journal of Statistics* 32:227–230, URL <http://dx.doi.org/10.1111/j.1467-842X.1990.tb01015.x>