Numerically Efficient Fully Orthogonalized Single-step SNP-BLUP

Matti Taskinen, Esa Mäntysaari, Ismo Strandén
Natural Resources Institute Finland (Luke)

12.2.2018 WCGALP 2018 Auckland NZ

Luke
NATURAL RESOURCES INSTITUTE FINLAND

Challenge

Computations of large animal breeding evaluations become **numerially challenging** when number of genotyped animals increases.

Background: Single-step Genomic BLUP (ssGBLUP)

- A **mixed model equation** (MME).
- Combines **pedigree** ($\mathbf{A}$) and **genomic marker** relationship information ($\mathbf{G}_g$) through Single-step relationship matrix $\mathbf{H}$.
- Requires inversion of **full genomic relationship matrix $\mathbf{G}_g$**.
- Inversion $\mathbf{G}_g^{-1}$ becomes a bottleneck when number of genotyped increases.

Luke NATURAL RESOURCES INSTITUTE FINLAND

Background: Single-step Genomic BLUP (ssGBLUP)

- A **mixed model equation** (MME).
- Combines **pedigree** ($\mathbf{A}$) and **genomic marker** relationship information ($\mathbf{G}_g$) through Single-step relationship matrix $\mathbf{H}$.
- Requires inversion of **full genomic relationship matrix** $\mathbf{G}_g$.
- Inversion $\mathbf{G}_g^{-1}$ becomes a bottleneck when number of genotyped increases.

Single-step relationship matrix $\mathbf{H}$

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

Pedigree relationship matrix:

$$\mathbf{A}^{-1} = \begin{bmatrix} \mathbf{A}^{11} & \mathbf{A}^{12} \\ \mathbf{A}^{21} & \mathbf{A}^{22} \end{bmatrix}$$

(1 = non-genotyped)
(2 = genotyped)

Adjusted genomic relationship matrix:

$$\mathbf{G}_w = (1-w)\mathbf{G}_g + w\mathbf{A}_{22}$$

Genomic relationship matrix:

$$\mathbf{G}_g = \mathbf{Z}_m \mathbf{Z}_m'$$

$\mathbf{Z}_m$ is centered and scaled marker matrix

Background: Single-step Genomic BLUP (ssGBLUP)

- A **mixed model equation** (MME).
- Combines **pedigree** ($\mathbf{A}$) and **genomic marker** relationship information ($\mathbf{G}_g$) through Single-step relationship matrix $\mathbf{H}$.
- Requires inversion of **full genomic relationship matrix $\mathbf{G}_g$**.
- Inversion $\mathbf{G}_g^{-1}$ becomes a bottleneck when number of genotyped increases.

Single-step relationship matrix $\mathbf{H}$

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

Pedigree relationship matrix:

$$\mathbf{A}^{-1} = \begin{bmatrix} \mathbf{A}^{11} & \mathbf{A}^{12} \\ \mathbf{A}^{21} & \mathbf{A}^{22} \end{bmatrix}$$

(1 = non-genotyped)
(2 = genotyped)

Adjusted genomic relationship matrix:

$$\mathbf{G}_w = (1-w)\mathbf{G}_g + w\mathbf{A}_{22}$$

Genomic relationship matrix:

$$\mathbf{G}_g = \mathbf{Z}_m\mathbf{Z}_m'$$

$\mathbf{Z}_m$ is centered and scaled marker matrix

Luke NATURAL RESOURCES INSTITUTE FINLAND

Background: Single-step Genomic BLUP (ssGBLUP)

- A **mixed model equation** (MME).
- Combines **pedigree** ($\mathbf{A}$) and **genomic marker** relationship information ($\mathbf{G}_g$) through Single-step relationship matrix $\mathbf{H}$.
- Requires inversion of **full genomic relationship matrix $\mathbf{G}_g$**.
- Inversion $\mathbf{G}_g^{-1}$ becomes a bottleneck when number of genotyped increases.

Single-step relationship matrix $\mathbf{H}$

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

Pedigree relationship matrix:

$$\mathbf{A}^{-1} = \begin{bmatrix} \mathbf{A}^{11} & \mathbf{A}^{12} \\ \mathbf{A}^{21} & \mathbf{A}^{22} \end{bmatrix} \quad \begin{array}{l} (1 = \text{non-genotyped}) \\ (2 = \text{genotyped}) \end{array}$$

Adjusted genomic relationship matrix:

$$\mathbf{G}_w = (1 - w)\mathbf{G}_g + w\mathbf{A}_{22}$$

Genomic relationship matrix:

$$\mathbf{G}_g = \mathbf{Z}_m \mathbf{Z}_m'$$

$\mathbf{Z}_m$ is centered and scaled marker matrix

Previously: Linearly Equivalent Mixed Model Equations

[1] M. Taskinen, E. A. Mäntysaari, and I. Strandén. "Single-step SNP-BLUP with on-the-fly imputed genotypes and residual polygenic effects". In: *Genet. Sel. Evol.* 49.1 (2017), p. 15.

- **Linearly equivalent MMEs** [1]: decomposition of relationship matrix $\mathbf{G}$ of random effects $\hat{\mathbf{u}}$ as

$$\mathbf{G} = \mathbf{M}\widetilde{\mathbf{G}}\mathbf{M}'$$

- Alternative MME: parts of decomposition attached to new model matrix $\widetilde{\mathbf{Z}} = \mathbf{Z}\mathbf{M}$ and new set of random effects $\widetilde{\mathbf{u}}$ related through $\widetilde{\mathbf{G}}$.
- Original MME solved from $\hat{\mathbf{u}} = \mathbf{M}\widetilde{\mathbf{u}}$.
- Inversion $\mathbf{G}^{-1}$ avoided if $\widetilde{\mathbf{G}} = \mathbf{I}$, i.e. $\widetilde{\mathbf{u}}$ **"orthogonal"**.
- GBLUP $\Leftrightarrow$ SNP-BLUP: $\widetilde{\mathbf{u}}_m$ are **marker effects**.

Previously: Linearly Equivalent Mixed Model Equations

[1] M. Taskinen, E. A. Mäntysaari, and I. Strandén. "Single-step SNP-BLUP with on-the-fly imputed genotypes and residual polygenic effects". In: *Genet. Sel. Evol.* 49.1 (2017), p. 15.

- **Linearly equivalent MMEs** [1]: decomposition of relationship matrix $\mathbf{G}$ of random effects $\hat{\mathbf{u}}$ as

$$\mathbf{G} = \mathbf{M}\widetilde{\mathbf{G}}\mathbf{M}'$$

- Alternative MME: parts of decomposition attached to new model matrix $\widetilde{\mathbf{Z}} = \mathbf{Z}\mathbf{M}$ and new set of random effects $\tilde{\mathbf{u}}$ related through $\widetilde{\mathbf{G}}$.
- Original MME solved from $\hat{\mathbf{u}} = \mathbf{M}\tilde{\mathbf{u}}$.
- Inversion $\mathbf{G}^{-1}$ avoided if $\widetilde{\mathbf{G}} = \mathbf{I}$, i.e. $\tilde{\mathbf{u}}$ **"orthogonal"**.
- GBLUP ⇔ SNP-BLUP: $\tilde{\mathbf{u}}_m$ are **marker effects.**

Original MME

$$\begin{bmatrix}\hat{\mathbf{b}}\\ \hat{\mathbf{u}}\end{bmatrix} = \begin{bmatrix}\mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}\\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}+\mathbf{G}^{-1}\end{bmatrix}^{-1}\begin{bmatrix}\mathbf{X}'\mathbf{R}^{-1}\mathbf{y}\\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y}\end{bmatrix}$$

Luke NATURAL RESOURCES INSTITUTE FINLAND

Previously: Linearly Equivalent Mixed Model Equations

[1] M. Taskinen, E. A. Mäntysaari, and I. Strandén. “Single-step SNP-BLUP with on-the-fly imputed genotypes and residual polygenic effects”. In: *Genet. Sel. Evol.* 49.1 (2017), p. 15.

- **Linearly equivalent MMEs** [1]: decomposition of relationship matrix $\mathbf{G}$ of random effects $\hat{\mathbf{u}}$ as

$$\mathbf{G} = \mathbf{M}\widetilde{\mathbf{G}}\mathbf{M}'$$

- Alternative MME: parts of decomposition attached to new model matrix $\widetilde{\mathbf{Z}} = \mathbf{Z}\mathbf{M}$ and new set of random effects $\widetilde{\mathbf{u}}$ related through $\widetilde{\mathbf{G}}$.
- Original MME solved from $\hat{\mathbf{u}} = \mathbf{M}\widetilde{\mathbf{u}}$.
- Inversion $\mathbf{G}^{-1}$ avoided if $\widetilde{\mathbf{G}} = \mathbf{I}$, i.e. $\widetilde{\mathbf{u}}$ **“orthogonal”**.
- GBLUP ⇔ SNP-BLUP: $\widetilde{\mathbf{u}}_m$ are **marker effects**.

Original MME

$$\begin{bmatrix}\hat{\mathbf{b}}\\ \hat{\mathbf{u}}\end{bmatrix} = \begin{bmatrix}\mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}\\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}+\mathbf{G}^{-1}\end{bmatrix}^{-1}\begin{bmatrix}\mathbf{X}'\mathbf{R}^{-1}\mathbf{y}\\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y}\end{bmatrix}$$

Linearly equivalent MME

$$\begin{bmatrix}\hat{\mathbf{b}}\\ \widetilde{\mathbf{u}}\end{bmatrix} = \begin{bmatrix}\mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\widetilde{\mathbf{Z}}\\ \widetilde{\mathbf{Z}}'\mathbf{R}^{-1}\mathbf{X} & \widetilde{\mathbf{Z}}'\mathbf{R}^{-1}\widetilde{\mathbf{Z}}+\widetilde{\mathbf{G}}^{-1}\end{bmatrix}^{-1}\begin{bmatrix}\mathbf{X}'\mathbf{R}^{-1}\mathbf{y}\\ \widetilde{\mathbf{Z}}'\mathbf{R}^{-1}\mathbf{y}\end{bmatrix}$$

Calculation of original effects:

$$\begin{bmatrix}\hat{\mathbf{b}}\\ \hat{\mathbf{u}}\end{bmatrix} = \begin{bmatrix}\mathbf{I}_b & \mathbf{0}\\ \mathbf{0} & \mathbf{M}\end{bmatrix}\begin{bmatrix}\hat{\mathbf{b}}\\ \widetilde{\mathbf{u}}\end{bmatrix}$$

NATURAL RESOURCES INSTITUTE FINLAND

Previously: Linearly Equivalent Mixed Model Equations

[1] M. Taskinen, E. A. Mäntysaari, and I. Strandén. "Single-step SNP-BLUP with on-the-fly imputed genotypes and residual polygenic effects". In: *Genet. Sel. Evol.* 49.1 (2017), p. 15.

- **Linearly equivalent MMEs** [1]: decomposition of relationship matrix $\mathbf{G}$ of random effects $\hat{\mathbf{u}}$ as

$$\mathbf{G} = \mathbf{M}\tilde{\mathbf{G}}\mathbf{M}'$$

- Alternative MME: parts of decomposition attached to new model matrix $\tilde{\mathbf{Z}} = \mathbf{Z}\mathbf{M}$ and new set of random effects $\tilde{\mathbf{u}}$ related through $\tilde{\mathbf{G}}$.
- Original MME solved from $\hat{\mathbf{u}} = \mathbf{M}\tilde{\mathbf{u}}$.
- Inversion $\mathbf{G}^{-1}$ avoided if $\tilde{\mathbf{G}} = \mathbf{I}$, i.e. $\tilde{\mathbf{u}}$ **"orthogonal"**.
- GBLUP ⇔ SNP-BLUP: $\tilde{\mathbf{u}}_m$ are **marker effects**.

Original MME

$$\begin{bmatrix}\hat{\mathbf{b}}\\ \hat{\mathbf{u}}\end{bmatrix} = \begin{bmatrix}\mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}\\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}+\mathbf{G}^{-1}\end{bmatrix}^{-1}\begin{bmatrix}\mathbf{X}'\mathbf{R}^{-1}\mathbf{y}\\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y}\end{bmatrix}$$

Linearly equivalent MME

$$\begin{bmatrix}\hat{\mathbf{b}}\\ \tilde{\mathbf{u}}\end{bmatrix} = \begin{bmatrix}\mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\tilde{\mathbf{Z}}\\ \tilde{\mathbf{Z}}'\mathbf{R}^{-1}\mathbf{X} & \tilde{\mathbf{Z}}'\mathbf{R}^{-1}\tilde{\mathbf{Z}}+\tilde{\mathbf{G}}^{-1}\end{bmatrix}^{-1}\begin{bmatrix}\mathbf{X}'\mathbf{R}^{-1}\mathbf{y}\\ \tilde{\mathbf{Z}}'\mathbf{R}^{-1}\mathbf{y}\end{bmatrix}$$

Calculation of original effects:

$$\begin{bmatrix}\hat{\mathbf{b}}\\ \hat{\mathbf{u}}\end{bmatrix} = \begin{bmatrix}\mathbf{I}_b & \mathbf{0}\\ \mathbf{0} & \mathbf{M}\end{bmatrix}\begin{bmatrix}\hat{\mathbf{b}}\\ \tilde{\mathbf{u}}\end{bmatrix}$$

Previously: Linearly Equivalent Mixed Model Equations

[1] M. Taskinen, E. A. Mäntysaari, and I. Strandén. "Single-step SNP-BLUP with on-the-fly imputed genotypes and residual polygenic effects". In: *Genet. Sel. Evol.* 49.1 (2017), p. 15.

- **Linearly equivalent MMEs** [1]: decomposition of relationship matrix $\mathbf{G}$ of random effects $\hat{\mathbf{u}}$ as

$$\mathbf{G} = \mathbf{M}\widetilde{\mathbf{G}}\mathbf{M}'$$

- Alternative MME: parts of decomposition attached to new model matrix $\widetilde{\mathbf{Z}} = \mathbf{Z}\mathbf{M}$ and new set of random effects $\widetilde{\mathbf{u}}$ related through $\widetilde{\mathbf{G}}$.
- Original MME solved from $\hat{\mathbf{u}} = \mathbf{M}\widetilde{\mathbf{u}}$.
- Inversion $\mathbf{G}^{-1}$ avoided if $\widetilde{\mathbf{G}} = \mathbf{I}$, i.e. $\widetilde{\mathbf{u}}$ **"orthogonal"**.
- GBLUP $\Leftrightarrow$ SNP-BLUP: $\widetilde{\mathbf{u}}_m$ are **marker effects.**

Original MME

$$\begin{bmatrix}\hat{\mathbf{b}}\\ \hat{\mathbf{u}}\end{bmatrix} = \begin{bmatrix}\mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}\\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}+\mathbf{G}^{-1}\end{bmatrix}^{-1}\begin{bmatrix}\mathbf{X}'\mathbf{R}^{-1}\mathbf{y}\\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y}\end{bmatrix}$$

Linearly equivalent MME

$$\begin{bmatrix}\hat{\mathbf{b}}\\ \widetilde{\mathbf{u}}\end{bmatrix} = \begin{bmatrix}\mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\widetilde{\mathbf{Z}}\\ \widetilde{\mathbf{Z}}'\mathbf{R}^{-1}\mathbf{X} & \widetilde{\mathbf{Z}}'\mathbf{R}^{-1}\widetilde{\mathbf{Z}}+\widetilde{\mathbf{G}}^{-1}\end{bmatrix}^{-1}\begin{bmatrix}\mathbf{X}'\mathbf{R}^{-1}\mathbf{y}\\ \widetilde{\mathbf{Z}}'\mathbf{R}^{-1}\mathbf{y}\end{bmatrix}$$

Calculation of original effects:

$$\begin{bmatrix}\hat{\mathbf{b}}\\ \hat{\mathbf{u}}\end{bmatrix} = \begin{bmatrix}\mathbf{I}_b & \mathbf{0}\\ \mathbf{0} & \mathbf{M}\end{bmatrix}\begin{bmatrix}\hat{\mathbf{b}}\\ \widetilde{\mathbf{u}}\end{bmatrix}$$

Previously: Linearly Equivalent Mixed Model Equations

[1] M. Taskinen, E. A. Mäntysaari, and I. Strandén. "Single-step SNP-BLUP with on-the-fly imputed genotypes and residual polygenic effects". In: *Genet. Sel. Evol.* 49.1 (2017), p. 15.

- **Linearly equivalent MMEs** [1]: decomposition of relationship matrix $\mathbf{G}$ of random effects $\hat{\mathbf{u}}$ as

$$\mathbf{G} = \mathbf{M}\tilde{\mathbf{G}}\mathbf{M}'$$

- Alternative MME: parts of decomposition attached to new model matrix $\tilde{\mathbf{Z}} = \mathbf{Z}\mathbf{M}$ and new set of random effects $\tilde{\mathbf{u}}$ related through $\tilde{\mathbf{G}}$.
- Original MME solved from $\hat{\mathbf{u}} = \mathbf{M}\tilde{\mathbf{u}}$.
- Inversion $\mathbf{G}^{-1}$ avoided if $\tilde{\mathbf{G}} = \mathbf{I}$, i.e. $\tilde{\mathbf{u}}$ **"orthogonal"**.
- GBLUP $\Leftrightarrow$ SNP-BLUP: $\tilde{\mathbf{u}}_m$ are **marker effects.**

Original MME

$$\begin{bmatrix}\hat{\mathbf{b}}\\ \hat{\mathbf{u}}\end{bmatrix} = \begin{bmatrix}\mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}\\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}+\mathbf{G}^{-1}\end{bmatrix}^{-1}\begin{bmatrix}\mathbf{X}'\mathbf{R}^{-1}\mathbf{y}\\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y}\end{bmatrix}$$

Linearly equivalent MME

$$\begin{bmatrix}\hat{\mathbf{b}}\\ \tilde{\mathbf{u}}\end{bmatrix} = \begin{bmatrix}\mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\tilde{\mathbf{Z}}\\ \tilde{\mathbf{Z}}'\mathbf{R}^{-1}\mathbf{X} & \tilde{\mathbf{Z}}'\mathbf{R}^{-1}\tilde{\mathbf{Z}}+\tilde{\mathbf{G}}^{-1}\end{bmatrix}^{-1}\begin{bmatrix}\mathbf{X}'\mathbf{R}^{-1}\mathbf{y}\\ \tilde{\mathbf{Z}}'\mathbf{R}^{-1}\mathbf{y}\end{bmatrix}$$

Calculation of original effects:

$$\begin{bmatrix}\hat{\mathbf{b}}\\ \hat{\mathbf{u}}\end{bmatrix} = \begin{bmatrix}\mathbf{I}_b & \mathbf{0}\\ \mathbf{0} & \mathbf{M}\end{bmatrix}\begin{bmatrix}\hat{\mathbf{b}}\\ \tilde{\mathbf{u}}\end{bmatrix}$$

Previously: Linearly Equivalent ssGBLUP Formulations

[1] M. Taskinen, E. A. Mäntysaari, and I. Strandén. "Single-step SNP-BLUP with on-the-fly imputed genotypes and residual polygenic effects". In: *Genet. Sel. Evol.* 49.1 (2017), p. 15.

- Unlimited number of possible decompositions of ssGBLUP relationship matrix [1]:

$$\mathbf{H} = \mathbf{M}_i \tilde{\mathbf{G}}_i \mathbf{M}_i'$$

- Linearly equivalent ssGBLUP formulations with varying number of random effects $\tilde{\mathbf{u}}$ without or with marker effects (ssSNP-BLUP).

- Numerical solutions same as with original ssGBLUP.

- Fully orthogonalized ($\tilde{\mathbf{G}} = \mathbf{I}$) ssSNP-BLUPs:
  - almost identical **number of iterations** when solved using **iterative methods**, e.g. PCG.

Previously: Linearly Equivalent ssGBLUP Formulations

[1] M. Taskinen, E. A. Mäntysaari, and I. Strandén. “Single-step SNP-BLUP with on-the-fly imputed genotypes and residual polygenic effects”. In: *Genet. Sel. Evol.* 49.1 (2017), p. 15.

- Unlimited number of possible decompositions of ssGBLUP relationship matrix [1]:

$$\mathbf{H} = \mathbf{M}_i \widetilde{\mathbf{G}}_i \mathbf{M}_i'$$

- Linearly equivalent ssGBLUP formulations with varying number of random effects $\widetilde{\mathbf{u}}$ without or with marker effects (ssSNP-BLUP).
- Numerical solutions same as with original ssGBLUP.
- Fully orthogonalized ($\widetilde{\mathbf{G}} = \mathbf{I}$) ssSNP-BLUPs:
  - almost identical **number of iterations** when solved using **iterative methods**, e.g. PCG.

$$\mathbf{M}_1 = \begin{bmatrix} \mathbf{I}_1 & \mathbf{A}_{imp} \\ \mathbf{0} & \mathbf{I}_2 \end{bmatrix} \qquad \widetilde{\mathbf{G}}_1 = \begin{bmatrix} (\mathbf{A}^{11})^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}_w \end{bmatrix} \qquad \mathbf{G}_w = w\mathbf{A}_{22} + (1-w)\mathbf{Z}_m\mathbf{Z}_m'$$

$$\mathbf{M}_2 = \begin{bmatrix} \mathbf{I}_1 & \sqrt{w}\mathbf{A}_{imp} & \sqrt{1-w}\mathbf{Z}_{imp} \\ \mathbf{0} & \sqrt{w}\mathbf{I}_2 & \sqrt{1-w}\mathbf{Z}_m \end{bmatrix} \qquad \widetilde{\mathbf{G}}_2 = \begin{bmatrix} (\mathbf{A}^{11})^{-1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{22} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_m \end{bmatrix} \qquad \mathbf{Z}_{imp} = \mathbf{A}_{imp}\mathbf{Z}_m$$

$$\mathbf{M}_3 = \begin{bmatrix} \sqrt{1-w}\mathbf{I}_1 & \sqrt{w}\mathbf{J}_1 & \sqrt{1-w}\mathbf{Z}_{imp} \\ \mathbf{0} & \sqrt{w}\mathbf{J}_2 & \sqrt{1-w}\mathbf{Z}_m \end{bmatrix} \qquad \widetilde{\mathbf{G}}_3 = \begin{bmatrix} (\mathbf{A}^{11})^{-1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_m \end{bmatrix}$$

Alternative RPG form
$\mathbf{J}_1$ picks non-genot.
$\mathbf{J}_2$ picks genotyped

$$\mathbf{M}_4 = \begin{bmatrix} (\mathbf{L}_1')^{-1} & \sqrt{w}\mathbf{A}_{imp}\mathbf{J}_2(\mathbf{L}')^{-1} & \sqrt{1-w}\mathbf{Z}_{imp} \\ \mathbf{0} & \sqrt{w}\mathbf{J}_2(\mathbf{L}')^{-1} & \sqrt{1-w}\mathbf{Z}_m \end{bmatrix} \qquad \widetilde{\mathbf{G}}_4 = \begin{bmatrix} \mathbf{I}_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{all} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_m \end{bmatrix}$$

Orthogonalized $\widetilde{\mathbf{G}}_2$
$\mathbf{L}$ = `chol(`$\mathbf{A}^{-1}$`)`
$\mathbf{L}_1$ = `chol(`$\mathbf{A}^{11}$`)`

$$\mathbf{M}_5 = \begin{bmatrix} \sqrt{1-w}(\mathbf{L}_1')^{-1} & \sqrt{w}\mathbf{J}_1(\mathbf{L}')^{-1} & \sqrt{1-w}\mathbf{Z}_{imp} \\ \mathbf{0} & \sqrt{w}\mathbf{J}_2(\mathbf{L}')^{-1} & \sqrt{1-w}\mathbf{Z}_m \end{bmatrix} \qquad \widetilde{\mathbf{G}}_5 = \begin{bmatrix} \mathbf{I}_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{all} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_m \end{bmatrix}$$

Orthogonalized $\widetilde{\mathbf{G}}_3$

$$\mathbf{M}_6 = \begin{bmatrix} (\mathbf{L}_1')^{-1} & \sqrt{w}\mathbf{A}_{imp}\hat{\mathbf{J}}_2(\hat{\mathbf{L}}')^{-1} & \sqrt{1-w}\mathbf{Z}_{imp} \\ \mathbf{0} & \sqrt{w}\hat{\mathbf{J}}_2(\hat{\mathbf{L}}')^{-1} & \sqrt{1-w}\mathbf{Z}_m \end{bmatrix} \qquad \widetilde{\mathbf{G}}_6 = \begin{bmatrix} \mathbf{I}_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{ganc} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_m \end{bmatrix}$$

Reduced $\mathbf{A}_{22}$ of $\widetilde{\mathbf{G}}_4$
$\hat{\mathbf{L}}$ = `chol(`$\hat{\mathbf{A}}^{-1}$`)`, where
$\hat{\mathbf{A}}$: genot. and ancestors

Previously: Linearly Equivalent ssGBLUP Formulations

[1] M. Taskinen, E. A. Mäntysaari, and I. Strandén. “Single-step SNP-BLUP with on-the-fly imputed genotypes and residual polygenic effects”. In: *Genet. Sel. Evol.* 49.1 (2017), p. 15.

- Unlimited number of possible decompositions of ssGBLUP relationship matrix [1]:

$$\mathbf{H} = \mathbf{M}_i \widetilde{\mathbf{G}}_i \mathbf{M}_i'$$

- Linearly equivalent ssGBLUP formulations with varying number of random effects $\widetilde{\mathbf{u}}$ without or with marker effects (ssSNP-BLUP).
- Numerical solutions same as with original ssGBLUP.
- Fully orthogonalized ($\widetilde{\mathbf{G}} = \mathbf{I}$) ssSNP-BLUPs:
  - almost identical **number of iterations** when solved using **iterative methods**, e.g. PCG.

$$\mathbf{M}_1 = \begin{bmatrix} \mathbf{I}_1 & \mathbf{A}_{imp} \\ \mathbf{0} & \mathbf{I}_2 \end{bmatrix} \qquad \widetilde{\mathbf{G}}_1 = \begin{bmatrix} (\mathbf{A}^{11})^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}_w \end{bmatrix} \qquad \mathbf{G}_w = w\mathbf{A}_{22} + (1-w)\mathbf{Z}_m\mathbf{Z}_m'$$

$$\mathbf{M}_2 = \begin{bmatrix} \mathbf{I}_1 & \sqrt{w}\mathbf{A}_{imp} & \sqrt{1-w}\mathbf{Z}_{imp} \\ \mathbf{0} & \sqrt{w}\mathbf{I}_2 & \sqrt{1-w}\mathbf{Z}_m \end{bmatrix} \qquad \widetilde{\mathbf{G}}_2 = \begin{bmatrix} (\mathbf{A}^{11})^{-1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{22} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_m \end{bmatrix} \qquad \mathbf{Z}_{imp} = \mathbf{A}_{imp}\mathbf{Z}_m$$

$$\mathbf{M}_3 = \begin{bmatrix} \sqrt{1-w}\mathbf{I}_1 & \sqrt{w}\mathbf{J}_1 & \sqrt{1-w}\mathbf{Z}_{imp} \\ \mathbf{0} & \sqrt{w}\mathbf{J}_2 & \sqrt{1-w}\mathbf{Z}_m \end{bmatrix} \qquad \widetilde{\mathbf{G}}_3 = \begin{bmatrix} (\mathbf{A}^{11})^{-1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_m \end{bmatrix}$$

Alternative RPG form
$\mathbf{J}_1$ picks non-genot.
$\mathbf{J}_2$ picks genotyped

$$\mathbf{M}_4 = \begin{bmatrix} (\mathbf{L}_1')^{-1} & \sqrt{w}\mathbf{A}_{imp}\mathbf{J}_2(\mathbf{L}')^{-1} & \sqrt{1-w}\mathbf{Z}_{imp} \\ \mathbf{0} & \sqrt{w}\mathbf{J}_2(\mathbf{L}')^{-1} & \sqrt{1-w}\mathbf{Z}_m \end{bmatrix} \qquad \widetilde{\mathbf{G}}_4 = \begin{bmatrix} \mathbf{I}_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{all} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_m \end{bmatrix}$$

Orthogonalized $\widetilde{\mathbf{G}}_2$
$\mathbf{L} = \mathtt{chol}(\mathbf{A}^{-1})$
$\mathbf{L}_1 = \mathtt{chol}(\mathbf{A}^{11})$

$$\mathbf{M}_5 = \begin{bmatrix} \sqrt{1-w}(\mathbf{L}_1')^{-1} & \sqrt{w}\mathbf{J}_1(\mathbf{L}')^{-1} & \sqrt{1-w}\mathbf{Z}_{imp} \\ \mathbf{0} & \sqrt{w}\mathbf{J}_2(\mathbf{L}')^{-1} & \sqrt{1-w}\mathbf{Z}_m \end{bmatrix} \qquad \widetilde{\mathbf{G}}_5 = \begin{bmatrix} \mathbf{I}_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{all} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_m \end{bmatrix}$$

Orthogonalized $\widetilde{\mathbf{G}}_3$

$$\mathbf{M}_6 = \begin{bmatrix} (\mathbf{L}_1')^{-1} & \sqrt{w}\mathbf{A}_{imp}\widehat{\mathbf{J}}_2(\widehat{\mathbf{L}}')^{-1} & \sqrt{1-w}\mathbf{Z}_{imp} \\ \mathbf{0} & \sqrt{w}\widehat{\mathbf{J}}_2(\widehat{\mathbf{L}}')^{-1} & \sqrt{1-w}\mathbf{Z}_m \end{bmatrix} \qquad \widetilde{\mathbf{G}}_6 = \begin{bmatrix} \mathbf{I}_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{ganc} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_m \end{bmatrix}$$

Reduced $\mathbf{A}_{22}$ of $\widetilde{\mathbf{G}}_4$
$\widehat{\mathbf{L}} = \mathtt{chol}(\widehat{\mathbf{A}}^{-1})$, where
$\widehat{\mathbf{A}}$: genot. and ancestors

Previously: Linearly Equivalent ssGBLUP Formulations

[1] M. Taskinen, E. A. Mäntysaari, and I. Strandén. "Single-step SNP-BLUP with on-the-fly imputed genotypes and residual polygenic effects". In: *Genet. Sel. Evol.* 49.1 (2017), p. 15.

- Unlimited number of possible decompositions of ssGBLUP relationship matrix [1]:

$$\mathbf{H} = \mathbf{M}_i \widetilde{\mathbf{G}}_i \mathbf{M}_i'$$

- Linearly equivalent ssGBLUP formulations with varying number of random effects $\tilde{\mathbf{u}}$ without or with marker effects (ssSNP-BLUP).
- Numerical solutions same as with original ssGBLUP.
- Fully orthogonalized ($\widetilde{\mathbf{G}} = \mathbf{I}$) ssSNP-BLUPs:
  - almost identical **number of iterations** when solved using **iterative methods**, e.g. PCG.

$$\mathbf{M}_1 = \begin{bmatrix} \mathbf{I}_1 & \mathbf{A}_{imp} \\ \mathbf{0} & \mathbf{I}_2 \end{bmatrix} \qquad \widetilde{\mathbf{G}}_1 = \begin{bmatrix} (\mathbf{A}^{11})^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}_w \end{bmatrix} \qquad \mathbf{G}_w = w\mathbf{A}_{22} + (1-w)\mathbf{Z}_m\mathbf{Z}_m'$$

$$\mathbf{M}_2 = \begin{bmatrix} \mathbf{I}_1 & \sqrt{w}\mathbf{A}_{imp} & \sqrt{1-w}\mathbf{Z}_{imp} \\ \mathbf{0} & \sqrt{w}\mathbf{I}_2 & \sqrt{1-w}\mathbf{Z}_m \end{bmatrix} \qquad \widetilde{\mathbf{G}}_2 = \begin{bmatrix} (\mathbf{A}^{11})^{-1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{22} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_m \end{bmatrix} \qquad \mathbf{Z}_{imp} = \mathbf{A}_{imp}\mathbf{Z}_m$$

$$\mathbf{M}_3 = \begin{bmatrix} \sqrt{1-w}\mathbf{I}_1 & \sqrt{w}\mathbf{J}_1 & \sqrt{1-w}\mathbf{Z}_{imp} \\ \mathbf{0} & \sqrt{w}\mathbf{J}_2 & \sqrt{1-w}\mathbf{Z}_m \end{bmatrix} \qquad \widetilde{\mathbf{G}}_3 = \begin{bmatrix} (\mathbf{A}^{11})^{-1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_m \end{bmatrix}$$

Alternative RPG form
$\mathbf{J}_1$ picks non-genot.
$\mathbf{J}_2$ picks genotyped

$$\mathbf{M}_4 = \begin{bmatrix} (\mathbf{L}_1')^{-1} & \sqrt{w}\mathbf{A}_{imp}\mathbf{J}_2(\mathbf{L}')^{-1} & \sqrt{1-w}\mathbf{Z}_{imp} \\ \mathbf{0} & \sqrt{w}\mathbf{J}_2(\mathbf{L}')^{-1} & \sqrt{1-w}\mathbf{Z}_m \end{bmatrix} \qquad \widetilde{\mathbf{G}}_4 = \begin{bmatrix} \mathbf{I}_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{all} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_m \end{bmatrix}$$

Orthogonalized $\widetilde{\mathbf{G}}_2$
$\mathbf{L} = \texttt{chol}(\mathbf{A}^{-1})$
$\mathbf{L}_1 = \texttt{chol}(\mathbf{A}^{11})$

$$\mathbf{M}_5 = \begin{bmatrix} \sqrt{1-w}(\mathbf{L}_1')^{-1} & \sqrt{w}\mathbf{J}_1(\mathbf{L}')^{-1} & \sqrt{1-w}\mathbf{Z}_{imp} \\ \mathbf{0} & \sqrt{w}\mathbf{J}_2(\mathbf{L}')^{-1} & \sqrt{1-w}\mathbf{Z}_m \end{bmatrix} \qquad \widetilde{\mathbf{G}}_5 = \begin{bmatrix} \mathbf{I}_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{all} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_m \end{bmatrix}$$

Orthogonalized $\widetilde{\mathbf{G}}_3$

$$\mathbf{M}_6 = \begin{bmatrix} (\mathbf{L}_1')^{-1} & \sqrt{w}\mathbf{A}_{imp}\hat{\mathbf{J}}_2(\hat{\mathbf{L}}')^{-1} & \sqrt{1-w}\mathbf{Z}_{imp} \\ \mathbf{0} & \sqrt{w}\hat{\mathbf{J}}_2(\hat{\mathbf{L}}')^{-1} & \sqrt{1-w}\mathbf{Z}_m \end{bmatrix} \qquad \widetilde{\mathbf{G}}_6 = \begin{bmatrix} \mathbf{I}_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{ganc} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_m \end{bmatrix}$$

Reduced $\mathbf{A}_{22}$ of $\widetilde{\mathbf{G}}_4$
$\hat{\mathbf{L}} = \texttt{chol}(\hat{\mathbf{A}}^{-1})$, where
$\hat{\mathbf{A}}$: genot. and ancestors

NATURAL RESOURCES INSTITUTE FINLAND

Previously: Linearly Equivalent ssGBLUP Formulations

[1] M. Taskinen, E. A. Mäntysaari, and I. Strandén. “Single-step SNP-BLUP with on-the-fly imputed genotypes and residual polygenic effects”. In: *Genet. Sel. Evol.* 49.1 (2017), p. 15.

- Unlimited number of possible decompositions of ssGBLUP relationship matrix [1]:

$$\mathbf{H} = \mathbf{M}_i \widetilde{\mathbf{G}}_i \mathbf{M}_i'$$

- Linearly equivalent ssGBLUP formulations with varying number of random effects $\tilde{\mathbf{u}}$ without or with marker effects (ssSNP-BLUP).
- Numerical solutions same as with original ssGBLUP.
- Fully orthogonalized ($\widetilde{\mathbf{G}} = \mathbf{I}$) ssSNP-BLUPs:
  - almost identical **number of iterations** when solved using **iterative methods**, e.g. PCG.

$$\mathbf{M}_1 = \begin{bmatrix} \mathbf{I}_1 & \mathbf{A}_{imp} \\ \mathbf{0} & \mathbf{I}_2 \end{bmatrix} \qquad \widetilde{\mathbf{G}}_1 = \begin{bmatrix} (\mathbf{A}^{11})^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}_w \end{bmatrix} \qquad \mathbf{G}_w = w\mathbf{A}_{22} + (1-w)\mathbf{Z}_m\mathbf{Z}_m'$$

$$\mathbf{M}_2 = \begin{bmatrix} \mathbf{I}_1 & \sqrt{w}\mathbf{A}_{imp} & \sqrt{1-w}\mathbf{Z}_{imp} \\ \mathbf{0} & \sqrt{w}\mathbf{I}_2 & \sqrt{1-w}\mathbf{Z}_m \end{bmatrix} \qquad \widetilde{\mathbf{G}}_2 = \begin{bmatrix} (\mathbf{A}^{11})^{-1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{22} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_m \end{bmatrix} \qquad \mathbf{Z}_{imp} = \mathbf{A}_{imp}\mathbf{Z}_m$$

$$\mathbf{M}_3 = \begin{bmatrix} \sqrt{1-w}\mathbf{I}_1 & \sqrt{w}\mathbf{J}_1 & \sqrt{1-w}\mathbf{Z}_{imp} \\ \mathbf{0} & \sqrt{w}\mathbf{J}_2 & \sqrt{1-w}\mathbf{Z}_m \end{bmatrix} \qquad \widetilde{\mathbf{G}}_3 = \begin{bmatrix} (\mathbf{A}^{11})^{-1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_m \end{bmatrix}$$

Alternative RPG form
$\mathbf{J}_1$ picks non-genot.
$\mathbf{J}_2$ picks genotyped

$$\mathbf{M}_4 = \begin{bmatrix} (\mathbf{L}_1')^{-1} & \sqrt{w}\mathbf{A}_{imp}\mathbf{J}_2(\mathbf{L}')^{-1} & \sqrt{1-w}\mathbf{Z}_{imp} \\ \mathbf{0} & \sqrt{w}\mathbf{J}_2(\mathbf{L}')^{-1} & \sqrt{1-w}\mathbf{Z}_m \end{bmatrix} \qquad \widetilde{\mathbf{G}}_4 = \begin{bmatrix} \mathbf{I}_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{all} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_m \end{bmatrix}$$

Orthogonalized $\widetilde{\mathbf{G}}_2$
$\mathbf{L} = \texttt{chol}(\mathbf{A}^{-1})$
$\mathbf{L}_1 = \texttt{chol}(\mathbf{A}^{11})$

$$\mathbf{M}_5 = \begin{bmatrix} \sqrt{1-w}(\mathbf{L}_1')^{-1} & \sqrt{w}\mathbf{J}_1(\mathbf{L}')^{-1} & \sqrt{1-w}\mathbf{Z}_{imp} \\ \mathbf{0} & \sqrt{w}\mathbf{J}_2(\mathbf{L}')^{-1} & \sqrt{1-w}\mathbf{Z}_m \end{bmatrix} \qquad \widetilde{\mathbf{G}}_5 = \begin{bmatrix} \mathbf{I}_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{all} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_m \end{bmatrix}$$

Orthogonalized $\widetilde{\mathbf{G}}_3$

$$\mathbf{M}_6 = \begin{bmatrix} (\mathbf{L}_1')^{-1} & \sqrt{w}\mathbf{A}_{imp}\hat{\mathbf{J}}_2(\hat{\mathbf{L}}')^{-1} & \sqrt{1-w}\mathbf{Z}_{imp} \\ \mathbf{0} & \sqrt{w}\hat{\mathbf{J}}_2(\hat{\mathbf{L}}')^{-1} & \sqrt{1-w}\mathbf{Z}_m \end{bmatrix} \qquad \widetilde{\mathbf{G}}_6 = \begin{bmatrix} \mathbf{I}_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{ganc} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_m \end{bmatrix}$$

Reduced $\mathbf{A}_{22}$ of $\widetilde{\mathbf{G}}_4$
$\hat{\mathbf{L}} = \texttt{chol}(\hat{\mathbf{A}}^{-1})$, where
$\hat{\mathbf{A}}$: genot. and ancestors

MME	Heritability $h^2 = 0.1$							
	polygenic proportion w							
	no preconditioning				with preconditioning			
	0.00	0.01	0.20	1.00	0.00	0.01	0.20	1.00
orig.	-	619	624	623	-	178	126	119
1	-	758	701	744	-	191	148	139
2	736	1005	907	748	438	491	429	139
3	736	3422	3154	623	438	1732	1416	119
4	74	74	72	71	107	105	151	82
5	74	73	70	70	107	107	151	49
6	74	74	72	71	107	105	151	82

Result 1: Iteration Convergence of Fully Orthogonalized ($\mathbf{G} = \mathbf{M}_i\mathbf{M}_i'$) MMEs

- Assuming **singular values** of (here usually) **wide** matrices $\mathbf{M}_i$ to be known: diagonal $\mathbf{D}_i$.
- Since $\mathbf{M}_i$ are all decompositions of same $\mathbf{G}$, singular values are shared: $\mathbf{D}_i = \mathbf{D}$.
- If fixed effects are neglected, $\mathbf{Z} = \mathbf{I}$, and single trait case is assumed, all fully orthogonalized MME share eigenvalues $\mathbf{D}^2 + \lambda\mathbf{I}$ and rest equal λ.
- Numbers of **distinct** (approximate) eigenvalues are thus same $\Rightarrow$ this explains same iteration counts.

Singular value decomposition of $\mathbf{M}_i$

$$\mathbf{M}_i = \mathbf{U}_i \begin{bmatrix} \mathbf{D}_i & \mathbf{0} \end{bmatrix} \mathbf{V}_i'$$

Luke NATURAL RESOURCES INSTITUTE FINLAND

Result 1: Iteration Convergence of Fully Orthogonalized ($\mathbf{G} = \mathbf{M}_i\mathbf{M}_i'$) MMEs

- Assuming **singular values** of (here usually) **wide** matrices $\mathbf{M}_i$ to be known: diagonal $\mathbf{D}_i$.
- Since $\mathbf{M}_i$ are all decompositions of same $\mathbf{G}$, singular values are shared: $\mathbf{D}_i = \mathbf{D}$.
- If fixed effects are neglected, $\mathbf{Z} = \mathbf{I}$, and single trait case is assumed, all fully orthogonalized MME share eigenvalues $\mathbf{D}^2 + \lambda\mathbf{I}$ and rest equal λ.
- Numbers of **distinct** (approximate) eigenvalues are thus same $\Rightarrow$ this explains same iteration counts.

Singular value decomposition of $\mathbf{M}_i$

$$\mathbf{M}_i = \mathbf{U}_i \begin{bmatrix} \mathbf{D}_i & \mathbf{0} \end{bmatrix} \mathbf{V}_i'$$

Eigendecomposition of $\mathbf{G}$

$$\mathbf{G} = \mathbf{M}_i\mathbf{M}_i' = \mathbf{U}_i\mathbf{D}_i^2\mathbf{U}_i' = \mathbf{U}\mathbf{D}^2\mathbf{U}'$$

Result 1: Iteration Convergence of Fully Orthogonalized ($\mathbf{G} = \mathbf{M}_i\mathbf{M}_i'$) MMEs

- Assuming **singular values** of (here usually) **wide** matrices $\mathbf{M}_i$ to be known: diagonal $\mathbf{D}_i$.
- Since $\mathbf{M}_i$ are all decompositions of same $\mathbf{G}$, singular values are shared: $\mathbf{D}_i = \mathbf{D}$.
- If fixed effects are neglected, $\mathbf{Z} = \mathbf{I}$, and single trait case is assumed, all fully orthogonalized MME share eigenvalues $\mathbf{D}^2 + \lambda\mathbf{I}$ and rest equal λ.
- Numbers of **distinct** (approximate) eigenvalues are thus same ⇒ this explains same iteration counts.

Singular value decomposition of $\mathbf{M}_i$

$$\mathbf{M}_i = \mathbf{U}_i \begin{bmatrix}\mathbf{D}_i & \mathbf{0}\end{bmatrix} \mathbf{V}_i'$$

Eigendecomposition of $\mathbf{G}$

$$\mathbf{G} = \mathbf{M}_i\mathbf{M}_i' = \mathbf{U}_i\mathbf{D}_i^2\mathbf{U}_i' = \mathbf{U}\mathbf{D}^2\mathbf{U}'$$

Eigendecomposition of "MME"

$$\widetilde{\mathbf{Z}}_i'\widetilde{\mathbf{Z}}_i + \lambda\mathbf{I} \approx \mathbf{V}_i \begin{bmatrix}\mathbf{D}^2 + \lambda\mathbf{I} & \mathbf{0} \\ \mathbf{0} & \lambda\mathbf{I}\end{bmatrix} \mathbf{V}_i'$$

Luke NATURAL RESOURCES INSTITUTE FINLAND

Result 1: Iteration Convergence of Fully Orthogonalized ($\mathbf{G} = \mathbf{M}_i\mathbf{M}_i'$) MMEs

- Assuming **singular values** of (here usually) **wide** matrices $\mathbf{M}_i$ to be known: diagonal $\mathbf{D}_i$.
- Since $\mathbf{M}_i$ are all decompositions of same $\mathbf{G}$, singular values are shared: $\mathbf{D}_i = \mathbf{D}$.
- If fixed effects are neglected, $\mathbf{Z} = \mathbf{I}$, and single trait case is assumed, all fully orthogonalized MME share eigenvalues $\mathbf{D}^2 + \lambda\mathbf{I}$ and rest equal λ.
- Numbers of **distinct** (approximate) eigenvalues are thus same ⇒ this explains same iteration counts.

Singular value decomposition of $\mathbf{M}_i$

$$\mathbf{M}_i = \mathbf{U}_i \begin{bmatrix} \mathbf{D}_i & \mathbf{0} \end{bmatrix} \mathbf{V}_i'$$

Eigendecomposition of $\mathbf{G}$

$$\mathbf{G} = \mathbf{M}_i\mathbf{M}_i' = \mathbf{U}_i\mathbf{D}_i^2\mathbf{U}_i' = \mathbf{U}\mathbf{D}^2\mathbf{U}'$$

Eigendecomposition of “MME”

$$\widetilde{\mathbf{Z}}_i'\widetilde{\mathbf{Z}}_i + \lambda\mathbf{I} \approx \mathbf{V}_i \begin{bmatrix} \mathbf{D}^2 + \lambda\mathbf{I} & \mathbf{0} \\ \mathbf{0} & \lambda\mathbf{I} \end{bmatrix} \mathbf{V}_i'$$

Numerical efficiency: Smaller Pedigree of Ancestors of Genotyped: $\hat{\mathbf{A}}^{-1}$

- Linearly equivalent ssSNP-BLUPs: same solution and convergence, which formulation to choose?
  - Select **numerically efficient** formulation.
  - How to get a more efficient formulation?

- Group of non-genotyped (1) split to **non-ancestors** (n) and **ancestors** (a) of genotyped (2).

- **Smaller pedigree** of non-genotyped ancestors (a) and genotyped (2) individuals can be formed with $\hat{\mathbf{A}}^{-1} = \hat{\mathbf{L}}\hat{\mathbf{L}}'$ as its pedigree relationship matrix.

Numerical efficiency: Smaller Pedigree of Ancestors of Genotyped: $\hat{\mathbf{A}}^{-1}$

- Linearly equivalent ssSNP-BLUPs: same solution and convergence, which formulation to choose?
  - Select **numerically efficient** formulation.
  - How to get a more efficient formulation?
- Group of non-genotyped (1) split to **non-ancestors** (n) and **ancestors** (a) of genotyped (2).
- **Smaller pedigree** of non-genotyped ancestors (a) and genotyped (2) individuals can be formed with $\hat{\mathbf{A}}^{-1} = \hat{\mathbf{L}}\hat{\mathbf{L}}'$ as its pedigree relationship matrix.

NATURAL RESOURCES INSTITUTE FINLAND

Numerical efficiency: Smaller Pedigree of Ancestors of Genotyped: $\hat{\mathbf{A}}^{-1}$

- Linearly equivalent ssSNP-BLUPs: same solution and convergence, which formulation to choose?
  - Select **numerically efficient** formulation.
  - How to get a more efficient formulation?
- Group of non-genotyped (1) split to **non-ancestors** (n) and **ancestors** (a) of genotyped (2).
- **Smaller pedigree** of non-genotyped ancestors (a) and genotyped (2) individuals can be formed with $\hat{\mathbf{A}}^{-1} = \hat{\mathbf{L}}\hat{\mathbf{L}}'$ as its pedigree relationship matrix.

Numerical efficiency: Smaller Pedigree of Ancestors of Genotyped: $\hat{\mathbf{A}}^{-1}$

- Linearly equivalent ssSNP-BLUPs: same solution and convergence, which formulation to choose?
  - Select **numerically efficient** formulation.
  - How to get a more efficient formulation?
- Group of non-genotyped (1) split to **non-ancestors** (n) and **ancestors** (a) of genotyped (2).
- **Smaller pedigree** of non-genotyped ancestors (a) and genotyped (2) individuals can be formed with $\hat{\mathbf{A}}^{-1} = \hat{\mathbf{L}}\hat{\mathbf{L}}'$ as its pedigree relationship matrix.

“Cholesky” of $\mathbf{A}^{-1} = \mathbf{L}\mathbf{L}'$

$$\mathbf{L} = \begin{bmatrix} \mathbf{L}_{nn} & \mathbf{0} & \mathbf{0} \\ \mathbf{L}_{an} & \mathbf{L}_{aa} & \mathbf{L}_{a2} \\ \mathbf{L}_{2n} & \mathbf{L}_{2a} & \mathbf{L}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{L}_{nn} & \mathbf{0} \quad \mathbf{0} \\ \mathbf{L}_{an} & \\ \mathbf{L}_{2n} & \begin{bmatrix} \hat{\mathbf{L}} \end{bmatrix} \end{bmatrix}$$

Numerical efficiency: Smaller Pedigree of Ancestors of Genotyped: $\hat{\mathbf{A}}^{-1}$

- Linearly equivalent ssSNP-BLUPs: same solution and convergence, which formulation to choose?
  - Select **numerically efficient** formulation.
  - How to get a more efficient formulation?
- Group of non-genotyped (1) split to **non-ancestors** (n) and **ancestors** (a) of genotyped (2).
- **Smaller pedigree** of non-genotyped ancestors (a) and genotyped (2) individuals can be formed with $\hat{\mathbf{A}}^{-1} = \hat{\mathbf{L}}\hat{\mathbf{L}}'$ as its pedigree relationship matrix.

"Cholesky" of $\mathbf{A}^{-1} = \mathbf{L}\mathbf{L}'$

$$\mathbf{L} = \begin{bmatrix} \mathbf{L}_{nn} & \mathbf{0} & \mathbf{0} \\ \mathbf{L}_{an} & \mathbf{L}_{aa} & \mathbf{L}_{a2} \\ \mathbf{L}_{2n} & \mathbf{L}_{2a} & \mathbf{L}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{L}_{nn} & \mathbf{0} & \mathbf{0} \\ \mathbf{L}_{an} & \multicolumn{2}{c}{\multirow{2}{*}{$[\hat{\mathbf{L}}]$}} \\ \mathbf{L}_{2n} & & \end{bmatrix}$$

"Cholesky" of $\hat{\mathbf{A}}^{-1} = \hat{\mathbf{L}}\hat{\mathbf{L}}'$

Smaller pedigree of non-genotyped ancestors (a) and genotyped (2):

$$\hat{\mathbf{L}} = \begin{bmatrix} \hat{\mathbf{L}}_{a} \\ \hat{\mathbf{L}}_{2} \end{bmatrix} = \begin{bmatrix} \mathbf{L}_{aa} & \mathbf{L}_{a2} \\ \mathbf{L}_{2a} & \mathbf{L}_{22} \end{bmatrix},$$

NATURAL RESOURCES INSTITUTE FINLAND

Partial Orthogonalization of ssGBLUP Relationship Matrix $\mathbf{H}$

- "Cholesky" matrix $\mathbf{L}$ (of $\mathbf{A}^{-1} = \mathbf{L}\mathbf{L}'$) naturally orthogonalizes non-genotyped non-ancestors (n) from ssGBLUP relationship matrix $\mathbf{H}$.

Orthogonalizing group (n) in $\mathbf{H}$

$$\mathbf{L}'\mathbf{H}\mathbf{L} = \begin{bmatrix} \mathbf{I}_n & \mathbf{0} \\ \mathbf{0} & \hat{\mathbf{L}}'\hat{\mathbf{H}}\hat{\mathbf{L}} \end{bmatrix}$$

- Smaller pedigree individuals are related through:

$$\hat{\mathbf{L}}'\hat{\mathbf{H}}\hat{\mathbf{L}} = \hat{\mathbf{I}} + (1-w)\mathbf{P}^{\perp}_{\hat{\mathbf{L}}'_a}(\hat{\mathbf{L}}'_2\mathbf{Z}_m\mathbf{Z}'_m\hat{\mathbf{L}}_2 - \hat{\mathbf{I}})\mathbf{P}^{\perp}_{\hat{\mathbf{L}}'_a}$$

- Former **"on-the-fly" imputation operations** of genomic information are now part of **orthogonal projection $\mathbf{P}^{\perp}_{\hat{\mathbf{L}}'_a}$**.

Partial Orthogonalization of ssGBLUP Relationship Matrix $\mathbf{H}$

- "Cholesky" matrix $\mathbf{L}$ (of $\mathbf{A}^{-1} = \mathbf{L}\mathbf{L}'$) naturally orthogonalizes non-genotyped non-ancestors (n) from ssGBLUP relationship matrix $\mathbf{H}$.

- Smaller pedigree individuals are related through:

$$\hat{\mathbf{L}}'\hat{\mathbf{H}}\hat{\mathbf{L}} = \hat{\mathbf{I}} + (1-w)\mathbf{P}^{\perp}_{\hat{\mathbf{L}}_a'}(\hat{\mathbf{L}}_2'\mathbf{Z}_m\mathbf{Z}_m'\hat{\mathbf{L}}_2 - \hat{\mathbf{I}})\mathbf{P}^{\perp}_{\hat{\mathbf{L}}_a'}$$

- Former "**on-the-fly" imputation operations** of genomic information are now part of **orthogonal projection** $\mathbf{P}^{\perp}_{\hat{\mathbf{L}}_a'}$.

Orthogonalizing group (n) in $\mathbf{H}$

$$\mathbf{L}'\mathbf{H}\mathbf{L} = \begin{bmatrix} \mathbf{I}_n & \mathbf{0} \\ \mathbf{0} & \hat{\mathbf{L}}'\hat{\mathbf{H}}\hat{\mathbf{L}} \end{bmatrix}$$

Partial Orthogonalization of ssGBLUP Relationship Matrix $\mathbf{H}$

- "Cholesky" matrix $\mathbf{L}$ (of $\mathbf{A}^{-1} = \mathbf{L}\mathbf{L}'$) naturally orthogonalizes non-genotyped non-ancestors (n) from ssGBLUP relationship matrix $\mathbf{H}$.

- Smaller pedigree individuals are related through:

$$\hat{\mathbf{L}}'\hat{\mathbf{H}}\hat{\mathbf{L}} = \hat{\mathbf{I}} + (1-w)\mathbf{P}^{\perp}_{\hat{\mathbf{L}}'_a}(\hat{\mathbf{L}}'_2\mathbf{Z}_m\mathbf{Z}'_m\hat{\mathbf{L}}_2 - \hat{\mathbf{I}})\mathbf{P}^{\perp}_{\hat{\mathbf{L}}'_a}$$

- Former **"on-the-fly" imputation operations** of genomic information are now part of **orthogonal projection $\mathbf{P}^{\perp}_{\hat{\mathbf{L}}'_a}$**.

Orthogonalizing group (n) in $\mathbf{H}$

$$\mathbf{L}'\mathbf{H}\mathbf{L} = \begin{bmatrix} \mathbf{I}_n & \mathbf{0} \\ \mathbf{0} & \hat{\mathbf{L}}'\hat{\mathbf{H}}\hat{\mathbf{L}} \end{bmatrix}$$

Orthogonal projection of $\hat{\mathbf{L}}'_a$

$$\mathbf{P}^{\perp}_{\hat{\mathbf{L}}'_a} = \hat{\mathbf{I}} - \hat{\mathbf{L}}'_a(\hat{\mathbf{L}}_a\hat{\mathbf{L}}'_a)^{-1}\hat{\mathbf{L}}_a$$

$$= \hat{\mathbf{I}} - \hat{\mathbf{L}}'_a(\hat{\mathbf{A}}^{aa})^{-1}\hat{\mathbf{L}}_a$$

Luke NATURAL RESOURCES INSTITUTE FINLAND

Result 2: New ssSNP-BLUP Formulation using the Smaller Pedigree

- Let $\widetilde{\mathbf{L}}_a$ be **sparsity preserving Cholesky factorization** of $\widehat{\mathbf{A}}^{aa}$.

- New fully orthogonalized ssSNP-BLUP: $\mathbf{H} = \mathbf{M}\widetilde{\mathbf{G}}\mathbf{M}'$, where

$$\mathbf{M} = (\mathbf{L}')^{-1}\begin{bmatrix}\mathbf{I}_n & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \sqrt{1-w}\widehat{\mathbf{M}}_a & \sqrt{w}\widehat{\mathbf{I}} & \sqrt{1-w}(\widehat{\mathbf{I}} - \widehat{\mathbf{M}}_a\widehat{\mathbf{M}}_a')\widehat{\mathbf{L}}_2'\mathbf{Z}_m\end{bmatrix}$$

$\widetilde{\mathbf{G}} = \mathbf{I}$, and $\widehat{\mathbf{M}}_a = \widehat{\mathbf{L}}_a'(\widetilde{\mathbf{L}}_a')^{-1}$.

- Main advantage of new formulation: size of numerical sparse Cholesky factorization matrix $\widetilde{\mathbf{L}}_a$ is number of non-genotyped ancestors ($\widehat{\mathbf{A}}^{aa}$) instead of all non-genotyped ($\mathbf{A}^{11}$).

- Four groups of random effects $\widetilde{\mathbf{u}}$: non-genotyped ancestors (a) have two sets of animals effects.

Sparse Cholesky $\widetilde{\mathbf{L}}_a$ of $\widehat{\mathbf{A}}^{aa}$

$$\widehat{\mathbf{A}}^{aa} = \widehat{\mathbf{L}}_a\widehat{\mathbf{L}}_a' = \widetilde{\mathbf{L}}_a\widetilde{\mathbf{L}}_a'$$

Result 2: New ssSNP-BLUP Formulation using the Smaller Pedigree

- Let $\widetilde{\mathbf{L}}_a$ be **sparsity preserving Cholesky factorization** of $\widehat{\mathbf{A}}^{aa}$.

- New fully orthogonalized ssSNP-BLUP: $\mathbf{H} = \mathbf{M}\widetilde{\mathbf{G}}\mathbf{M}'$, where

$$\mathbf{M} = (\mathbf{L}')^{-1}\begin{bmatrix} \mathbf{I}_n & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \sqrt{1-w}\,\widehat{\mathbf{M}}_a & \sqrt{w}\,\widehat{\mathbf{I}} & \sqrt{1-w}\,(\widehat{\mathbf{I}} - \widehat{\mathbf{M}}_a\widehat{\mathbf{M}}'_a)\widehat{\mathbf{L}}'_2\mathbf{Z}_m \end{bmatrix}$$

$\widetilde{\mathbf{G}} = \mathbf{I}$, and $\widehat{\mathbf{M}}_a = \widehat{\mathbf{L}}'_a(\widetilde{\mathbf{L}}'_a)^{-1}$.

- Main advantage of new formulation: size of numerical sparse Cholesky factorization matrix $\widetilde{\mathbf{L}}_a$ is number of non-genotyped ancestors ($\widehat{\mathbf{A}}^{aa}$) instead of all non-genotyped ($\mathbf{A}^{11}$).

- Four groups of random effects $\widetilde{\mathbf{u}}$: non-genotyped ancestors (a) have two sets of animals effects.

Sparse Cholesky $\widetilde{\mathbf{L}}_a$ of $\widehat{\mathbf{A}}^{aa}$

$$\widehat{\mathbf{A}}^{aa} = \widehat{\mathbf{L}}_a\widehat{\mathbf{L}}'_a = \widetilde{\mathbf{L}}_a\widetilde{\mathbf{L}}'_a$$

Result 2: New ssSNP-BLUP Formulation using the Smaller Pedigree

- Let $\widetilde{\mathbf{L}}_a$ be **sparsity preserving Cholesky factorization** of $\widehat{\mathbf{A}}^{aa}$.

- New fully orthogonalized ssSNP-BLUP: $\mathbf{H} = \mathbf{M}\widetilde{\mathbf{G}}\mathbf{M}'$, where

$$\mathbf{M} = (\mathbf{L}')^{-1}\begin{bmatrix}\mathbf{I}_n & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \sqrt{1-w}\widehat{\mathbf{M}}_a & \sqrt{w}\widehat{\mathbf{I}} & \sqrt{1-w}(\widehat{\mathbf{I}} - \widehat{\mathbf{M}}_a\widehat{\mathbf{M}}'_a)\widehat{\mathbf{L}}'_2\mathbf{Z}_m\end{bmatrix}$$

$\widetilde{\mathbf{G}} = \mathbf{I}$, and $\widehat{\mathbf{M}}_a = \widehat{\mathbf{L}}'_a(\widetilde{\mathbf{L}}'_a)^{-1}$.

- Main advantage of new formulation: size of numerical sparse Cholesky factorization matrix $\widetilde{\mathbf{L}}_a$ is number of non-genotyped ancestors ($\widehat{\mathbf{A}}^{aa}$) instead of all non-genotyped ($\mathbf{A}^{11}$).

- Four groups of random effects $\widetilde{\mathbf{u}}$: non-genotyped ancestors (a) have two sets of animals effects.

Sparse Cholesky $\widetilde{\mathbf{L}}_a$ of $\widehat{\mathbf{A}}^{aa}$

$$\widehat{\mathbf{A}}^{aa} = \widehat{\mathbf{L}}_a\widehat{\mathbf{L}}'_a = \widetilde{\mathbf{L}}_a\widetilde{\mathbf{L}}'_a$$

Luke NATURAL RESOURCES INSTITUTE FINLAND

Result 2: New ssSNP-BLUP Formulation using the Smaller Pedigree

- Let $\widetilde{\mathbf{L}}_a$ be **sparsity preserving Cholesky factorization** of $\widehat{\mathbf{A}}^{aa}$.

- New fully orthogonalized ssSNP-BLUP: $\mathbf{H} = \mathbf{M}\widetilde{\mathbf{G}}\mathbf{M}'$, where

$$\mathbf{M} = (\mathbf{L}')^{-1}\begin{bmatrix} \mathbf{I}_n & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \sqrt{1-w}\widehat{\mathbf{M}}_a & \sqrt{w}\widehat{\mathbf{I}} & \sqrt{1-w}(\widehat{\mathbf{I}} - \widehat{\mathbf{M}}_a\widehat{\mathbf{M}}_a')\widehat{\mathbf{L}}_2'\mathbf{Z}_m \end{bmatrix}$$

$\widetilde{\mathbf{G}} = \mathbf{I}$, and $\widehat{\mathbf{M}}_a = \widehat{\mathbf{L}}_a'(\widetilde{\mathbf{L}}_a')^{-1}$.

- Main advantage of new formulation: size of numerical sparse Cholesky factorization matrix $\widetilde{\mathbf{L}}_a$ is number of non-genotyped ancestors ($\widehat{\mathbf{A}}^{aa}$) instead of all non-genotyped ($\mathbf{A}^{11}$).

- Four groups of random effects $\widetilde{\mathbf{u}}$: non-genotyped ancestors (a) have two sets of animals effects.

Sparse Cholesky $\widetilde{\mathbf{L}}_a$ of $\widehat{\mathbf{A}}^{aa}$

$$\widehat{\mathbf{A}}^{aa} = \widehat{\mathbf{L}}_a\widehat{\mathbf{L}}_a' = \widetilde{\mathbf{L}}_a\widetilde{\mathbf{L}}_a'$$

New random effects $\widetilde{\mathbf{u}}$

$$\widetilde{\mathbf{u}} = \begin{bmatrix} \widetilde{\mathbf{u}}_n \\ \widetilde{\mathbf{u}}_a \\ \widetilde{\mathbf{u}}_{a+2} \\ \widetilde{\mathbf{u}}_m \end{bmatrix}$$

Result 3: Numerical Feasibility with Large Number of Genotyped

- Sparse Cholesky $\widetilde{\mathbf{L}}_a$ (of $\widehat{\mathbf{A}}^{aa} = \widetilde{\mathbf{L}}_a \widetilde{\mathbf{L}}_a'$) has **fill-ins** depending on pedigree and proportion of genotyped.
- **Memory requirements** of $\widetilde{\mathbf{L}}_a$ must remain manageable with very large number of genotyped.
- Investigated by simulating different proportions of genotyped animals in Nordic dairy cattle population.
- Pedigree of 9.2 million was "genotyped" with four **scenarios** varying number of genotyped.
- Memory usage: 13 times size of pedigree when 2.8 million genotyped, occupying 5 GB memory.

Result 3: Numerical Feasibility with Large Number of Genotyped

- Sparse Cholesky $\widetilde{\mathbf{L}}_a$ (of $\widehat{\mathbf{A}}^{aa} = \widetilde{\mathbf{L}}_a \widetilde{\mathbf{L}}_a'$) has **fill-ins** depending on pedigree and proportion of genotyped.

- **Memory reguirements** of $\widetilde{\mathbf{L}}_a$ must remain manageable with very large number of genotyped.

- Investigated by simulating different proportions of genotyped animals in Nordic dairy cattle population.

- Pedigree of 9.2 million was "genotyped" with four **scenarios** varying number of genotyped.

- Memory usage: 13 times size of pedigree when 2.8 million genotyped, occupying 5 GB memory.

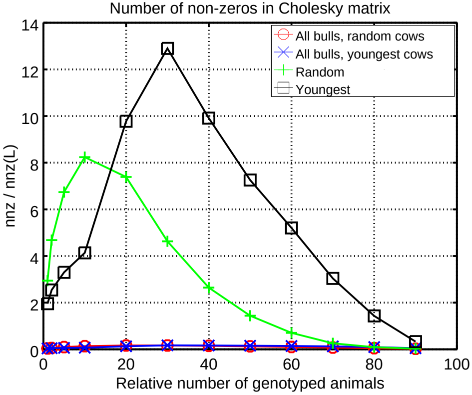
Result 3: Numerical Feasibility with Large Number of Genotyped

- Sparse Cholesky $\widetilde{\mathbf{L}}_a$ (of $\widehat{\mathbf{A}}^{aa} = \widetilde{\mathbf{L}}_a \widetilde{\mathbf{L}}'_a$) has **fill-ins** depending on pedigree and proportion of genotyped.
- **Memory reguirements** of $\widetilde{\mathbf{L}}_a$ must remain manageable with very large number of genotyped.
- Investigated by simulating different proportions of genotyped animals in Nordic dairy cattle population.
- Pedigree of 9.2 million was "genotyped" with four **scenarios** varying number of genotyped.
- Memory usage: 13 times size of pedigree when 2.8 million genotyped, occupying 5 GB memory.

NATURAL RESOURCES INSTITUTE FINLAND

Result 3: Numerical Feasibility with Large Number of Genotyped

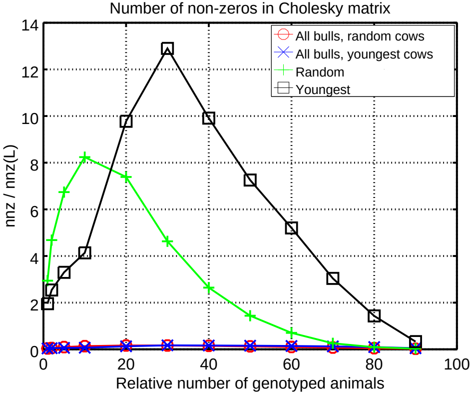
- Sparse Cholesky $\widetilde{\mathbf{L}}_a$ (of $\hat{\mathbf{A}}^{aa} = \widetilde{\mathbf{L}}_a \widetilde{\mathbf{L}}'_a$) has **fill-ins** depending on pedigree and proportion of genotyped.
- **Memory reguirements** of $\widetilde{\mathbf{L}}_a$ must remain manageable with very large number of genotyped.
- Investigated by simulating different proportions of genotyped animals in Nordic dairy cattle population.
- Pedigree of 9.2 million was "genotyped" with four **scenarios** varying number of genotyped.
- Memory usage: 13 times size of pedigree when 2.8 million genotyped, occupying 5 GB memory.



NATURAL RESOURCES INSTITUTE FINLAND

Result 3: Numerical Feasibility with Large Number of Genotyped

- Sparse Cholesky $\widetilde{\mathbf{L}}_a$ (of $\hat{\mathbf{A}}^{aa} = \widetilde{\mathbf{L}}_a \widetilde{\mathbf{L}}_a'$) has **fill-ins** depending on pedigree and proportion of genotyped.
- **Memory reguirements** of $\widetilde{\mathbf{L}}_a$ must remain manageable with very large number of genotyped.
- Investigated by simulating different proportions of genotyped animals in Nordic dairy cattle population.
- Pedigree of 9.2 million was "genotyped" with four **scenarios** varying number of genotyped.
- Memory usage: 13 times size of pedigree when 2.8 million genotyped, occupying 5 GB memory.



Luke
NATURAL RESOURCES INSTITUTE FINLAND

Conclusions

1. All fully orthogonalized ssSNP-BLUPs were shown to share same convergence properties with respect to iterative solution methods.

2. New ssSNP-BLUP formulation was presented in which Single-step relationship matrix $\mathbf{H}$ was expressed using smaller pedigree of genotyped and their ancestors.

3. Feasibility of ssSNP-BLUP was investigated with very large number of genotyped individuals.
   - Memory requirements were found to be manageable indicating that ssSNP-BLUP remains numerically efficient when number of genotyped animals increases.

Conclusions

1 All fully orthogonalized ssSNP-BLUPs were shown to share same convergence properties with respect to iterative solution methods.

2 New ssSNP-BLUP formulation was presented in which Single-step relationship matrix $\mathbf{H}$ was expressed using smaller pedigree of genotyped and their ancestors.

3 Feasibility of ssSNP-BLUP was investigated with very large number of genotyped individuals.

- Memory requirements were found to be manageable indicating that ssSNP-BLUP remains numerically efficient when number of genotyped animals increases.

NATURAL RESOURCES INSTITUTE FINLAND

Conclusions

1 All fully orthogonalized ssSNP-BLUPs were shown to share same convergence properties with respect to iterative solution methods.

2 New ssSNP-BLUP formulation was presented in which Single-step relationship matrix $\mathbf{H}$ was expressed using smaller pedigree of genotyped and their ancestors.

3 Feasibility of ssSNP-BLUP was investigated with very large number of genotyped individuals.

- Memory requirements were found to be manageable indicating that ssSNP-BLUP remains numerically efficient when number of genotyped animals increases.

NATURAL RESOURCES INSTITUTE FINLAND

Conclusions

1. All fully orthogonalized ssSNP-BLUPs were shown to share same convergence properties with respect to iterative solution methods.

2. New ssSNP-BLUP formulation was presented in which Single-step relationship matrix $\mathbf{H}$ was expressed using smaller pedigree of genotyped and their ancestors.

3. Feasibility of ssSNP-BLUP was investigated with very large number of genotyped individuals.
   - Memory requirements were found to be manageable indicating that ssSNP-BLUP remains numerically efficient when number of genotyped animals increases.

Conclusions

1. All fully orthogonalized ssSNP-BLUPs were shown to share same convergence properties with respect to iterative solution methods.

2. New ssSNP-BLUP formulation was presented in which Single-step relationship matrix $\mathbf{H}$ was expressed using smaller pedigree of genotyped and their ancestors.

3. Feasibility of ssSNP-BLUP was investigated with very large number of genotyped individuals.
   - Memory requirements were found to be manageable indicating that ssSNP-BLUP remains numerically efficient when number of genotyped animals increases.

Luke

NATURAL RESOURCES
INSTITUTE FINLAND