

Single-step method for genetic evaluation

- Ismo Strandén, MTT
- Ole Christensen, AU

- Main messages:
- Conceptually simple
- Computationally feasible in practice
- Tweaks are needed, but have (mostly) been solved

Single-step method

- **Traditional genetic evaluation**: pedigree relationship matrix, BLUP, multivariate models, etc
- **Genomic evaluation in cattle**: three steps
 1. Compute deregressed EBV (pseudo phenotype)
 2. Genomic prediction model
 3. Blending
- **Single-step**: similar to traditional genetic evaluation, with difference being that pedigree relationship is replaced by combined relationship (both genomic and pedigree).

Single step method

- Usual additive relationship matrix constructed from pedigree is replaced by a combined relationship matrix.
- Combined relationship matrix extends GBLUP to non-genotyped animals.
- Combined relationship matrix has a sparse inverse (good for fast computing !)
- BLUP, AI-REML, multivariate model, genetic groups, heterogeneous variance, maternal effects, etc.

Single-step in Danish pig breeding



- Genomic selection in pig project: a collaboration between VSP and AU
- Single-step used in routine evaluation for Duroc, Landrace and Yorkshire from October 2011.
- Single-step is more accurate than traditional method.
- Transition from traditional genetic evaluation to genomic evaluation went fairly smooth: all the old models still used, only difference is combined relationship.
- Now: VSP are improving old models
- Research: GS for crossbred performance

Single-step: Complete model solving (Nordic red)



- Routine TD evaluation, no heterogeneous var. correction
- Production traits:
 - runs took ~12 h
 - For maximum 1500 iterations rounds there was no time and convergence difference in TD and in single-step TD

		ROUND	CD
EBV _F	12h	1500	0.1077E-02
GEBV _F	12h 27 min	1500	0.1159E-02

- Udder health traits:
 - Runs took 17 – 31 h; single step models needed ~1600 iteration rounds and 14 h more for convergence than models without genomic data

		ROUND	CD
EBV _F	17h	2199	0.9892E-04
GEBV _F	31 h	3757	0.9944E-04

- Previous exercise used \mathbf{G}^{-1}
- Can be relaxed: single-step using \mathbf{G} or its components

Nordic red and single-step

Nordic red is composed of three 'populations'

- Different populations have different allele frequencies (AF)
- Different genomic relationship matrices (**G**):

Need to account **breed composition** of animal
- similar to accounting for multiple breeds
in traditional evaluation

Relationship matrix **A** using pedigree traces information to
base population

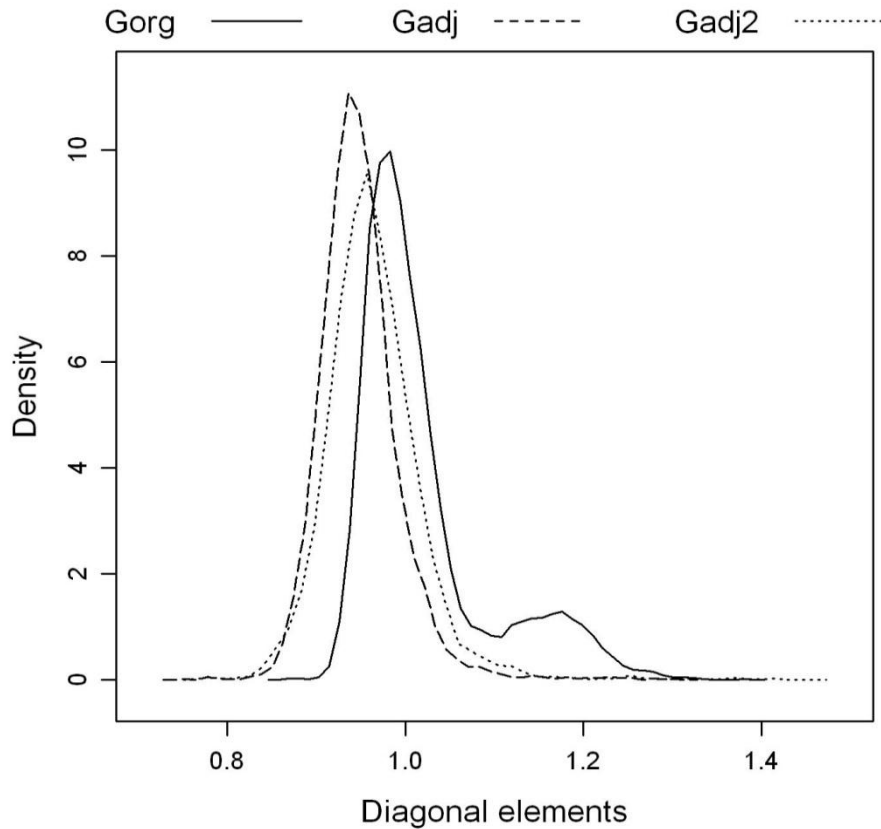
Used allele frequencies define base population in **G**.

- Use base population AF instead of current (genotyped)
population AF

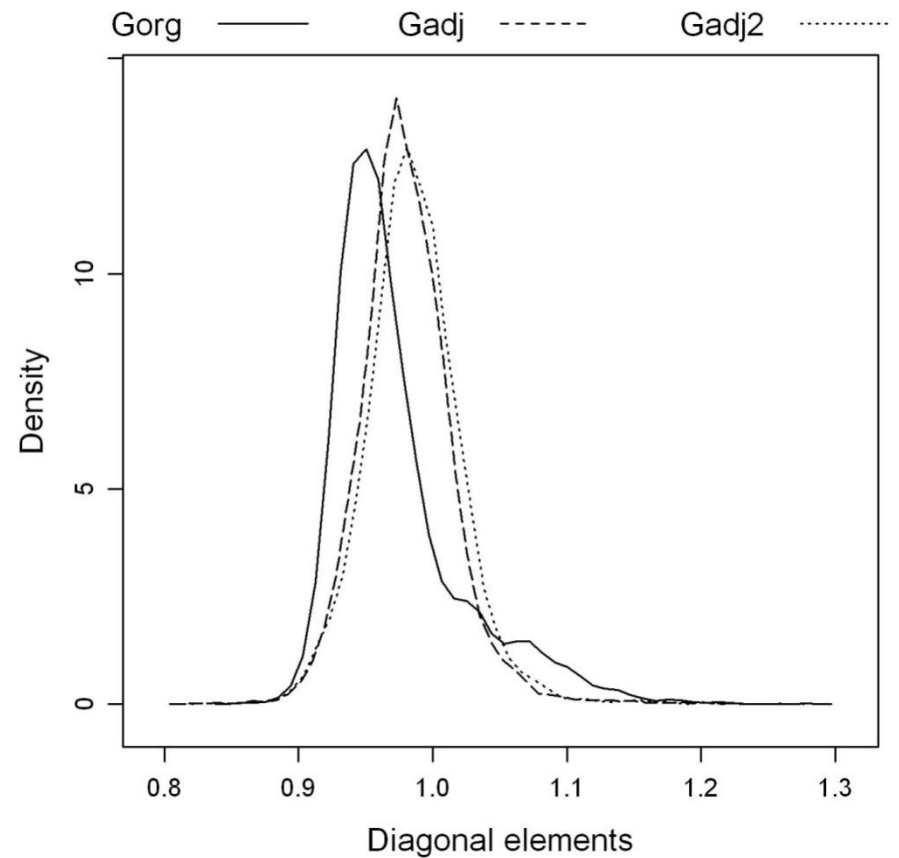
Diagonal elements of **G** matrix

Gorg: standard **G**

Gadj, Gadj2: account breed differences in **G**



Current population



Base population

Weighing matrices in single step

- Standard: $\mathbf{H}^{-1} = \mathbf{A}^{-1} + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}^{-1} - \mathbf{A}_{22}^{-1} \end{bmatrix}$
- Mixing \mathbf{G} with \mathbf{A} : $\mathbf{G}_{\alpha} = \alpha\mathbf{G} + (1-\alpha)\mathbf{A}_{22}$
- Give weights to inverses (I. Misztal):

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

Values of 1.5 and 0.6 by I. Misztal make magic.

Validation results (G-BLUP)

Target for b_1 is one (1), and validation reliability R^2 as high as possible

Base population AF	Milk			Protein			Fat		
	b_1	R^2	b_0	b_1	R^2	b_0	b_1	R^2	
Gorg	84.06	0.770	0.351	76.11	0.895	0.378	82.81	0.845	0.443
Gorg _(0.5, 0.5)	81.77	0.815	0.327	71.92	0.970	0.355	80.27	0.905	0.408
Gorg _(1.5, 0.6)	83.60	0.883	0.357	76.37	0.983	0.371	82.93	0.964	0.443
Gadj	84.84	0.750	0.341	77.49	0.875	0.353	83.43	0.837	0.434
Gadj _(0.5, 0.5)	82.64	0.786	0.319	74.53	0.898	0.335	80.88	0.886	0.400
Gadj _(1.5, 0.6)	84.59	0.846	0.348	79.19	0.894	0.345	83.46	0.949	0.435

Conclusions

- Genomic enhanced evaluation possible through use of genomic relationship matrix **G** in current evaluations
- Genomic data brings new challenge: **G** vs. **A**
 - Complexities of population structure
- Population structure in **G** through breed proportions
- How to weigh **G** and **A**, and their inverses?