

Assignment 3 - Single-step GBLUP

Download and extract **ssgblup.zip** file from wiki. The data for this lab was created using simulation for a single trait animal model. Simulation was done using QMSim (Sargolzaei, M. and F. S. Schenkel, 2009)

The simulated phenotype were generated using the following model:

$$\text{Phenotype} = \text{mean} + \text{true_ebv} + \text{residual}$$

Parameters:

Mean = 1.0

True variances:

direct genetic = 0.25

residual = 0.72

Description of files

data.txt:

1: Animal ID

2: Sire ID

3: Dam ID

4: Generation number

5: Sex code

6: Number of male progeny

7: Number of female progeny

8: Inbreeding coefficient

9: Homozogosity

10: Phenotype

11: Simulated residual

12: Polygenic effect

13: True EBV

14: Internal EBV from QMSim

15: Mean (column of ones to fit mean effect in BLUPF90)

ped.txt:

1: animal ID

2: sire ID

3: dam ID

snps.txt:

1: animal ID

2: marker information

1. From raw data modify renumf90 parameter file (renlab.par) according to the data file and to fit the following model for genomic selection:

$$y = \text{mean} + \text{animal} + e$$

2. Run renumf90 program to renumber data, pedigree file, and marker data.

3. Check the renf90.par, renf90.dat, and renaddxx.ped. From the renaddxx.ped file, identify genotyped animals, and check with wiki (<http://nce.ads.uga.edu/wiki/doku.php?id=readme.renumf90>) the content of each column.
4. Estimate variance components considering and ignoring marker information. From the airemlf90 output find the following statistics: number of genotyped animals, number of SNP markers
5. Run blupf90 without marker information using estimated variance components. Now run blupf90 using genomic information and compare cpu time and solutions. Obs: Check wiki to see how to read an external or pre-computed genomic matrix.
6. Validation on young candidates (individuals from 10th generation with no phenotypes).
 - a) Remove the phenotypic information from the 10th generation and obtain solutions from a model with marker information and with no marker information.

Hint: if generation column is number 4, new data can be created using the AWK Linux tool

```
awk "$4!=10" renf90.dat > renf90.dat.pred
```

- b) Compare correlations with true breeding values for genetic additive direct effect. Hint: have renumf90 passing to the data a column containing "generations".