GIBBSTHUR

Luis Varona[1] and Andrés Legarra[2]

[1] Departamento de Anatomía, Embriología y Genética Animal. Instituto Agroalimentario de Aragón (IA2). Universidad de Zaragoza. Zaragoza , SPAIN.

[2] Génétique Physiologie et Systèmes d'Elevage (GenPhySE), Institut National de la Recherche Agronomique de Toulouse, Castanet-Tolosan, France

The software GIBBSTHUR provides a Bayesian estimation of variance components and predictions of breeding values for ranking traits based on a Thurstonian approach (Gianola and Simianer, 2005). The GIBBSTHUR software was developed in FORTRAN 90 and based on the TM software (Legarra et al., 2010). It can be executed in UNIX, OSX, or WINDOWS environments. In addition, it also provides estimates of variance components and prediction of breeding values for additional threshold and continuous traits.

INPUT FILES

Parameters.txt

The execution requires a parameter file denoted as 'parameters.txt' with the following structure:

```
dat.txt ! name of the data file
genea.txt ! name of the pedigree file
10   ! Number of effects in the model (including additive genetic effects)
1 ! number of covariables
2 ! Number of genetic groups
3 ! Number of total traits (continous + threshold + thrustonian)
1 ! Number of threshold traits
2 ! Categories for the threshold traits (one number for each threshold trait)
1 ! Number of ranking traits
6 ! Location in the data file of the event effects (must be equal to number of ranking traits and it must be a systematic
                            effect, not random)
3 ! Number of random environmental effects
1 ! Number of genetic effects
1 3 10 50 50 50 100 100 100 1000 ! Levels for each effect (for animals do not include genetic group). First, as many
                            1s as the number of covariable. Second, the number of levels of the systematic
                            effects. Third, the number of levels of the random effects and, finally, the number
                            of levels of the genetic effects without the genetic group. The number of elements
                            must be identical to the number of effects.
1 1 1   1   0   0   1   1   1   1 !Model for trait 1 (1 the effect must be included or 0 the effect must not be included
                            for trait 1, the number of elements must be identical to the number of effects)
0 1 1   0   1   0   1   1   1   1 !Model for trait 2.
0 1 1   0   0   1   1   1   1   1 ! Model for trait 3
105000 ! Number of iterations of the Gibbs sampler
5000 ! Number of iterations to discard (burn in)
10 ! Thin interval between iterations.
```

Pedigree file

The pedigree must have the following structure for individuals with unknown sire and dam

1 1001 1002 ! "individual" " genetic group" "genetic group"

For the individuals with only one ancestor known.

2 1 1002 ! "individual" " sire" " genetic group"

And for the individuals with known sire and dam

3 1 2 ! "individual" "sire" "dam"

Genetic groups must be coded as number of animals in the pedigree + genetic group. Therefore, if there are 1000 animals in the pedigree, the first genetic group must be coded as 1001, the second as 1002, and so on.

Data file

The data file must have one row for each data that must be coded as:

130.3   2   4   5   7   12   90   39   87   991 56.0   2   2

Where the first elements correspond the levels of the effects of the model. First, it must include the value of the covariates:

130.3   2   4   5   7   12   90   39   87   991 56.0   2   2

Second, the levels of the systematic effects

130.3   2   4   5   7   12   90   39   87   991 56.0   2   2

Third, the levels of the random environmental effects

130.3   2   4   5   7   12   90   39   87   991 56.0   2   2

The levels for the animal genetic effects

130.3   2   4   5   7   12   90   39   87   991 56.0   2   2

The phenotype for the continuous traits

130.3   2   4   5   7   12   90   39   87   991 56.0   2   2

It must be a positive real number. If it is 0, it will be treated as a missing data and if it is a negative number, it will be considered as censored.

The phenotype for the threshold traits

130.3   2   4   5   7   12   90   39   87   991 56.0   2   2

It must be a positive integer between 1 and the number of categories defined in the parameter file. If it is 0, it will be treated as a missing data.

The phenotype for the ranking traits

130.3   2   4   5   7   12   90   39   87   991 56.0   2   2

It must be a positive integer. If it is 0, it will be treated as a missing data.

OUTPUT FILES

The program generates several output files denotes as "results.txt", "solutions.txt" and "samples.txt".

Results.txt

This file is generated every 100 iterations after burn-in, and gives the present estimates (posterior mean and standard deviation) for variance components, heritabilities and correlations. This is an example:

```
Iteration number:    1000
 Burn-in:     100

 Average additive variance
     0.96858025    0.16758666
     0.16758666    0.32270594
 Sd Additive variance
     0.03979433    0.01900286
     0.01900286    0.01875754

 Average residual variance
     1.99950363   -0.29061013
    -0.29061013    1.00000000
 Sd residual variance
     0.02138472    0.01170423
     0.01170423    0.00000000

 Average h2 and additive correlation
     0.32623058    0.29997320
     0.29997320    0.24382196
 Sd h2 and additive correlation
     0.00949030    0.03260313
     0.03260313    0.01073970

 Average he2 and residual cor
     0.67376942   -0.20551467
    -0.20551467    0.75617804
 Sd he2 and residual cor
     0.00949030    0.00804667
     0.00804667    0.01073970
```

Solutions.txt

The "solutions.txt" file provides the posterior estimates (and standard deviation) for the covariates, systematic effect and random environmental effects for all traits. Here, it is presented the head of the file for an example with three traits:

```
    0.03659825    0.08168124    1.08514199    0.48101735   -1.11392845    0.10974366
    0.04142401    0.06885735    1.08458681    0.47603070   -1.13280832    0.07971237
   -0.10875979    0.06141084    1.21138684    0.48711404   -0.81257421    0.08645512
    0.58074742    0.10670308    0.15238137    0.36754928   -0.29475642    0.13664090
    0.66822275    0.06870989    0.06178150    0.36355990   -0.45707502    0.08910092
    0.52330409    0.08304296    0.08403646    0.35529649   -0.19134721    0.11452671
    0.49168973    0.07219667   -0.01751253    0.36012791   -0.10949482    0.10980555
    0.45768417    0.07221291    0.09889356    0.35959759   -0.04073952    0.11113870
    0.45295796    0.06154615    0.01427350    0.37230582   -0.01474124    0.09941542
    0.56648412    0.07611892   -0.02076152    0.37584713   -0.20993120    0.12628192
```

Breeding.txt

The "breeding.txt" provides the posterior estimates (and standard deviations) of the breeding values and the genetic group estimates (at the end of the file). Here, it is presented the head of the file for an example with three traits:

```
-0.01232285    0.27515434    0.30658415    0.42668966    0.01490627    0.48916036
-0.14750920    0.24420097    0.14267855    0.38617217    0.28007673    0.43124017
-0.18271679    0.25463772    0.31545951    0.38548920    0.35296424    0.47305095
 0.00836714    0.25768420    0.01098382    0.40786698    0.00150559    0.45070685
 0.02505360    0.28575738    0.03533080    0.40274853    0.01015014    0.52537369
 0.23087852    0.30327152    0.05058291    0.38698281   -0.37254057    0.51105771
-0.04440926    0.04273097    0.30248813    0.14298000    0.07289879    0.03900783
 0.11019323    0.03417746   -0.74217173    0.11486472   -0.21107539    0.10693949
 0.11156882    0.04101711    0.10443986    0.10326349   -0.07209824    0.07450198
-0.02590168    0.05617096   -0.00998250    0.33528523    0.15406350    0.08817856
```

Samples.txt

The file "samples.txt" contains the results of the variance and covariance component for all gibbs sampler iterations. Each row includes: First, the additive genetic variance components. Second, the permanent environmental variance components and, finally, the residual variance components. Here it is an example with two traits and two variance components (additive genetic and residual). The first trait is continuous and the second is a ranking trait.

Va(1,1)	Va(1,2)	Va(2,2)	Ve(1,1)	Ve(1,2)	Ve(2,2)
1.54872163	0.06100350	0.36983780	1.92209426	-0.25681209	1.00000000
1.27037788	0.04836286	0.21047803	2.00249158	-0.37877683	1.00000000
1.06602889	0.04834945	0.15201653	1.98605149	-0.40349827	1.00000000
1.03443714	0.04707799	0.12287187	1.97199550	-0.39523052	1.00000000
1.03198105	0.05058611	0.10927489	1.96162081	-0.37314882	1.00000000
1.02881598	0.04481526	0.09266988	1.97259564	-0.39521293	1.00000000
0.99512659	0.04618707	0.08993458	1.97650036	-0.39495576	1.00000000
0.96676428	0.06091350	0.08824477	1.97788453	-0.36404886	1.00000000

REFERENCES

Gianola, D., Simianer, H. (2006). A thurstonian model for quantitative genetic analysis of ranks: a Bayesian approach. Genetics 174: 1613-1624.

Legarra A., Varona, L., López de Maturana, E. 2010. Threshold Model (TM). http://genoweb.toulouse.inra.fr/~alegarra/tm_folder/