

Imposing simple penalties in WOMBAT to improve estimates of genetic parameters

1 Introduction

Penalized REML estimation has been proposed to reduce sampling variation and thus reduce 'loss' in estimates of covariance matrices. Meyer (2016) proposed penalties on canonical eigenvalues or correlations, derived assuming Beta distributions targeted at shrinking eigenvalues towards their mean and correlations towards zero or their phenotypic counterparts. Further, it was suggested to regulate the strength of penalization by choosing an appropriate value representing the effective sample size of the prior. Values between 4 and 10 were suggested to achieve mild penalties with worthwhile reductions in loss with little risk of detrimental effects. This note describes how to invoke penalized estimation for these penalties in WOMBAT.

2 Specification in WOMBAT

Instructing WOMBAT to impose a penalty on the likelihood only requires an entry in SPECIAL block in the parameter file in addition to the specifications needed for a standard multivariate analysis.

This should consist of a single line, with the following, space separated entries:

- a) The code PENALTY at the start of the line.
- b) A code for the type of penalty to be fitted. Currently, three different codes are recognized:
 - KANEIG to impose a penalty on canonical eigenvalues (only valid for models without any additional random effects).
 - KORREL to impose a penalty on standard correlations.
 - PACORR to impose a penalty on partial correlations.
- c) Optional (for PACORR or KORREL only): The code PHENV. This specifies that correlations are to be shrunk towards their phenotypic counterparts, calculated from the current estimates of covariance components. If this code is omitted, all correlations are shrunk towards zero.
- d) For PACORR and KORREL only: The name of the random effect for which the correlations are to be penalized.
- e) A REAL variable giving the effective sample size for the Beta distribution.

For example, to impose a penalty on canonical eigenvalues for an effective sample size of 8.0:

```
SPECIAL
...
PENALTY KANEIG 8.0
...
END
```

Similarly, to shrink both partial genetic and partial residual correlations towards the corresponding phenotypic values, we need to include a separate line for each. If `animal` is the name of the random effect representing additive genetic effects,

```
SPECIAL
  ...
  PENALTY PACORR PHENV animal 8.0
  PENALTY PACORR PHENV residual 8.0
  ...
END
```

Convergence: Penalized estimation is only implemented for the average information algorithm and derivative-free maximization. The run option `--force` is supplied to enforce derivative-free search steps after the average information algorithm is deemed to have converged or stops because no increase in likelihood can be achieved.

3 Worked example

Use of WOMBAT with the simple penalties described is illustrated in Example 19. Data used are simulated records for 9 traits, recorded for all individuals, for a paternal half-sib design comprised of 400 independent sire families of size 10. The model of analysis is a simple animal model, fitting means as the only random effects.

Part “A” shows a standard, unpenalized multivariate analysis for a simulated data set. Part “B” demonstrates the increase in the number of iterates required for the same analysis when parameterising the analysis to the elements of the canonical decomposition. Part “C” contains the analysis imposing a penalty on canonical eigenvalues and part “D” illustrates the effect of a penalty on partial genetic correlations, shrinking them towards the phenotypic value. An example parameter file to simulate other data sets is given in directory “0”. Directory “Z” contains the script and parameter files for an example of a simulation study to investigate the effect of penalization (a simple FORTRAN program is supplied to summarize results).

References

Meyer K. Simple penalties on maximum likelihood estimates of genetic parameters to reduce sampling variation. *Genetics* 203 (2016) 1885–1900. doi: [10.1534/genetics.115.186114](https://doi.org/10.1534/genetics.115.186114).