"RRGIBBS" – A PROGRAM FOR SIMPLE RANDOM REGRESSION ANALYSES VIA GIBBS SAMPLING

K. Meyer

Animal Genetics and Breeding Unit, University of New England (jointly with NSW Agriculture), Armidale, NSW 2351, Australia

INTRODUCTION

Random regression (RR) models are a popular choice for the analysis of longitudinal data or 'repeated' records. Programs for estimation of the corresponding covariance functions via restricted maximum likelihood (REML) are available (e.g. Gilmour *et al.*, 1999; Meyer, 1998), but high computational demands of REML analyses severely limit their feasibility. Bayesian analysis using Gibbs sampling provides an alternative which is markedly simpler to implement and requires considerably less memory than REML, thus facilitating large scale analyses.

OBJECTIVE

RRGIBBS allows the Bayesian estimation of (co)variance components for RR models involving a polynomial regression on a continuous covariable, for multiple random effects and heterogeneous measurement error variances via Gibbs sampling.

COMPUTING METHODS

RRGIBBS generates a Markov Chain for the parameters of a RR model, sampling from their fully conditional posterior distributions. This comprises blockwise sampling from multivariate normal distributions for location parameters, with block size equal to the order of polynomial fit for regression coefficients, and sampling from inverted Wishart and χ^2 distributions for covariances among RR coefficients and measurement error variances, respectively. Full details are given, for instance, by Jamrozik and Schaeffer (1997) and Rekaya *et al.* (1999) for RR analyses of test-day records in dairy cattle. These extend readily to models involving additional random effects.

Model of analysis. RRGIBBS allows a range of simple RR models to be fitted. Model specification is through a parameter file. Currently, RRGIBBS performs univariate analyses and accommodates regression on a single, continuous covariable ("meta-meter") only. By default, regressions on Legendre polynomials of the meta-meter are fitted, but RRGIBBS can also accommodate other functions, supplied by the user.

Fixed effects allowed include standard cross-classified effects and covariables, as well as fixed regression(s) on Legendre polynomials of the meta-meter to model trends in means. The latter can be fitted as a nested regression within classes of fixed effects. Sets of regression coefficients for a chosen order of polynomial are fitted for levels of random effects. Orders of fit can be different for each random factor, and additional random effects can be restricted to a subset of the complete range of values of the meta-meter. Levels of random effects are assumed to be distributed either proportional to the numerator relationship matrix (genetic effects) or independently. Covariances between random sources of variations are considered zero. Residual or measurement error variances are assumed to be independently distributed, allowing for homogeneous variance or heterogeneous variances with changes in variance modelled as a step function of the meta-meter.

Options. The run time behaviour of RRGIBBS is modified by a number of command line options specifying, for instance, the number of samples to be drawn, the "burn-in" period, the interval at which complete samples are saved, whether this is a continuation run, and the summary information to be produced.

Output. The primary output from RRGIBBS is a binary file with the successive samples of (co)variance components drawn. Similar output for a selected range of location parameters is optional.

Limited summary information is produced. Estimates of covariance matrices among RR coefficients and measurement error variances are obtained as means over samples, discarding samples in the "burn-in" phase specified. Approximate 95% highest posterior density regions are determined as points between highest/lowest of the 2.5% lowest/highest valid sample values and the next value higher/lower than it (Chen and Shao, 1999). Estimates of variance components and variance ratios for the values of the meta-meter represented in the data are computed from the estimates and written to a separate, formatted file. Optionally, formatted files with individual samples for each parameter and running means and files with frequency distributions of samples can be written out for further, post-Gibbs analyses.

USE

RRGIBBS has been used for RR analyses of large sets of weight records from commercial beef cattle, fitting sets of RR coefficients for direct and maternal genetics and direct and maternal permanent environmental effects (Meyer, 2002).

COMPUTING ENVIRONMENT

RRGIBBS is written in standard FORTRAN 95 and is self-contained, except for some public domain routines to generate random samples from the normal and χ^2 distribution. Compiled versions for Linux (Redhat) and Compaq Alpha stations are available.

AVAILABILITY

RRGIBBS is available to the scientific community free of charge. Users are required, however, to credit its use in any publications. Material available comprises source code (complete with 'Makefile' for a UNIX environment), a manual and a worked example, as well as pre-compiled programs for several computing environments. RRGIBBS is available only by downloading the material required from its web page : **http://agbu.une.edu.au/∼kmeyer/rrgibbs.html**

REFERENCES

Chen, M.-H. and Shao, Q.-M. (1999) *J. Comp. Graph. Stat.* **8** : 69–92.

Gilmour, A., Cullis, B. R., Welham, S. J. and Thompson, R. (1999) Mimeo 177pp.

Jamrozik, J. and Schaeffer, L. R. (1997) *J. Dairy Sci.* **80** : 762–770.

Meyer, K. (1998) *Proc. 6th WCGALP* **27** : 465–466.

Meyer, K. (2002) ibid.

Rekaya, R., Carabaño, M. J. and Toro, M. A. (1999) *Livest. Prod. Sci.* **57** : 203–217.