Efficiency of alternative MCMC strategies illustrated using the reaction norm model

M. Shariati1, 2, 3 & D. Sorensen1

1 Department of Genetics and Biotechnology, Faculty of Agricultural Sciences, University of Aarhus, PB 50, DK-8830 Tjele, Denmark
2 Department of Large Animal Sciences, Faculty of Life Sciences, University of Copenhagen, Ridebanevej 12, DK-1870 Frederiksberg C, Denmark
3 Department of Animal Science, University of Yasuj, 75914-353 Yasuj, Iran

Introduction

In animal breeding the reaction norm model (RNM) is often used to study genotype by environment interactions (G × E). This model is a random regression of phenotypes on environmental gradients and it predicts a reaction norm specific for each individual (Kolmodin et al. 2002). The reaction norm or breeding value of an animal has two parts: (i) an environment-independent part, called the ‘genetic level’, and (ii) an environment-dependent part that includes the slope and is postulated

Keywords
MCMC strategies; posterior correlation, reaction norm

Correspondence
M. Shariati, Department of Genetics and Biotechnology, Faculty of Agricultural Sciences, University of Aarhus, PB 50, DK-8830 Tjele, Denmark.
E-mail: mohammad.shariati@agrsci.dk

Received: 17 February 2007; accepted: 15 November 2007

Summary

The Markov chain Monte Carlo (MCMC) strategy provides remarkable flexibility for fitting complex hierarchical models. However, when parameters are highly correlated in their posterior distributions and their number is large, a particular MCMC algorithm may perform poorly and the resulting inferences may be affected. The objective of this study was to compare the efficiency (in terms of the asymptotic variance of features of posterior distributions of chosen parameters, and in terms of computing cost) of six MCMC strategies to sample parameters using simulated data generated with a reaction norm model with unknown covariates as an example. The six strategies are single-site Gibbs updates (SG), single-site Gibbs sampler for updating transformed (a priori independent) additive genetic values (TSG), pairwise Gibbs updates (PG), blocked (all location parameters are updated jointly) Gibbs updates (BG), Langevin-Hastings (LH) proposals, and finally Langevin-Hastings proposals for updating transformed additive genetic values (TLH).

The ranking of the methods in terms of asymptotic variance is affected by the degree of the correlation structure of the data and by the true values of the parameters, and no method comes out as an overall winner across all parameters. TSG and BG show very good performance in terms of asymptotic variance especially when the posterior correlation between genetic effects is high. In terms of computing cost, TSG performs best except for dispersion parameters in the low correlation scenario where SG was the best strategy. The two LH proposals could not compete with any of the Gibbs sampling algorithms. In this study it was not possible to find an MCMC strategy that performs optimally across the range of target distributions and across all possible values of parameters. However, when the posterior correlation between parameters is high, TSG, BG and even PG show better mixing than SG.