An illustration of the use of reparameterisation methods for improving MCMC efficiency in crossed random effect models.

William J. Browne

University of Nottingham

1 Introduction

In this article we illustrate how the performance of MCMC methods can be improved by particular reparameterisation schemes from the MCMC literature. We use as an example a four way crossed classification model fitted to the Wytham Woods great tit dataset of bird nesting attempts recently analysed in Browne *et al.* (2004). We describe two methods, hierarchical centering (Gelfand, Sahu and Carlin, 1995) and parameter expansion (Liu, Rubin and Wu, 1998) and show how they can improve the efficiency of a Gibbs sampling algorithm. We then show how the methods can be combined to create a more efficient MCMC estimation algorithm.

2 Wytham Woods dataset

Random effect modelling can be used in many application areas and for our example we have chosen a dataset from bird ecology. Wytham Woods in Oxfordshire is a site where a long-term individual based study of great tits containing data of over 50 years has been carried out. Here we consider a dataset of 4165 observations taken over a 34-year period (1964-1997). Each observation is a breeding attempt for a pair of great tits and the dataset contains six response variables per attempt. From a substantive point of view interest lies in the relative importance of the genetic and environmental effects and Browne *et al.* (2004) consider fitting a multivariate response cross-classified model to the dataset.

For our purposes we will consider just one of the response variables, clutch size and examine the univariate normal response model fitted to it in Browne *et al.* (2004). The model can be written using the notation of Browne *et al.* (2001) as

$$\begin{aligned} y_i &= \beta_0 + u^{(2)}_{female(i)} + u^{(3)}_{male(i)} + u^{(4)}_{nestbox(i)} + u^{(5)}_{year(i)} + e_i, \\ u^{(2)}_{female(i)} &\sim N(0, \sigma^2_{u(2)}), u^{(3)}_{male(i)} \sim N(0, \sigma^2_{u(3)}), \\ u^{(4)}_{nestbox(i)} &\sim N(0, \sigma^2_{u(4)}), u^{(5)}_{year(i)} \sim N(0, \sigma^2_{u(5)}), e_i \sim N(0, \sigma^2_e), \\ \beta_0 &\propto 1, \sigma^2_{u(k)} \sim \Gamma^{-1}(\epsilon, \epsilon), k = 2, \ldots, 5, \sigma^2_e \sim \Gamma^{-1}(\epsilon, \epsilon). \end{aligned}$$

where y_i is the clutch size for observation i. The four sets of u's are random effects with the superscripts identifying the respective higher levels. The subscripts are functions that for each observation identify the corresponding higher level unit. We have added diffuse priors for all unknown parameters with $\epsilon = 10^{-3}$.

One difficulty in fitting and interpreting results from this model is the fact that for each male/female bird we only have a few observations (in the range 1-6 with a median of 1 observation per bird). This model can be fitted using a standard Gibbs sampling algorithm in either the MLwiN or WinBUGS software packages however the standard algorithm produces poorly mixing chains for many parameters.

We will here use MLwiN as it is faster and run for 50,000 iterations after a burnin of 5,000 iterations. This took 519 seconds in MLwiN. Table 1 gives estimates and effective sample sizes (ESS) per minute (see Kass *et al.* 1998 for details) for the fixed effect, β_0 and the 5 variances in the model. From the ESS values we see that for β_0 we have less than 100 effectively independent samples per minute and several of the variances have poor ESS values with the between males variance being the worst at 4.6 ESS per minute. We will therefore focus on β_0,$\sigma^2_{u(3)}$ and one of the male residuals $u^{(3)}_1$. The left hand columns of figures 1 and 2 give additional output from the MCMC chains, with figure 1 showing the actual chains and figure 2 showing plots of consecutive iterations plotted against each other. From figure 2 we see 'cigar' shapes that are typical for poorly mixing chains for β_0, and $\sigma^2_{u(3)}$. The plot for $u^{(3)}_1$ doesn't show similar evidence but the traces in figure 1 shows the clear dependence between $\sigma^2_{u(3)}$ and $u^{(3)}_1$.

Table 1: A comparison between fitting the original model in MLwiN and the reparameterised model in WinBUGS

Parameter	MLwiN Estimate (95% CI)	WinBUGS Estimate (95% CI)	MLwiN ESS/min	WinBUGS ESS/min
β_0	8.805 (8.589,9.025)	8.810 (8.593,9.024)	85.3	896.1
$\sigma^2_{u(5)}$ (Year)	0.365 (0.215,0.611)	0.365 (0.215,0.607)	3895.4	909.7
$\sigma^2_{u(4)}$ (Nestbox)	0.107 (0.059,0.158)	0.109 (0.061,0.162)	88.1	135.1
$\sigma^2_{u(3)}$ (Male)	0.045 (0.001,0.166)	0.070 (0.000,0.178)	4.6	14.6
$\sigma^2_{u(2)}$ (Female)	0.975 (0.854,1.101)	0.968 (0.848,1.089)	394.0	224.2
σ^2_e (Observation)	1.064 (0.952,1.173)	1.046 (0.935,1.157)	14.0	37.4

3 Reparameterisation methods

We will consider two methods from the MCMC literature to improve two aspects of the performance of our MCMC algorithm. Firstly hierarchical centering (Gelfand, Sahu and Carlin, 1995) can be used to improve the mixing of the β_0 parameter. Hierarchical centering is typically used in nested models but can also be used in crossed models by choosing one of the possible hierarchies. We will choose the year hierarchy as its variance had the largest ESS. We reparameterise the model by replacing the year random effects $u^{(5)}_j$ (with mean 0) by a set of random parameters $\beta^{(5)}_j$ that have mean β_0.

Our second method, parameter expansion (Liu, Rubin and Wu, 1998) is going to be used to improve the mixing of both $\sigma^2_{u(3)}$ and $u^{(3)}_1$. In fact we will use the parameter expansion method on three of the four higher classifications. The reason that we see poor mixing of $\sigma^2_{u(3)}$ and $u^{(3)}_1$ is that there is a large posterior probability of $\sigma^2_{u(3)}$ being close to zero. When the chain of $\sigma^2_{u(3)}$ moves close to zero all the random effects $u^{(3)}_j$ are also close to zero and the chain gets stuck for long periods in this area of the posterior.

Parameter expansion works by introducing an additional parameter α_3 that multiplies the random effects. This additional parameter is not identifiable but within the model there exists an 'embedded model' that is identifiable and is the original model that we wished to fit.

We can express our reparameterised model as follows:

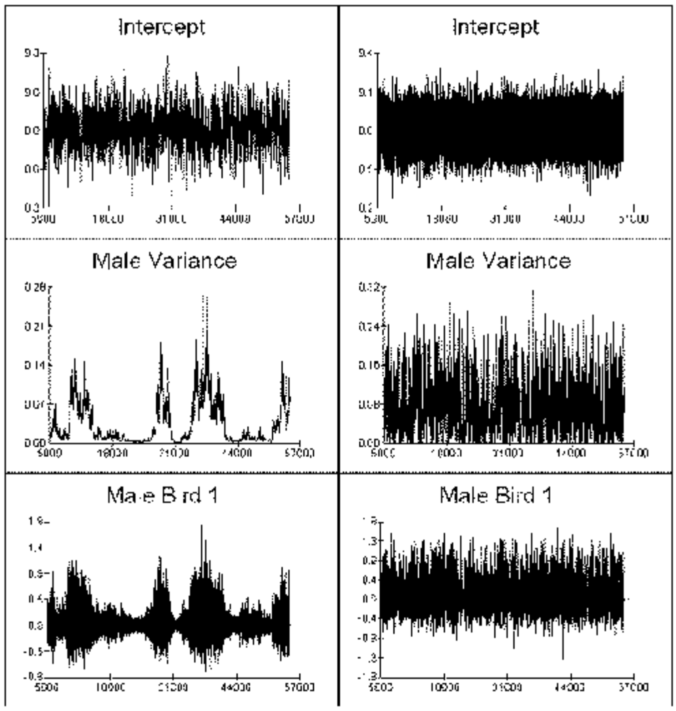

Figure 1: MCMC traces for three parameters for the standard parameterisation (left) and the reparameterisation (right)

$$
\begin{aligned}
y_i &= \alpha_2 v^{(2)}_{female(i)} + \alpha_3 v^{(3)}_{male(i)} + \alpha_4 v^{(4)}_{nestbox(i)} + \beta^{(5)}_{year(i)} + e_i, \\
v^{(2)}_{female(i)} &\sim N(0, \sigma^2_{v(2)}), v^{(3)}_{male(i)} \sim N(0, \sigma^2_{v(3)}), \\
v^{(4)}_{nestbox(i)} &\sim N(0, \sigma^2_{v(4)}), \beta^{(5)}_{year(i)} \sim N(\beta_0, \sigma^2_{v(5)}), e_i \sim N(0, \sigma^2_e), \\
\beta_0 &\propto 1, \alpha_k \propto 1, \sigma^2_{v(k)} \sim \Gamma^{-1}(\epsilon, \epsilon), k = 2, \ldots, 5, \sigma^2_e \sim \Gamma^{-1}(\epsilon, \epsilon).
\end{aligned}
$$

The original parameters can be found by $u^{(k)}_i = \alpha_k v^{(k)}_i$, $\sigma^2_{u(k)} = \alpha^2_k \sigma^2_{v(k)}, k = 2, \ldots, 4$ and $u^{(5)}_i = \beta^{(5)}_i - \beta_0$. It should be noted that although hierarchical centering doesn't change the model at all, parameter expansion has altered the 'diffuse' priors used for the random effect variances.

This reparameterised model can be easily fit in WinBUGS and 55,000 iterations took 2,526 seconds. The model gives similar estimates to the original model in table 1 with slight discrepancies for $\sigma^2_{u(3)}$ due to the change of priors. However the mixing of all parameters is greatly improved as is shown by the improvement in ESS for all parameters and the graphs in the right hand columns of figures 1 and 2.

4 Conclusions

In this article we have shown that the performance of MCMC algorithms can be improved greatly by reparameterisation. In our example we have combined two methods, hierarchical centering and parameter expansion that treat two separate aspects of the poor performance. These methods can be easily implemented in the software package WinBUGS. In future work

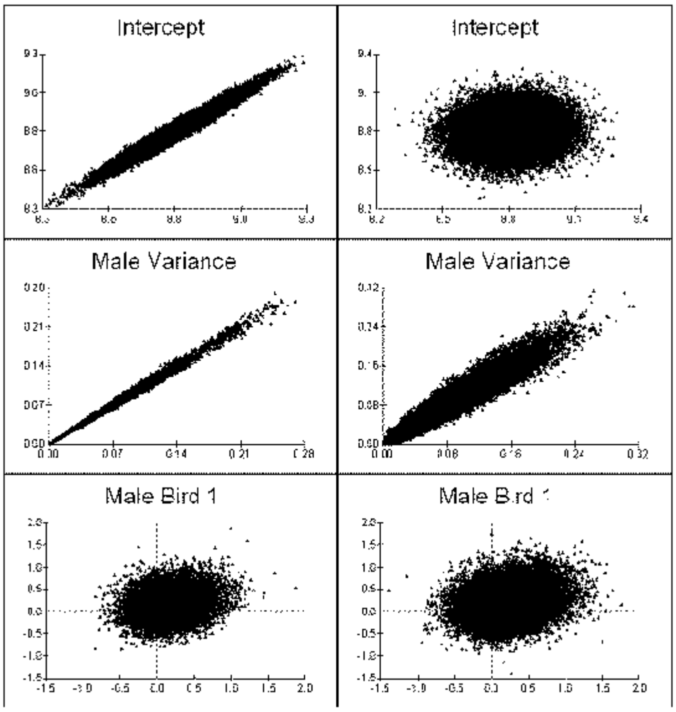

Figure 2: Plots of consecutive iterations for three parameters for the standard parameterisation (left) and the reparameterisation (right)

we are interested in incorporating these methods in MLwiN and also comparing the performance of these methods with block updating methods.

References

Browne, W.J., Goldstein, H. and Rasbash, J. (2001). Multiple membership multiple classification (MMMC) models. *Statistical Modelling* 1: 103-124.

Browne, W.J., Pettifor, R.A., Sheldon, B.C., Rasbash, J. and McCleery, R.H. (2004). Using cross-classified multivariate mixed response models with application to the reproductive success of great tits (Parus Major). *Nottingham Statistics Research report* 03-18.

Gelfand A.E., Sahu S.K., and Carlin B.P. (1995). Efficient Parametrizations For Normal Linear Mixed Models. *Biometrika* 82 (3): 479-488.

Kass, R.E., Carlin, B.P., Gelman, A. and Neal, R. (1998). Markov chain Monte Carlo in practice: a roundtable discussion. *American Statistician*, 52, 93-100.

Liu, C., Rubin, D.B., and Wu, Y.N. (1998) Parameter expansion to accelerate EM: The PX-EM algorithm. *Biometrika* 85 (4): 755-770.