Elena A. Erosheva and Stephen E. Fienberg (Carnegie Mellon University, Pittsburgh, PA, USA)

We congratulate the authors on their stimulating paper. We hope that they can suggest how to adapt their approach to a problem with which we have been struggling.

We are interested in what is known as the Grade of Membership (GoM) model (see Manton et al. (1994) and Erosheva et al. (2002)). For a random sample of subjects, we observe $J$ dichotomous responses, $x_1, \ldots, x_J$. We assume there are $K$ basis sub-populations, which are determined by the conditional (positive) response probabilities, $\lambda_{kj}$, $j = 1, \ldots, J$. The subjects are characterized by their degrees of membership in each of the sub-populations, $g = (g_1, \ldots, g_K)$, which are nonnegative and add to 1. Conditional on the subject’s membership scores, $g$, the subject’s response probability for item $j$ is given by a convex combination $\Pr(x_j = 1|g) = \sum_k g_k \lambda_{kj}$. We assume that the responses $x_1, \ldots, x_J$ are conditionally independent, given the membership scores, and that the membership scores, $g$, have a Dirichlet distribution with parameters $\alpha = (\alpha_1, \ldots, \alpha_K)$.

By using a data augmentation procedure, we obtain a posterior distribution of the parameters via a Metropolis-Hastings within Gibbs algorithm. The current implementation of the algorithm involves separate Metropolis-Hastings steps for the hyperparameters, reparameterized as $\alpha_0 = \sum_k \alpha_k$ and $\xi = \alpha/\alpha_0$, and a Gibbs sampler for the structural parameters, $\lambda = \{\lambda_{kj} : k = 1, \ldots, K; j = 1, \ldots, J\}$, membership scores, $g$, and the variables from data augmentation (Erosheva (2002)).

To date we have fit the model to a $2^{16}$ table, separately for $K = 2, 3, 4, 5$. Incrementing the number of sub-populations from $K$ to $K + 1$ produces 17 additional structural parameters, $\lambda_{K+1,j}$, $j = 1, \ldots, 16$ and $\alpha_{K+1}$, and also increases the number of incidental parameters linearly with the number of subjects. As we have increased $K$, we have observed a slowdown in mixing, especially for the hyperparameter $\alpha_0$, and the MCMC takes longer to achieve convergence.

The GoM model is a generalized mixture model which assumes partial instead of complete membership in a component. This gives us more incidental parameters. Since there is already poor mixing for the hyperparameters of the model, the question is whether the chain with a reversible jump can achieve reasonable mixing properties. Do the authors have any advice on how to adapt their approach in this circumstance?

We are interested in a number of related applications of essentially the same structure but where the number of variable and the number of sub-populations are considerably larger (and where the speed of convergence is also important). Is there really any hope for reversible jump in this context?

REFERENCES
