Mixed Model of Multinomial Discrete Choice: Update 4/26

1 Overview

1. Simulated data for multinomial logit using demographic variables as covariates.
2. Attempted additional alterations to the code including using Train's method for sampling β and the Polya-Gamma latent variables for sampling θ but did not observe any improvements.
3. Regularization/ variable selection with random coefficients - on reviewing the literature on selection of random effects, I found that the approach to applying sparse priors on random effects differs from that of fixed effects. Our present approach (Park and Casella, 2008) is used to address sparsity in fixed effects.

2 Possible Alternative Approaches

2.1 Multinomial logit

Consider the case where the vector of covariates pertains to demographic information about customers. We are interested in estimating the vector of parameters β_j which is unique to each product. The choice probability in such a case is,

$$P(y_{itqj} = j) = \frac{\exp(g\left(X_{itq}, \beta_j\right))}{\sum_{k=1}^{J} \exp(g\left(X_{itq}, \beta_k\right))}.$$

Polson and Scott (2011) describe the full set of conditionals for the parameters under such a specification. The large number of products in the supermarket application ensures that despite the change in specification, the problem continues to be a high-dimensional one. The use of sparsity priors can then be explored in this set-up.

2.2 Random Effects Selection

As detailed in Chen and Dunson (2003), sparsity in random coefficients is accounted for by applying priors that shrink the diagonal elements of the covariance matrix. The procedure involves the Cholesky decomposition of the covariance matrix and the specification of a prior that can generate positive probability values to variances of random effects that are zero. Kinney and Dunson (2007) also demonstrate the Bayesian procedure for selection in logistic mixed models.