Introduction to Bayesian Data Analysis

Tutorial 11

(1) Problem 11.2 (Hoff) Randomized block design: Researchers interested in identifying the optimal planting density for a type of perennial grass performed the following randomized experiment: Ten different plots of land were each divided into eight subplots, and planting densities of 2,4,6 and 8 plants per square metre were randomly assigned to the subplots, so that there are two subplots at each density in each plot. At the end of the growing season the amount of plant matter yield was recorded in metric tons per hectare. The data appear in the file `pdensity.dat.` The researchers want to fit a model like $y = \beta_1 + \beta_2 x + \beta_3 x^2 + \epsilon$, were y is the yield and x is the planting density, but worry that since soil conditions vary across plots they should allow for some across-plot heterogeneity in this relationship. To accommodate this possibility we will analyse these data using the hierarchical linear model.

(a) Before we do a Bayesian analysis we will get some ad hoc estimates of these parameters via least squares regression. Fit the model $y = \beta_1 + \beta_2 x + \beta_3 x^2 + \epsilon$ using OLS for each group, and make a plot of residuals versus fitted values for each of your models. From the least squares coefficients find ad hoc estimates of $\boldsymbol{\theta}$ and Σ. Also obtain an estimate of σ^2 by combining the information from the residuals across the groups.

(b) Now we will perform an analysis of the data using the following distributions as prior distributions:

$$\Sigma^{-1} \sim \text{Wishart}(4, \hat{\Sigma}^{-1})$$
$$\boldsymbol{\theta} \sim \text{MVN}(\hat{\boldsymbol{\theta}}, \hat{\Sigma})$$
$$\sigma^2 \sim \text{InvGamma}(1, \hat{\sigma}^2)$$

where $\hat{\boldsymbol{\theta}}, \hat{\Sigma}$ and $\hat{\sigma}^2$ are the estimates you obtained in a). Note that this analysis is not combining prior information with information from the data, as the "prior" distirbution is based on the observed data. However, such an analysis can be roughly interpreted as the Bayesian analysis of an individual who has weak but unbiased prior information.

(c) Use a Gibbs sampler to approximate posterior expectations of $\boldsymbol{\beta}$ for each group j, and plot the posterior expectations of residuals versus fitted values. Compare to the plot in a) and comment on any differences.

(d) From your posterior samples, plot the marginal posterior and prior densities $\hat{\boldsymbol{\theta}}$ and the elements of Σ. Discuss the evidence that the slopes or intercepts vary across groups.

(e) Suppose we want to identify the planting density that maximises average yield over a random sample of plots. Find the value $x_{\max}$ of x that maximises expected yield of a randomly sampled plot having planting density $x_{\max}$.

(2) Problem 11.3 (Hoff) The researchers in Problem 1 are worries that the plots are not just heterogeneous in their regression lines, but also in their variances. In this exercise we will consider the same hierarchical model as above except that the sampling variability within a group is given by $y_{i,j} \sim \text{normal}(\beta_{1,j} + \beta_{2,j}x_{i,j} + \beta_{3,j}x_{i,j}^2, \sigma_j^2)$, that is, the variances are allowed to differ across groups. We will model $\sigma_1^2, ..., \sigma_m^2 \overset{\text{iid}}{\sim} \text{inverse-gamma}(\nu_0/2, \nu_0\sigma_0^2/2)$ with $\sigma_0^2 \sim \text{gamma}(2, 2)$ and $p(\nu_0)$ uniform on the integers $\{1, 2, 3, ..., 100\}$.

(a) Obtain the full conditional distribution of σ_0^2.

(b) Obtain the full conditional distribution of σ_j^2.

(c) Obtain the full conditional distribution of $\boldsymbol{\beta}_j$.

(d) For two values $\nu_0^{(s)}$ and ν_0^* of ν_0, obtain the ratio $p(\nu_0^*|\sigma_0^2, \sigma_1^2, ..., \sigma_m^2)$ divided by $p(\nu_0^{(s)}|\sigma_0^2, \sigma_1^2, ..., \sigma_m^2)$ and simplify as much as possible.

(e) Implement a Metropolis-Hastings algorithm for obtaining the joint posterior distribution of all the unknown parameters. Plot values of σ_0^2 and ν_0 versus iteration number and describe the mixing of the Markov chain in terms of these parameters.

(f) Compare the prior and posterior distributions of ν_0. Comment on any evidence there is that the variances differ across groups.