A Comprehensive Project on Bayesian Hierarchical Models and MCMC Algorithms

Yuhang Liu

Aug 5, 2015

1 Question 1 (Homework from STA 5107)

Write a program implementing the **Metropolis-Hasting** algorithm to sample a random variable X with the density

$$f(x) = \frac{x^2|sin(\pi * x)|e^{-x^3}}{M}, \quad x > 0$$

where

$$M = \int_0^\infty x^2|sin(\pi * x)|e^{-x^3}dx$$

is the normalizing constant.
You have to decide what q (proposal density) you want to use. Choose positive numbers to start the Markov chain.
(a) Plot the density function $f(x)$.
(b) Histogram the values attained by Markov chain and compare it to the plot of $f(x)$.
(c) Estimate the value of $E[X]$ and $var(X)$ using values of the Markov chain.

2 Question 2 (A Famous Data Set Project and Package *SAT Coaching* in R, Bayesian Hierarchical Model)

Let y_j be the estimated coaching effects and σ_j be the corresponding standard error for school j as summarized in Table 1. We now perform a fully Bayesian analysis of the following hierarchical model

$$y_j \sim N(\theta_j, \sigma_j^2)$$

$$\theta_j \sim N(\mu, \tau^2)$$

where we consider the normal prior for μ: $\mu \sim N(0, 10000)$, and the Inverse Gamma prior for τ^2: $\tau^2 \sim IG(a, b)$, with $a = 0.01$ and $b = 0.01$. Answer the following questions and attach your R code.

School	Estimated treatment effect y_j	Standard error of effect estimate σ_j
A	28.39	14.9
B	7.94	10.2
C	-2.75	16.3
D	6.82	11.0
E	-0.64	9.4
F	0.63	11.4
G	18.01	10.4
H	12.16	17.6

Table 1: Observed effects of special preparation on SAT-V scores in eight randomized experiments. Rubin (1981)

(a) Do the priors for the hyperparameters μ and τ^2 express any strong prior belief? Carefully justify your answer.
(b) Find the full probability model.
(c) Derive the full conditionals for all the (hyper)parameters involved.
(d) Implement a Gibbs sampler in R and use the posterior simulations to estimate (i) for each school j, the probability that its coaching program is the best of the eight; and (ii) for each pair of schools, j and k, the probability that the coaching program in school j is better than that in school k.
(e) Investigate the sensitivity of the posterior probability that School C's coaching program is the best of the eight schools under two alternative prior specifications for τ^2; that is, choose two alternative settings for (a, b).
(f) Now let $\tau^2 = +\infty$. What does this really mean in data analysis? Use the posterior simulations to estimate probabilities (i) and (ii) in (d).
(g) Implement a classical random-effects analysis of variance (ANOVA) of the SAT coaching data, and estimate θ_js for $j = 1, 2, ..., 8$.
(h) Discuss how the answers in (d) and (f) differ. Based on the difference you observe, discuss the advantage of Bayesian inference, as opposed to the classical analysis in (g).