MAST30027: Modern Applied Statistics

Corrected Assignment 4, 2019

Due: 11:59pm Tuesday Oct 29th

This assignment is worth 13% of your total mark. To get full marks, show your working including derivation and the R code you use. You are not allowed to use STAN.

Data: 'Assign4Data.txt' contains 100 observations simulated from a normal distribution with mean = 5 and standard deviation = 2 by using the following code.

```
> set.seed(30027)
> x = rnorm(100, 5, 2)
```

Model: we consider a normal model.

$$x_i \sim \mathrm{N}(\mu, \frac{1}{\tau}) \quad \text{for} \quad i = 1, \ldots, 100.$$

Prior: we impose the following prior for mean and precision parameters.

$$p(\mu, \tau) \propto \frac{1}{\tau}$$

Problem 1: Posterior inference using Gibbs sampling

(a) Give the following conditional distributions including their parameters. For example, gamma distribution with shape = $\sum_i x_i$ and scale = $\sum_i x_i^2$. Show your work.

$$p(\mu|\tau, x_1, \ldots, x_{100}) \quad \text{and} \quad p(\tau|\mu, x_1, \ldots, x_{100}).$$

(b) Write a code that uses the Gibbs sampling to simulate samples from $p(\mu, \tau|x_1, \ldots, x_{100})$. Run at least two Gibbs sampling chains with different initial values. Please run with at least 500 iterations. Make a trace plot for each of parameters and see if samples from different chains are mixed well and behave similarly.

(c) Using the simulated samples, for each parameter 1) make a plot that shows empirical (estimated) marginal posterior distribution, 2) estimate marginal posterior mean, and 3) report a 90% credible interval for the marginal posterior distribution. You can find a 90% credible interval in a number of ways. For this assignment, use 5% in each tail.

Problem 2: Posterior inference using the Metropolis-Hastings (MH) algorithm

(a) Write a code that uses the MH algorithm to simulate samples from $p(\mu, \tau|x_1, \ldots, x_{100})$. For the current values of parameters (μ_c, τ_c), we propose new values (μ_n, τ_n) as follows. $\tau_n \sim$ gamma(shape = $5\tau_c$, rate = 5) and $\mu_n \sim$ Normal(mean = μ_c, variance = τ_n). Run at least two MH chains with different initial values. Please run with at least 10000 iterations. Make a trace plot for each of parameters and see if samples from different chains are mixed well and behave similarly.

(b) Repeat (c) in the Problem 1.