ISYE 6420 – HW #4

Jing Ma

2024-10-12

Problem 1:

In this question, we are given the unnormalized posterior:

$$p(\theta|y) \propto e^{-\frac{1}{2}\{\theta_1^2\theta_2^2+\theta_1^2+\theta_2^2-2\theta_1\theta_2-4\theta_1-4\theta_2\}}$$

1.1

Please see the accompanying Jupyter Notebook for the implementation of Metropolis algorithm based on the illustration in Unit 5.9 Weibull Lifetimes and during 10/2 OH by Aaron Reding. We manually tuned the scale until the acceptance rate reached 0.40. The scale that we chose is 0.95 with the actual acceptance rate of 0.40.

1.2

Please see the accompanying Jupyter Notebook for the plot of the sample points over the two-dimensional image of the distribution. From the plot, we observed that the sample points mainly concentrate in the high-density area.

1.3

Please see the accompanying Jupyter Notebook for the plots of he marginal densities of parameters, θ_1 and θ_2. As expected, the plots highly resemble each other since they are generated from the bivariate normal distribution.

1.4

We used np.percentile() function to compute the 95% equi-tailed credible intervals for both parameters:

$$\theta_1 : [0.014348733795099915, 3.310998701887477]$$

$$\theta_2 : [-0.0069567445246182 91, 3.2376253506865567]$$

Problem 2:

We are given two priors and one likelihood distribution below:

$$y|\theta_1, \theta_2 \sim N(\theta_1 + \theta_2, 1)$$

$$\theta_i \sim^{\text{iid}} N(0, \nu^2)$$

$$\nu^2 \sim \text{ Inv-Gamma}(10, 10)$$

2.1

Based on the given distribution above, we know that:

$$f(y|\theta_1, \theta_2) = \frac{1}{\sqrt{2\pi 1}} \exp\left(-\frac{(y-\theta_1-\theta_2)^2}{2*1}\right)$$

$$\propto \exp\left(-\frac{(y-\theta_1-\theta_2)^2}{2}\right)$$

θ_1 and θ_2 have the same normal distribution:

$$\pi(\theta_1) = \frac{1}{\sqrt{2\pi\nu^2}} \exp\left(-\frac{\theta_1^2}{2\nu^2}\right)$$
$$\propto \nu^{-1} \exp\left(-\frac{\theta_1^2}{2\nu^2}\right)$$

$$\pi(\theta_2) = \frac{1}{\sqrt{2\pi\nu^2}} \exp\left(-\frac{\theta_2^2}{2\nu^2}\right)$$
$$\propto \nu^{-1} \exp\left(-\frac{\theta_2^2}{2\nu^2}\right)$$

Finally, we have ν^2:

$$\pi(\nu^2) = \frac{10^{10}}{\Gamma(10)(\nu^2)^{10+1}} \exp\left(-\frac{10}{\nu^2}\right)$$
$$\propto \nu^{-22} \exp\left(-\frac{10}{\nu^2}\right)$$

Our joint distribution is:

$$f(y, \theta_1, \theta_2, \nu^2) \propto \exp\left(-\frac{(y-\theta_1-\theta_2)^2}{2}\right) \nu^{-1} \exp\left(-\frac{\theta_1^2}{2\nu^2}\right) \nu^{-1} \exp\left(-\frac{\theta_2^2}{2\nu^2}\right) \nu^{-22} \exp\left(-\frac{10}{\nu^2}\right)$$
$$\propto \nu^{-24} \exp\left(-\frac{(y-\theta_1-\theta_2)^2}{2}\right) \exp\left(-\frac{\theta_1^2}{2\nu^2}\right) \exp\left(-\frac{\theta_2^2}{2\nu^2}\right) \exp\left(-\frac{10}{\nu^2}\right)$$

To find the full conditional for θ_1, we select the terms from $f(y, \theta_1, \theta_2, \nu^2)$ that contain θ_1. We have:

$$\pi(\theta_1|\theta_2, y, \nu^2) = \exp\left(-\frac{(y-\theta_1-\theta_2)^2}{2}\right) \exp\left(-\frac{\theta_1^2}{2\nu^2}\right)$$
$$= \exp\left\{-\left[\frac{(y-\theta_1-\theta_2)^2}{2} + \frac{\theta_1^2}{2\nu^2}\right]\right\}$$
$$= \exp\left\{-\frac{[\nu^2(y-\theta_1-\theta_2)^2 + \theta_1^2]}{2\nu^2}\right\}$$

To simply the computation, we will substitute $y - \theta_2$ with h and we will replace h at the end:

$$
\begin{aligned}
\pi(\theta_1|\theta_2, y, \nu^2) &= \exp\left\{-\frac{[\nu^2(h-\theta_1)^2 + \theta_1^2]}{2\nu^2}\right\} \\
&= \exp\left\{-\frac{(\nu^2h^2 - 2\nu^2h\theta_1 + \nu^2\theta_1^2) + \theta_1^2}{2\nu^2}\right\} \\
&= \exp\left\{-\frac{(\nu^2+1)\theta_1^2 - 2\nu^2h\theta_1 + \nu^2h^2}{2\nu^2}\right\} \\
&\propto \exp\left\{-\frac{(\nu^2+1)\theta_1^2 - 2\nu^2h\theta_1}{2\nu^2}\right\} \\
&= \exp\left\{-\frac{(\nu^2+1)\theta_1^2 - 2\frac{\nu^2+1}{\nu^2+1}\nu^2h\theta_1}{2\nu^2}\right\} \\
&= \exp\left\{-\frac{(\nu^2+1)\left(\theta_1^2 - 2\frac{\nu^2h}{\nu^2+1}\theta_1\right)}{2\nu^2}\right\} \\
&\propto \exp\left\{-\frac{(\nu^2+1)\left[\theta_1^2 - 2\frac{\nu^2h}{\nu^2+1}\theta_1 + \left(\frac{\nu^2h}{\nu^2+1}\right)^2\right]}{2\nu^2}\right\} \\
&= \exp\left\{-\frac{(\nu^2+1)\left(\theta_1 - \frac{\nu^2h}{\nu^2+1}\right)^2}{2\nu^2}\right\} \\
&= \exp\left\{-\frac{\left(\theta_1 - \frac{\nu^2h}{\nu^2+1}\right)^2}{2\left(\frac{\nu^2}{\nu^2+1}\right)}\right\} \\
&= \exp\left\{-\frac{\left(\theta_1 - \frac{\nu^2(y-\theta_2)}{\nu^2+1}\right)^2}{2\left(\frac{\nu^2}{\nu^2+1}\right)}\right\}
\end{aligned}
$$

which is the kernel of a Normal:

$$
N\left(\frac{\nu^2(y-\theta_2)}{\nu^2+1}, \frac{\nu^2}{\nu^2+1}\right) \Longrightarrow N\left(\frac{y-\theta_2}{1+\nu^{-2}}, \frac{1}{1+\nu^{-2}}\right)
$$

The full conditional for θ_2 follows the same distribution except that we replace θ_1 with θ_2:

$$
\pi(\theta_2|y, \theta_1, \nu^2) \sim N\left(\frac{y-\theta_1}{1+\nu^{-2}}, \frac{1}{1+\nu^{-2}}\right)
$$

Lastly, the full conditional for ν^2 is:

$$\pi(\nu^2|\theta_1,\theta_2,y) \propto \nu^{-24}\exp\left(-\frac{\theta_1^2}{2\nu^2}\right)\exp\left(-\frac{\theta_2^2}{2\nu^2}\right)\exp\left(-\frac{10}{\nu^2}\right)$$
$$= \nu^-24\exp\left\{-\frac{\theta_1^2+\theta_2^2+20}{2\nu^2}\right\}$$
$$= \left(\nu^2\right)^{-12}\exp\left\{-\frac{\theta_1^2+\theta_2^2+20}{2\nu^2}\right\}$$
$$= \left(\nu^2\right)^{-(11+1)}\exp\left\{-\frac{\theta_1^2+\theta_2^2+20}{2\nu^2}\right\}$$

which is the kernel of an Inverse-Gamma:

$$\text{Inv-Gamma}\left(11, \frac{\theta_1^2+\theta_2^2+20}{2}\right)$$

Please see the accompanying Jupyter Notebook for the implementation of Gibbs sampling based on the illustration in Unit 5.11 Normal-Cauchy Gibbs Sampler by Aaron Reding.

2.2

Please see the accompanying Jupyter Notebook for the plots of the marginal posterior densities of θ_1, θ_2, and ν^2.

By using np.percentile() function, we found the 95% credible intervals below:

$$\theta_1 : [-1.0296756595906118, 1.8407254560401045]$$

$$\theta_2 : [-1.0296756595906118, 1.1108362365338769]$$

$$\nu^2 : [0.5627993547475181, 1.9302897056213753]$$

Further, we found the mean of the three parameters below:

$$\bar{\theta_1} = 0.39513482147703755$$

$$\bar{\theta_2} = 0.39570468137156445$$

$$\bar{\nu^2} = 1.049465244212932$$

2.3

Regarding the trace plots, please see the accompanying Jupyter Notebook. We observed that all three plots show good mixing by the rapid oscillations around the central region throughout the iterations.