ST451 Bayesian Machine Learning
Week 8

Exercises

1. Verify that each step of a Gibbs Sampler can be viewed as a Metropolis Hastings algorithm with acceptance probability of 1. For each step you can consider the proposal that draws θ_i from $\pi(\theta_i|y, \theta_{-1})$ and keeps θ_{-i} to its current value, hence $q(\theta^{(*)}|\theta^{(t)}) = \pi(\theta_i^{(*)}|\theta_{-i}, y)$.

   Note: The notation in this exercise may be complicated. Feel free to skip it and move to the next one, the main purpose is to complement the slides.

   Answer: In each update of each $\theta^{(t)}$ we will propose $\theta_i^{(*)}$ from $\pi(\theta_i|\theta_{-i}^{(t)}, x)$ setting $\theta^{(*)} = (\theta_i^{*}, \theta_{-1}^{(t)})$, whereas $\theta^{(t)} = (\theta_i^{(t)}, \theta_{-1}^{(t)})$.

   Note that then $q(\theta^{(*)}|\theta^{(t)}) = \pi(\theta_i^{(*)}|\theta_{-i}, x)$. We will therefore accept with probability 1:

$$\alpha(\theta^{(t)}, \theta^{(*)}) = \min\left(1,\ \frac{\pi(\theta^{(*)}|x)\pi(\theta_i^{(t)}|\theta_{-i}^{(t)}, x)}{\pi(\theta^{(t)}|x)\pi(\theta_i^{(*)}|\theta_{-i}^{(t)}, x)}\right)$$
$$= \min\left(1,\ \frac{\pi(\theta_i^{(*)}|\theta_{-i}^{(t)}, x)\pi(\theta_{-i}^{(t)}|x)\pi(\theta_i^{(t)}|\theta_{-i}^{(t)}, x)}{\pi(\theta_i^{(t)}|\theta_{-i}^{(t)}, x)\pi(\theta_{-i}^{(t)}|x)\pi(\theta_i^{(*)}|\theta_{-i}^{(t)}, x)}\right) = 1$$

2. Let $y = (y_1, \ldots, y_n)$ be a r.s. from a $\mathrm{N}(\mu, v)$ where v is distributed according to an $\mathrm{IGamma}(\frac{\nu}{2}, \frac{\nu}{2}\sigma^2)$. The parameter ν is assumed to be known. Finalise the model with an improper prior on μ, σ^2, $\pi(\mu, \sigma^2) \propto 1$.

   (a) Write down the posterior up to proportionality.
   Answer: The posterior can be written as

$$\pi(v, \mu, \sigma^2|y) \propto f(y_1, \ldots, y_n|v, \mu) f(v|\sigma^2)$$
$$\propto \prod_{i=1}^{n} \left\{ v^{-1/2} \exp\left(-\frac{(y_i - \mu)^2}{2v}\right)\right\} \frac{\left(\frac{\nu}{2}\sigma^2\right)^{\nu/2}}{\Gamma(\nu/2)} v^{-\nu/2-1} \exp\left(-\frac{\frac{\nu}{2}\sigma^2}{v}\right)$$
$$\propto v^{-\frac{n+\nu}{2}-1} \exp\left(-\frac{\sum_{i=1}^{n}(y_i - \mu)^2}{2v}\right) (\sigma^2)^{\nu/2} \exp\left(-\frac{\frac{\nu}{2}\sigma^2}{v}\right)$$

   (b) Specify the details needed to construct a Gibbs sampler to draw from the posterior of μ, v and σ^2.
   Answer: The full conditional posterior for v is

$$\pi(v|\mu, \sigma^2, y) \propto v^{-\frac{n+\nu}{2}-1} \exp\left(-\frac{\frac{\sum_{i=1}^{n}(y_i-\mu)^2}{2}}{v}\right) \exp\left(-\frac{\frac{\nu}{2}\sigma^2}{v}\right)$$
$$= v^{-\frac{n+\nu}{2}-1} \exp\left(-\frac{\frac{\nu\sigma^2+\sum_{i=1}^{n}(y_i-\mu)^2}{2}}{v}\right)$$
$$\stackrel{\mathcal{D}}{=} \mathrm{IGamma}\left(\frac{n+\nu}{2}, \frac{\nu\sigma^2 + \sum_{i=1}^{n}(y_i - \mu)^2}{2}\right),$$

whereas for μ is

$$\pi(\mu|v,\sigma^2,y) \propto \exp\left(-\sum_{i=1}^{n}\frac{(\mu^2-2\mu x_i+y_i^2)}{2v}\right) \propto \exp\left(-\frac{n\mu^2-2\mu\sum_{i=1}^{n}y_i}{2v}\right)$$
$$= \exp\left(-\frac{\mu^2-2\mu\bar{x}}{2v/n}\right) \stackrel{\mathcal{D}}{=} \mathrm{N}\left(\bar{y},\frac{v}{n}\right),$$

and for σ^2

$$\pi(\sigma^2|v,\mu,y) \propto (\sigma^2)^{\nu/2+1-1}\exp\left(-\frac{\nu}{2v}\sigma^2\right) \stackrel{\mathcal{D}}{=} \text{Gamma}\left(\nu/2+1,\frac{\nu}{2v}\right),$$

A Gibbs Sampler initiates μ, v and σ^2 and then draws from the three conditional posterior distributions in turn at each iteration.

(c) Generate 200 numbers from the model above with $\mu = 0$, $\sigma^2 = 1$, $v = 1$ and $\nu = 20$, and set these numbers as $y = (y_1, \ldots, y_{200})$. Write a Python script to run $10,000$ iterations of the Gibbs sampler derived in the previous part based on the data you generated.

(d) Provide posterior summaries and traceplots for μ, v and σ^2.

(e) Repeat with PyMC3 and compare the results with the previous part.

Answer: For parts (c), (d) and (e), see the jupyter notebook 'exercise2cde.ipynb'