Homework 7 for Bayesian Data Analysis

Fan JIN (2015011506)

May 22, 2018

Importance Resampling

Suppose you want to estimate $\mathrm{E}_f\, h(\theta)$, with f being some posterior distribution $P(\theta|y)$. Further suppose that you choose a proposal distribution $g(\theta)$, and get a sample $(x_1, \cdots, x_m)$ by importance resampling. Please prove that the average of $(h(x_1), \cdots, h(x_m))$ can be used as an estimator of $\mathrm{E}_f\, h(\theta)$.

Proof: Note that the probability density for x_i is $w(x_i)g(x_i)$ for any i, which does not depend on the order in the sampling with replacement. It follows that

$$\mathop{\mathrm{E}}_g (h(x_i)) = \int h(x_i)w(x_i)g(x_i)\,\mathrm{d}\,x_i$$

$$= \int h(x_i)\frac{P(x_i|y)}{g(x_i)}g(x_i)\,\mathrm{d}\,x_i = \int h(x_i)P(x_i|y)\,\mathrm{d}\,x_i = \mathop{\mathrm{E}}_f h(\theta)$$

for any i. Therefore, the average of $(h(x_1), \cdots, h(x_m))$ is an unbiased estimator of $\mathrm{E}_f\, h(\theta)$:

$$\mathop{\mathrm{E}}_g \frac{1}{m}\sum_{i=1}^{m}(h(x_i)) = \frac{1}{m}\sum_{i=1}^{m}\mathop{\mathrm{E}}_g h(x_i) = \frac{1}{m}\sum_{i=1}^{m}\mathop{\mathrm{E}}_f h(\theta) = \mathop{\mathrm{E}}_f h(\theta).$$

Question 10.6d

$$p(\theta|y) = f(\theta) = N(0,3) = \frac{1}{\sqrt{6\pi}}\exp\left(-\frac{\theta^2}{3}\right).$$

$$g(\theta) = t_3 = \frac{2}{\pi\sqrt{3}}\left(1+\frac{\theta^2}{3}\right)^{-2}.$$

It follows that

$$\mathop{\mathrm{E}}_g\left[(\frac{f(\theta)}{g(\theta)})^2\right] = \int (\frac{f(\theta)}{g(\theta)})^2 g(\theta)\,\mathrm{d}\,\theta$$

$$= \int \frac{1}{6\pi}\exp\left(-\frac{2\theta^2}{3}\right)\cdot\frac{\pi\sqrt{3}}{2}\left(1+\frac{\theta^2}{3}\right)^2\mathrm{d}\,\theta$$

$$= \frac{\sqrt{3}}{12}\int \exp\left(-\frac{2\theta^2}{3}\right)\cdot\left(1+\frac{\theta^2}{3}\right)^2\mathrm{d}\,\theta$$

$$= \frac{\sqrt{2}}{12}\int \exp\left(-t^2\right)\cdot\left(1+\frac{t^2}{2}\right)^2\mathrm{d}\,t$$

$$= \frac{\sqrt{2}}{12}\left[\int \exp\left(-t^2\right)\mathrm{d}\,t + \int t^2\cdot\exp\left(-t^2\right)\mathrm{d}\,t + \frac{1}{4}\int t^4\cdot\exp\left(-t^2\right)\mathrm{d}\,t\right]$$

$$= \frac{\sqrt{2}}{12}\left[\sqrt{\pi} + \frac{1}{2}\sqrt{\pi} + \frac{3}{16}\sqrt{\pi}\right] = \frac{9\sqrt{2\pi}}{64} \approx 0.3525.$$

The effective sample size for $n = 10000$ is

$$n_{\text{eff}} = \frac{n}{\mathrm{E}_g\left[(\frac{f(\theta)}{g(\theta)})^2\right]} = 10000/\frac{9\sqrt{2\pi}}{64} \approx 28369.$$

Question 11.1

Lemma: (Detailed Balance condition) If a Markov chain with transition probability $p(\cdot|\cdot)$ that satisfies

$$\pi(\theta_a) \cdot p(\theta_b|\theta_a) = \pi(\theta_b) \cdot p(\theta_a|\theta_b)$$

for some distribution $\pi(\cdot)$, then $\pi(\cdot)$ is the stationary distribution of this Markov chain.

Using the lemma above, we only need to verify that the Detailed Balance condition is satisfied when $\pi(\cdot) = p(\cdot|y)$.

Note that $r(\theta_a, \theta_b) \cdot r(\theta_b, \theta_a) = 1$, for

$$r(\theta_a, \theta_b) = \frac{p(\theta_b|y) \cdot g(\theta_a|\theta_b)}{p(\theta_a|y) \cdot g(\theta_b|\theta_a)}$$

and

$$r(\theta_b, \theta_a) = \frac{p(\theta_a|y) \cdot g(\theta_b|\theta_a)}{p(\theta_b|y) \cdot g(\theta_a|\theta_b)},$$

which means it is safe to assume that $r(\theta_a, \theta_b) \geq 1$ *without loss of generality*. Thus, θ_b is always accepted after generated from the previous value θ_a, with the probability of 1. On the contrary, θ_a is accepted after generated from θ_b with the probability of $r(\theta_b, \theta_a)$. It follows that

$$p(\theta_b|\theta_a) = g(\theta_b|\theta_a) \cdot 1$$

and

$$p(\theta_a|\theta_b) = g(\theta_a|\theta_b) \cdot r(\theta_b, \theta_a).$$

Plug them all in, and we obtain

$$\text{LHS} = p(\theta_a|y) \cdot p(\theta_b|\theta_a) = p(\theta_a|y) \cdot g(\theta_b|\theta_a) \cdot 1$$
$$= p(\theta_a|y) \cdot g(\theta_b|\theta_a),$$

and

$$\text{RHS} = p(\theta_b|y) \cdot p(\theta_a|\theta_b) = p(\theta_b|y) \cdot g(\theta_a|\theta_b) \cdot r(\theta_b, \theta_a)$$
$$= p(\theta_b|y) \cdot g(\theta_a|\theta_b) \cdot \frac{p(\theta_a|y) \cdot g(\theta_b|\theta_a)}{p(\theta_b|y) \cdot g(\theta_a|\theta_b)}$$
$$= p(\theta_a|y) \cdot g(\theta_b|\theta_a),$$

which gives LHS = RHS and proves the Detailed Balance condition. Q.E.D.