Homework 10
Stat 697U, Spring 2019

Instructor: Brian Van Koten

Due Wednesday, May 1

Here is a basic problem on rejection sampling that I forgot to assign last week.

Problem 1 (3 points)**.** *Let $p(x)$ be a probability density on $\mathbb{R}^d$. Assume that it is not possible to compute $p(x)$, but that it is possible to compute a function $\ell(x)$ so that*

$$\ell(x) = cp(x)$$

for some unknown constant c. Let g be a proposal density for use in the rejection sampling method, and suppose that the envelope condition

$$Mg(x) \geq \ell(x)$$

holds with constant M for all x.

What is the distribution of the number of proposals required to generate one sample from p by rejection sampling with proposal density g? What is the expected number of proposals?

Hint*: The distribution depends only on c and M.*

Problem 2 (3 points)**.** *Suppose that you want to generate an i.i.d. sample from the uniform distribution on the unit ball*

$$B_d := \{x \in \mathbb{R}^d; \|x\|_2 = 1\}$$

in d-dimensions. One possibility is to perform rejection sampling with proposals from the uniform distribution on the cube

$$C_d := [-1, 1]^d.$$

Is this a good method when d is large?

Hint: *You will probably want to use the famous volume formula*

$$|B_d| = \frac{\pi^{d/2}}{\Gamma(n/2+1)},$$

where

$$\Gamma(n/2+1) = \pi^{\frac{1}{2}}(n-1/2)(n-3/2)\dots(1/2).$$

The following problem illustrates how the bottleneck inequality may be used to verify that MCMC converges slowly for multimodal problems. Let

$$E = \{-N+1, -N+2\dots, N\},$$

choose $\beta > 0$, and define the target distribution $\pi : E \to [0,1]$ by

$$\pi(x) = Z^{-1}\exp(-\beta\cos(-2\pi x/N)), \text{ where } Z = \sum_{i\in E}\exp(-\beta\cos(-2\pi x/N)).$$

To design a chain sampling π, we find it convenient to equip E with periodic boundary conditions, identifying $-N$ with N and $N+1$ with $-N+1$, for example. Now consider the transition matrix $P \in [0,1]^E$ with entries

$$\begin{aligned}
P_{ii} &:= \frac{1}{2}\left(\frac{\pi(i)}{\pi(i)+\pi(i+1)} + \frac{\pi(i)}{\pi(i)+\pi(i-1)}\right) && \text{for all } i \in E\\
P_{i,i+1} &:= \frac{1}{2}\frac{\pi(i+1)}{\pi(i)+\pi(i+1)} && \text{for all } i \in E\\
P_{i,i-1} &:= \frac{1}{2}\frac{\pi(i-1)}{\pi(i)+\pi(i-1)} && \text{for all } i \in E\\
P_{ij} &:= 0 && \text{if } j \notin \{i-1, i, i+1\}.
\end{aligned}$$

It should seem roughly plausible that P samples π, since moves in the direction of higher values of π are favored.

Problem 3 (1+1+3 points)**.** *Consider the matrix P and distribution π defined above.*

1. *Show that P is in detailed balance with π.*
2. *Draw a picture showing the distribution π for two values of β, one large and one small.*
3. *Prove using the bottleneck inequality that t_{mix} increases exponentially as β tends to infinity.*

Next, I would like you to prove some of the basic results on the total variation distance and its relation to Markov chains.

Problem 4 (2+2+0.5 points)**.** *Recall the fundamental definition*

$$\|\mu - \nu\|_{\text{TV}} = \max_{A\subset E}|\mu(A) - \nu(A)|.$$

1. *Show that*
$$\|\mu - \nu\|_{\text{TV}} = \frac{1}{2}\sum_{i\in E}|\mu_i - \nu_i|.$$
2. *Show that*
$$\|\mu - \nu\|_{\text{TV}} = \sum_{\substack{i\in E\\ \mu_i \geq \nu_i}} \mu_i - \nu_i.$$

3. *Let P be a stochastic matrix on E, and let μ, ν be measures on E. You have already shown, more or less, that $\|\mu P - \nu P\|_{\mathrm{TV}} \leq \|\mu - \nu\|_{\mathrm{TV}}$. Which of your old homework problems implies this inequality and why?*

Finally, here is a little problem related to a question from class.

Problem 5 (2 points)**.** *Let P be a stochastic matrix on E with invariant distribution π. Define the maximal distance at time t by*

$$d(t) := \max_{x \in E} \|\delta_x P^t - \pi\|_{\mathrm{TV}}.$$

Prove that

$$d(t) = \max_{\lambda} \|\lambda P^t - \pi\|_{\mathrm{TV}},$$

where the maximum above is over all probability distributions λ on E.