

### HOMEWORK 3 - V3

STA414/STA2104 WINTER 2019

*University of Toronto*

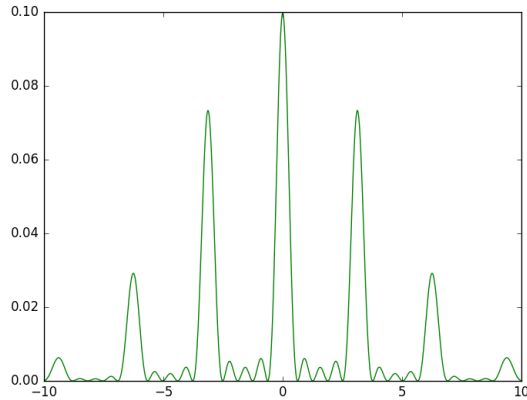
**1. Nature's rejection sampler, 30 points.** A light source emits photons whose angle  $x$  is normally distributed around  $\theta$ :

$$p(x|\theta) = \mathcal{N}(x|\mu = \theta, \sigma^2 = 4^2)$$

Then the photons pass through a grating which absorbs photons with a probability that depends on their angle. We denote the event that a photon makes it through the grating by saying that the Bernoulli variable  $g = 1$ . For this particular grating,  $p(g|x, \theta) = \text{Ber}(g|f(x, \theta))$  where

$$(1.1) \quad f(x, \theta) = \frac{\sin^2(5(x - \theta))}{25 \sin^2(x - \theta)}.$$

The resulting unnormalized probability  $p(x, g = 1|\theta = 0)$ , is shown below as a function of  $x$ .



- (a) **[4 points]** Use numerical integration (i.e. a brute-force sum over many values) to estimate the fraction of photons that get absorbed on average,  $p(g = 0|\theta = 0)$ . That is, estimate  $\int p(x, g = 0|\theta = 0)dx$  by summing over 10,000 values of  $x$ , equally spaced from  $-20$  to  $20$ .
- (b) **[4 points]** Implement a rejection sampler to sample from  $p(x|g = 1, \theta = 0)$ . Hint: Since it's a pmf,  $p(g|x, \theta = 0) \leq 1$ . Another hint: doing ancestral sampling in the above model forms a rejection sampler. **Plot a histogram 10,000 accepted samples, and print the fraction of accepted samples.**
- (c) **[4 points]** Implement a self-normalized importance sampler to estimate the fraction of all photons that get absorbed,  $p(g = 0|\theta = 0)$ . Use  $p(x|\theta = 0)$  as a proposal distribution. As a reminder, the formula for a self-normalized importance sampler for the expectation of a function  $f$  over an (unnormalized) target distribution  $\tilde{p}(z)$  and proposal distribution  $q(z)$

is:

$$(1.2) \quad \hat{e}(x_1, x_2, \dots, x_K) = \frac{1}{K} \sum_{i=1}^K f(x_i) \frac{\frac{\tilde{p}(x_i)}{q(x_i)}}{\frac{1}{K} \sum_{j=1}^K \frac{\tilde{p}(x_j)}{q(x_j)}} \quad \text{where } x_i \sim_{\text{iid}} q(x)$$

If you can derive a simpler expression by canceling terms, you don't have to implement the entire formula. Print the estimate of the fraction of photons that are absorbed, using 1000 samples.

- (d) **[4 points]** A physicist is trying to estimate the location  $\theta$  of the center of a light source. She has a Cauchy prior  $p(\theta) = \frac{1}{10\pi(1+(\frac{\theta}{10})^2)}$ . She observes a photon emitted from the particle at position  $x = 1.7$ . Plot the unnormalized density  $p(x = 1.7, g = 1, \theta)$ , as a function of  $\theta$ .
- (e) **[10 points]** Write a Metropolis-Hastings sampler to sample from  $p(\theta|x = 1.7, g = 1)$ . For the proposal distribution, use a Gaussian centered around the current sample. Fiddle with the proposal distribution and the number of iterations of your sampler until the histogram of samples looks roughly like the true unnormalized density. Plot the histogram of samples from your chain.
- (f) **[4 points]** The scientist wants to know how confident she should be, based on her observation of the photon, that her instrument is calibrated to within 3 degrees around zero. Use samples from your MH chain to estimate the posterior probability  $p(-3 < \theta < 3|x = 1.7, g = 1)$ .

**2. Gradient estimators, 20 points.** All else being equal, it's useful for a gradient estimator to be unbiased. The unbiasedness of a gradient estimator guarantees that, if we decay the step size and run stochastic gradient descent for long enough (see Robbins & Monroe), it will converge to a local optimum.

The standard REINFORCE, or score-function estimator is defined as:

$$(2.1) \quad \hat{g}_{\text{SF}}[f] = f(b) \frac{\partial}{\partial \theta} \log p(b|\theta), \quad b \sim p(b|\theta)$$

- (a) **[5 points]** First, let's warm up with the score function. Prove that the score function has zero expectation, i.e.  $\mathbb{E}_{p(x|\theta)}[\nabla_{\theta} \log p(x|\theta)] = 0$ . Assume that you can swap the derivative and integral operators.
- (b) **[5 points]** Show that REINFORCE is unbiased:  $\mathbb{E}_{p(b|\theta)}[f(b) \frac{\partial}{\partial \theta} \log p(b|\theta)] = \frac{\partial}{\partial \theta} \mathbb{E}_{p(b|\theta)}[f(b)]$ .
- (c) **[5 points]** Show that REINFORCE with a fixed baseline is still unbiased, i.e. show that  $\mathbb{E}_{p(b|\theta)}[(f(b) - c) \frac{\partial}{\partial \theta} \log p(b|\theta)] = \frac{\partial}{\partial \theta} \mathbb{E}_{p(b|\theta)}[f(b)]$  for any fixed  $c$ .
- (d) **[5 points]** If the baseline depends on  $b$ , then REINFORCE will in general give biased gradient estimates. Give an example where  $\mathbb{E}_{p(b|\theta)}[(f(b) - c(b)) \frac{\partial}{\partial \theta} \log p(b|\theta)] \neq \frac{\partial}{\partial \theta} \mathbb{E}_{p(b|\theta)}[f(b)]$  for some function  $c(b)$ , and show that it is biased.

The takeaway is that you can use a baseline to reduce the variance of REINFORCE, but not one that depends on the current action.

**3. Comparing variances of gradient estimators, 25 points.** If we restrict ourselves to consider only unbiased gradient estimators, then the main (and perhaps only) property we need to worry about is the variance of our estimators. In general, optimizing with a lower-variance unbiased estimator will converge faster than a high-variance unbiased estimator. However, which estimator has the lowest variance can depend on the function being optimized. In this question,

we'll look at which gradient estimators scale to large numbers of parameters, by computing their variance as a function of dimension.

For simplicity, we'll consider a toy problem. The goal will be to estimate the gradients of the expectation of a sum of  $D$  independent one-dimensional Gaussians. Each Gaussian has unit variance, and its mean is given by an element of the  $D$ -dimensional parameter vector  $\theta$ :

$$(3.1) \quad f(\mathbf{x}) = \sum_{d=1}^D x_d$$

$$(3.2) \quad L(\theta) = \mathbb{E}_{\mathbf{x} \sim p(\mathbf{x}|\theta)}[f(\mathbf{x})] = \mathbb{E}_{\mathbf{x} \sim \text{iid } \mathcal{N}(\theta, I)} \left[ \sum_{d=1}^D x_d \right]$$

- (a) **[4 points]** As a warm-up, compute the variance of a single-sample simple Monte Carlo estimator of the objective  $L(\theta)$ :

$$(3.3) \quad \hat{L}_{MC} = \sum_{d=1}^D x_d, \quad \text{where each } x_d \sim \text{iid } \mathcal{N}(\theta_d, 1)$$

That is, compute  $\mathbb{V}[\hat{L}_{MC}]$  as a function of  $D$ .

- (b) **[5 points]** The score-function, or REINFORCE estimator with a baseline has the form:

$$(3.4) \quad \hat{g}^{\text{SF}}[f] = [f(x) - c(\theta)] \frac{\partial}{\partial \theta} \log p(x|\theta), \quad x \sim p(x|\theta)$$

Derive a closed form for this gradient estimator for the objective defined above as a deterministic function of  $\epsilon$ , a  $D$ -dimensional vector of standard normals. Set the baseline to  $c(\theta) = \sum_{d=1}^D \theta_d$ . Hint: When simplifying  $\frac{\partial}{\partial \theta} \log p(x|\theta)$ , you shouldn't take the derivative through  $x$ , even if it depends on  $\theta$ . To help keep track of what is being differentiated, you can use the notation  $\frac{\partial}{\partial \bar{\theta}} g(\bar{\theta}, \theta)$  to denote taking the derivative only w.r.t. the second  $\theta$ .

- (c) **[8 points]** Derive the variance of the above gradient estimator. Because gradients are  $D$ -dimensional vectors, their covariance is a  $D \times D$  matrix. To make things easier, we'll consider only the variance of the gradient with respect to the first element of the parameter vector,  $\theta_1$ . That is, derive the scalar value  $\mathbb{V}[\hat{g}_1^{\text{SF}}]$  as a function of  $D$ . Hint: The third moment of a standard normal is 0, and the fourth moment is 3. As a sanity check, consider the case where  $D = 1$ .
- (d) **[8 points]** Next, let's look at the gold standard of gradient estimators, the reparameterization gradient estimator, where we reparameterize  $x = T(\theta, \epsilon)$ :

$$(3.5) \quad \hat{g}^{\text{REPARAM}}[f] = \frac{\partial f}{\partial x} \frac{\partial x}{\partial \theta}, \quad \epsilon \sim p(\epsilon)$$

In this case, we can use the reparameterization  $x = \theta + \epsilon$ , with  $\epsilon \sim \mathcal{N}(0, I)$ . Derive this gradient estimator for  $\nabla_{\theta} L(\theta)$ , and give  $\mathbb{V}[\hat{g}_1^{\text{REPARAM}}]$  as a function of  $D$ .

**4. Complete course quiz, 5 points.** We're planning to update the prerequisites for this course next year. To help us plan, go to this course's page on quercus, and fill out the short quiz on which courses you took before this one, and how well they prepared you for the different parts of this course.