Homework 5

Deadline: December 7, 13:15

Exercise 1

Consider the Johnson's S_u distribution with PDF

$$p(x) = \frac{\sqrt{2}}{\sqrt{\pi(1+(x-1)^2)}} e^{-\frac{1}{2}(3+2sinh^{-1}(x-1))^2}$$

The corresponding Python code for this PDF can be implemented as following using Numpy:

$np.sqrt(2)/np.sqrt(np.pi * (1 + (x - 1) * *2)) * np.exp(-.5 * (3 + 2 * np.arcsinh(x - 1)) * *2)$

1. Implement an importance sampler with a standard normal distribution ($mean = 0$, $variance = 1$) as proposal to generate (weighted) samples from $p(x)$. Draw $L = 1000$ samples.

2. Plot a histogram of the samples together with the target PDF and the proposal PDF and inspect your obtained Monte Carlo approximation. What can you say about the quality of the approximation you have obtained? How does it relate to the shape of the proposal?

3. Try using more samples L. Does the Monte Carlo approximation improve?

4. Limit yourself to $L = 1000$ samples again, and try to instead adjust the proposal to improve the approximation.

5. Use the samples to estimate the mean and variance of the target. (One can analytically show that the mean is -1.41 and the variance 1.98). Explore how the quality of these estimates changes with different L.

6. Change the proposal to a uniform distribution on the interval $[-4, 1]$. Is it still a valid importance sampler?

Exercise 2

Consider the following Markov chain whose stationary distribution is a standard normal distribution

$$x[k+1] = 0.9x[k] + v[k], \; v[k] \propto \mathcal{N}(0, 0.19) \tag{1}$$

1. Verify that the equation 1 is a Markov chain?

2. Implement 1. Start in $x[1] = 5$ and simulate it for $k = 2, 3, \cdots, 300$. Since this is a Markov chain that is simulated, it is a Markov chain Monte Carlo (MCMC) method. Plot a histogram of x and compare it to the PDF of a standard normal distribution.

3. Plot your obtained $x[1], x[2], \cdots, x[300]$. How long appears your burn-in period to be, for this example?

4. Prove that the stationary distribution of 1 is $\mathcal{N}(0, 1)$.

Exercise 3

Consider the Kullback-Leibler divergence

$$KL(p||q) = -\int p(x) ln \frac{q(x)}{p(x)} dx$$

evaluate $KL(p(x)||q(x))$ where $p(x)$ and $q(x)$ are

1. two scalar Gaussians

$$p(x) = \mathcal{N}(x; \mu, \sigma^2) \; q(x) = \mathcal{N}(x; m, s^2)$$

2. two multivariate Gaussians

$$p(\boldsymbol{x}) = \mathcal{N}(\boldsymbol{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) \; q(\boldsymbol{x}) = \mathcal{N}(\boldsymbol{x}; \boldsymbol{m}, \boldsymbol{S})$$

Hint: Use the following identity; assume $\boldsymbol{x} \sim \mathcal{N}(\boldsymbol{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$, then it holds that

$$\mathbb{E}\Big[(\boldsymbol{x} - \boldsymbol{a})^T \boldsymbol{B} (\boldsymbol{x} - \boldsymbol{a})\Big] = (\boldsymbol{\mu} - \boldsymbol{a})^T \boldsymbol{B} (\boldsymbol{\mu} - \boldsymbol{a}) + \mathrm{TR}(\boldsymbol{B\Sigma})$$

Exercise 4

Consider the dynamical model

$$x_{n+1} = \gamma x_n + v_n, \qquad v_n \sim \mathcal{N}(0, \beta_v^2)$$

$$y_n = \frac{1}{2} x_n + e_n, \qquad e_n \sim \mathcal{N}(0, \sigma_e^2)$$

with the initial state and prior

$$x_0 \sim \mathcal{N}(\bar{\boldsymbol{x}}_{\boldsymbol{0}}, \boldsymbol{\Sigma_0})$$

$$\gamma \sim \mathcal{N}(0, \sigma_\gamma^2)$$

We observe $y_1, \cdots, y_N$ and consider $x_1, \cdots, x_N$ and as our latent variables.

1. Derive an expression of the joint distribution

$$p(y_{1:N}, x_{1:N}, \gamma)$$

2. Consider the variational approximation

$$q(\gamma, x_1, \cdots, x_N) = q(\gamma) \prod_{n=0}^{N} q(x_n)$$

   Use Variational inference to derive the equations for estimating the posterior $q(\gamma) \approx p(\gamma|y_1, \cdots y_N)$.

3. We can also solve this problem without factorizing the terms x_n, i.e by considering

$$q(\gamma, x_1, \cdots, x_N) = q(\gamma) q(x_{0:n})$$

   Use variational inference to estimate the posterior of $q(\gamma) \approx p(\gamma|y_1, \cdots y_N)$ using this variational approximation.