Data assimilation and inverse problems – Homework Set 5

Importance sampling

1. Let the target distribution be $p(x) = \mathcal{N}(0, 4)$. Compute $E\left[f(x)\right]$, where

$$f(x) = \begin{cases} 0 & \text{if } x < 4, \\ 1 & \text{if } x \geq 4, \end{cases}$$

   using the proposal distribution $q(x) = \mathcal{N}(0, 1)$. You estimator is

$$\hat{E} = \frac{1}{N_e} \sum_{j=1}^{N_e} f(x_j) w(x_j),$$

   Define an error of your estimator by

$$e = \frac{|E[f(x)] - \hat{E}|}{|E[f(x)]|}.$$

   How many samples do you need to get an average error less than 20%? The average error is defined by averaging the error over 1000 "experiments" of computing the error.

2. Repeat 1, but use the proposal distribution $q(x) = \mathcal{N}(2, 1)$.

3. Let the target distribution be $p = \mathcal{N}(0, I)$, where I is the identity matrix of size n. Let the proposal distribution be $q = \mathcal{N}(0, (1 + \varepsilon)I)$, where $\varepsilon = 0.1$. Compute $\rho = E[w^2]/E[w]^2$ as a function of dimension, n, as follows. For a fixed n perform 100 experiments to compute ρ using $N_e = 10^5$ samples. Use the average of ρ over the 100 experiments as your estimate of ρ. Vary n between 10 and 10^3 and plot ρ as a function of n. What do you observe?

4. Let the target distribution be $p(x) = 0.7p_1(x) + 0.3p_2(x)$, where $p_1(x) = \mathcal{N}(0, 1)$ and $p_2(x) = \mathcal{N}(4, 1)$. Construct a proposal distribution to draw weighted samples from $p(x)$. How large is $\rho = E[w^2]/E[w]^2$ for your choice? Draw a histogram of your resampled samples. This requires that you write code for resampling. You can find ideas for a good algorithm in Arulampalam et al., *A Tutorial on Particle Filters for Online Nonlinear/Non-Gaussian Bayesian Tracking*, IEEE Transaction on Signal Processing, Vol. 50, No. 2, 2002.