E0 270 Machine Learning *Due:* 11.59 pm, April 17, 2017; Total: 100 Mark

Assignment 3

Submission Instructions

1. Please submit the assignment 3 on submission page linked from the course webpage as a single pdf file.
2. No hard copy submission is permitted. You can either submit a scanned copy of your handwritten solutions or solution typset using latex.
3. Code for question 5 b should be included in the report(as text). Dont submit separately.
4. Use the pdf submission link in the assignment submission page to submit your pdf.
5. Upon multiple submissions till the deadline, the most recent submission will overwrite the previous one.
6. Questions about the assignment can be posted on the group email if or directed to the TAs {annervaz.km, aadirupa.saha} @csa.iisc.ernet.in, sharmistha@grad.cds.iisc.ac.in.

Questions

1. Compute the bias and variance of the importance sampling estimator.

2. Let U be a uniformly distributed random variable in $[0, 1]$. Let F_X be the distribution function of a random variable X with density p. Assume that F_X be invertible. Show that $F_X^{-1}(U)$ has density p.

3. Let q be p be two probability densities. Consider the following sampling scheme:

$$X \sim q$$
$$U \sim Uniform[0, 1]$$

If $U < p(X)/q(X)$, then X is accepted, else X is rejected. What is the marginal distribution of drawing a sample using this scheme.

4. Let p be the distribution from which we have to generate samples. Let $p(x) = \frac{\tilde{p}(x)}{Z}$, where $\tilde{p}$ is the unnormalized probability distribution. We perform rejection sampling using $\tilde{p}$ rather than p as follows:

   (a) Find a distribution q and a constant K, such that $Kq(x) \geq \tilde{p}(x)$ for all x.
   (b) Sample $U \sim Uniform[0, Kq(x)]$.
   (c) Sample $X \sim q(x)$
   (d) If $U < \tilde{p}(X)$, accept the point, else reject it.

   Find the marginal distribution of X.

5. A restricted Boltzmann machine defines a joint distribution over $\{0, 1\}^d \times \{0, 1\}^m$. Two boolean random vectors X and H are said to be distributed as a restricted Boltzmann machine if for any x in $\{0, 1\}^d$ and h in $\{0, 1\}^m$, the joint probability distribution of X and H has the form

$$p(X = x, H = h) = \frac{\exp\left(\sum_{i=1}^{d} \sum_{j=1}^{m} x_i h_j w_{ij}\right)}{Z}, \tag{1}$$

   where Z is the normalizing constant.

   (a) Derive the conditional distributions: $p(X_i = x_i | X_{-i} = x_{-i}, H = h)$ and $p(H_j = h_j | X = x, H_{-j} = h_{-j})$.
   (b) Implement a Gibbs sampler for a restricted Boltzmann machine, with $d = 10$ and $m = 20$ using either Matlab or Python.

6. Explain the similarities and dissimilarities between Long Short Term Memory(LSTM) and Gated Recurrent Unit(GRU) architectures.