T.61.5140 Machine Learning: Advanced Probablistic Methods
Hollmén, Raiko (Spring 2008)
Problem session, 4th and 11th of April, 2008
http://www.cis.hut.fi/Opinnot/T-61.5140/

These are programming excercises that are discussed on 4th of April and demonstrated on 11th of April. The Matlab files are on the course webpage.

Let us use a notation $N(x \mid \mu, \sigma^2)$ to denote a Gaussian distribution of x with mean μ and variance σ^2.

1. Study the unknown distribution $p_u(x)$ by drawing samples from it. You can get values of $p_u(x)$ by calling the Matlab function unknown_ p.m with the argument x which may be a scalar or a vector. Use rejection sampling (page 528) with a proposal distribution $q(x) = N(0, 2)$ (Gaussian with mean 0 and variance 2) and scaling k=1.5 (You get values of $kq(x)$ as 1.5*gaussian(x,0,2)). (a) Plot a figure that shows the accepted samples with a black dot ('.k') and rejected samples with a cyan dot ('.c'). Use both 100 and 1000 samples. (b) Estimate the expected values $E(x)$ and $E(\tanh(x))$ over distribution $p_u(x)$ by using the accepted samples.

2. Study the same distribution $p_u(x)$ by using importance sampling (page 532). (a) Plot the importance weights $p_u(x)/q(x)$. (b) Estimate $E(x)$ and $E(\tanh(x))$ over distribution $p_u(x)$ by weighting all samples. (c) Was it more accurate than rejection sampling?

3. Let us study a model with two variables: $x_1 \sim N(x_1 \mid 0, 1.0)$ and $x_2 \sim N(x_2 \mid x_1, 0.1)$ where the two Gaussians are independent. (a) Solve $p(x_1 \mid x_2)$. (b) Initialize $x_1^{(0)} = x_2^{(0)} = 0$ and plot 50 samples from the model by Gibbs sampling (page 542) with a line connecting consequtive samples.

4. Let us study a model with three variables, x_1, x_2, and x_3. The model is such that x_1 and x_2 are drawn from the same distribution $p_u(x)$ as before, and $x_3 \sim N(x_3 \mid x_1 + x_2, 1.0)$. Draw samples from $p(x_1, x_2 \mid x_3 = 1.0)$ by using Metropolis algorithm (page 538). Use a proposal distribution $x_1^* \sim N(x_1^* \mid x_1^{(\tau)}, 1.0)$ and $x_2^* \sim N(x_2^* \mid x_2^{(\tau)}, 1.0)$ (simultaneously, not

alternately as in Gibbs sampling). Plot 50 samples with a line connecting consequtive samples ′-′ and rejected proposals with a cross ′x′.