Problems for week 7. Total value is 6.

1. PRML 11.1

2. The text for 11.1.1 is arguably awkward, but the notion of how to sample from a probability distribution an important. Fill in the following details.

   a) The blue curve in figure 11.2 has a special name. What is it?

   b) Referring to figure 11.2, can you describe intuitively how to sample from the red distribution? (Hint: Suppose you have a dart that can land on the Y axis, uniformly between the 0 and the 1). Explain intuitively why more of your samples end up near the peaks of the red curve.

   c) Show that 11.5 is true (no need to be rigorous).

2. PRML 11.2

3. Using the material in 11.2.1, AND, what is shown in problem 5 below, write an argument that the Metropolis-Hastings has the desired properties (start by saying what they are) under reasonable conditions (list them as relevant). Note that what you are doing in this question is explaining, in your own style, why MH works. As part of this, you should fill in a few details of 11.45, to verify for yourself that you understand it.

5. (Double value).

Even if you don't do this problem, you should understand what it is about.

(Based on Neal, 93).

An important feature of MCMC sampling is that if you run it for long enough, the samples come from the desired distribution, regardless of the starting state. This is ergodicity property mentioned in the book. The intuition is that each transition preserves the stationary part, but tends to kill the rest. This is often analyzed using the eigenvectors of the transition matrix, but here we will consider doing it algebraically:

Consider some value (not specified), *v*, that is in (0,1). Let the stationary distribution be $\pi(x)$. The initial value will be drawn from a distribution $r_o(x)$.

Consider writing the distribution at time *n* as:

$$p_n(x) = \left(1 - (1-v)^n\right)\pi(x) + (1-v)^n r_n(x)$$

Where $r_n(x)$ is a probability distribution.

Observe that for *n*=0, the above is trivially true (explain why).

Using the formula 11.39, and the definition of stationary, show that the above holds for all *n* (use induction).

What is an expression for $r_n(x)$? Argue that for suitable *v*, $r_n(x)$ will always a probability distribution. (Hint: consider that all transition probabilities are strictly positive)

Write an expression for $|\pi(x) - p_n(x)|$ that can be used to argue that $p_n(x) \Rightarrow \pi(x)$