EE3731C Tutorial - Statistical Signal 3.5

Department of Electrical and Computer Engineering

1. Let's further motivate why we might want to generate samples from a probability distribution. Let $\vec{x}$ be a random vector consisting of N random variables: $\{x_1, \cdots, x_N\}$. Suppose each random variable x_n can take on values from 1 to K. Assume $N = 20$ and $K = 10$.

   (a) Given an observation y, to find the MAP estimate of $\vec{x}$ given y, we want:

   $$\vec{x}_{MAP} = \underset{\vec{x}}{\operatorname{argmax}}\, p(\vec{x}|y) \tag{1}$$

   To evaluate Eq. (1), we can iterate through every possible value of $\vec{x}$ and keep track of the value of $\vec{x}$ with the highest posterior probability. Suppose evaluating $p(\vec{x}|y)$ for a particular pair of $\vec{x}$ and y takes 1ms. How long would this exhaustive search approach take to compute the MAP estimate?

   (b) Given an observation y, suppose we want to find the MMSE estimate of $\vec{x}$ given y:

   $$\vec{x}^{MMSE} = E_{p(\vec{x}|y)}(\vec{x})$$

   We can compute the above MMSE estimate for each x_n independently. In other words

   $$x_n^{MMSE} = E_{p(x_n|y)}(x_n) = \sum_{x_n=1}^{K} x_n p(x_n|y)$$

   Suppose evaluating $p(\vec{x}|y)$ for a particular pair of $\vec{x}$ and y takes 1ms. How long would it take to compute x_n^{MMSE}? How long would it take to compute $\vec{x}^{MMSE}$?

2. Suppose we want to sample $x \in \{1, 2, 3, 4\}$,

   $$\begin{aligned} p(x = 1) &= 0.4 \\ p(x = 2) &= 0.1 \\ p(x = 3) &= 0.2 \\ p(x = 4) &= 0.3 \end{aligned}$$

   with the Metropolis algorithm. Suppose the proposal distribution is

   $$\begin{aligned} q(x' = x - 1|x) = q(x' = x|x) = q(x' = x + 1|x) = 1/3 &\qquad \text{for } x \in \{2, 3\} \\ q(x' = 1|x) = 2/3; q(x' = 2|x) = 1/3 &\qquad \text{for } x = 1 \\ q(x' = 3|x) = 1/3; q(x' = 4|x) = 2/3 &\qquad \text{for } x = 4 \end{aligned}$$

(a) Suppose current x is 2 and x' is 3, what is the probability of accepting x' as the new x?

(b) Suppose current x is 3 and x' is 2, what is the probability of accepting x' as the new x?

(c) What is the equivalent Markov chain transition matrix in this Metropolis algorithm?

(d) Show that this particular Metropolis algorithm will converge to $p(x)$ by showing that the equivalent markov chain satisfy the conditions of the Fundamental Theorem of Markov Chain.