EE 641 Final Exam

Fall 2014

Name: **Key** ________________________________

Instructions

- This exam contains 4 problems worth a total of 105 points.
- You may not use any notes, textbooks, or calculators.
- Answer questions precisely and completely. Credit will be subtracted for vague answers.

Good luck.

Problem 1. (25pt)

Let x and y be vectors in R^N.

a) Calculate a closed form expression for the ICD update when the MAP cost function has the form.

$$f(x;y) = ||y - x||^2 + ||x||_1$$

where $||x||_1$ is the L_1 norm of x.

b) Prove that $f(x;y)$ is a strictly convex function of x.

c) Prove that this function takes on a unique global minimum.

Solution:

a) Let $f(x_i)$ denote the value of $f(x;y)$ as a function of x_i. Then we have that

$$f(x_i) = (x_i - y_i)^2 + |x_i| \ .$$

If $x_i > 0$, then we have that

$$\frac{\partial f(x_i)}{\partial x_i} = 2(x_i - y_i) + 1 = 0 \ .$$

Solving for x_i results in $x_i = y_i - \frac{1}{2}$. If $x_i < 0$, then we have that

$$\frac{\partial f(x_i)}{\partial x_i} = 2(x_i - y_i) - 1 = 0 \ .$$

Solving for x_i results in $x_i = y_i + \frac{1}{2}$. If $x_i = 0$, then we have that $|x_i - y_i| \leq \frac{1}{2}$.

Putting this together results in

$$x_i \leftarrow \begin{cases} y_i - \frac{1}{2} & \text{if } y_i > 1/2 \\ 0 & \text{if } |y_i| \leq 1/2 \\ y_i + \frac{1}{2} & \text{if } y_i < 1/2 \end{cases}$$

This can also be written as

$$x_i \leftarrow \text{sign}(y_i) \left[|y_i| - \frac{1}{2} \right]^+$$

where the notation $[x]^+ = \max\{0, x\}$.

b) The function $||y - x||^2$ is a strikely convex function of x and the function $||x||_1$ is a convex function of x. So therefore, the sum of the two functions is a strickly convex function of x.

c) First, the function $f(x;y)$ is strickly convex on $x \in R^N$, so if it has a local minimum, then this must be the unique global minimum.

Now in order to prove that there is at least one local minimum, define the compact set

$$B = \{x \in R^N : ||x|| \leq 2||y|| + 1\} \ .$$

Then since B is a convex and compact set and $f(x; y)$ is a strictly convex function, we know that $f(x; y)$ must take on a unique global minimum on B. So more formally, we know that for all $x \in B$, there exists a $x^* \in B$ such at $f(x^*; y) < f(x; y)$.
However, we also know that for all $||x|| > 2||y|| + 1$,

$$f(x; y) = ||x - y||^2 + ||x||_1 > ||y||^2 = f(0; y) \leq f(x^*; y) \ .$$

So therefore, we know that for all $x \in R^N$, $f(x^*; y) < f(x; y)$. So in fact, x^* is the unique global minimum in R^N.

Problem 2. (30pt) Consider the homogeneous Markov chain $\{X_n\}_{n=0}^{\infty}$ with parameters

$$\begin{aligned} \tau_j &= P\{X_0 = j\} \\ P_{i,j} &= P\{X_n = j, X_{n-1} = i\} \end{aligned}$$

where $i, j \in \{0, \cdots, M-1\}$. Furthermore, assume that the transition parameters are given by

$$P_{i,j} = \begin{cases} 1/2 & \text{if } j = (i+1) \text{mod} M \\ 1/2 & \text{if } j = i \\ 0 & \text{otherwise} \end{cases}$$

a) Write out the transition matrix P for the special case of $M = 4$. (But solve the remaining problems for any M.)

b) Is the Markov chain irreducible? Prove or give a counter example.

c) Is the Markov chain periodic? Prove or give a counter example.

d) Is the Markov chain ergodic? Prove or give a counter example.

e) Determine the value of the following matrix

$$\lim_{n \to \infty} P^n$$

f) Is the Markov chain reversible? Prove or give a counter example.

Solution:

a)

$$P = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \end{bmatrix}$$

b) Yes, it is irreducible because it is homogeneous because for $i \neq j$ we have that

$$P\left\{X_{(i-j)\text{mod}M} = j | X_0 = i\right\} = \left(\frac{1}{2}\right)^{(i-j)\text{mod}M} > 0 \ .$$

c) No, it is not period because for all states i,

$$P\{X_1 = i | X_0 = i\} = 1/2$$

d) Yes, because the Markov chain has a finite number of states, it is irreducible, and it is aperiod; therefore it must be ergodic.

e) Since the Markov chain is ergodic, then the limit

$$P^{\infty} = \lim_{n \to \infty} P^n$$

must exist, and every row of the matrix P^∞ is equal to the ergodic distribution of the MC. The ergodic distribution is the unique solution to the full balance equations given by

$$\pi P = \pi$$

where $\sum_{i=0}^{M-1} \pi_i = 1$. In particular, the distribution

$$\pi_i = \frac{1}{M} \ ,$$

solves the full balance equations; so therefore

$$[P^\infty]_{i,j} = \frac{1}{M} \ .$$

f) No, the MC is not reversible because it does not satisfy the detailed balance equations. In particular, for $i = 0, \cdots, M-2$, the detailed balance equations are given by

$$\pi_i P_{i,i+1} = \pi_{i+1} P_{i+1,i} \ ,$$

but substituting in with $i = 0$, we have that

$$\frac{1}{M}\frac{1}{2} \neq \frac{1}{M} 0 \ .$$

Problem 3. (25pt) Let X_n be a sequence of i.i.d. random variables for $n = 1, \cdots, N$, and let Y_n be a sequence of random variables that are conditionally independent and exponentially distributed with mean μ_i for $i = 0, \cdots, M-1$ given the corresponding values of X_n. More specifically, the distributions of X and Y are given by

$$\begin{aligned} P\{X_n = i\} &= \pi_i \\ p(y_n|x_n) &= \frac{1}{\mu_{x_n}} e^{-y_n/\mu_{x_n}} u(y_n) \end{aligned}$$

where $i \in \{0, \cdots, M-1\}$ and $u(\cdot)$ is the unit step function. Furthermore, let $\theta = [\pi_0, \mu_0, \cdots, \pi_{M-1}, \mu_{M-1}]$ be the parameter vector of the distribution.

a) Derive a closed form expression for the ML estimate of θ given both X and Y (i.e., the complete data).

b) Find the natural sufficient statistics for the exponential distribution of (X, Y).

c) Derive the EM algorithm for computing the ML estimate of θ from Y. (Feel free to use the shortcuts we learned in class.)

Solution:

a) Define

$$\begin{aligned} N_i &= \sum_{n=1}^{N} \delta(X_n = i) \\ S_i &= \sum_{n=1}^{N} Y_n \delta(X_n = i) \ . \end{aligned}$$

Then the ML estimate is given by

$$\begin{aligned} \pi_i &= \frac{N_i}{N} \\ \mu_i &= \frac{S_i}{N_i} \ . \end{aligned}$$

b) The natural sufficient statistics are given by $\{N_i, S_i\}_{i=0}^{M-1}$.

c) The posterior distribution is given by

$$f(x = i|y, \theta) = \frac{\frac{\pi_i e^{-y/\mu_i}}{\mu_i}}{\sum_{i=0}^{M} \frac{\pi_i e^{-y/\mu_i}}{\mu_i}}$$

So then the E-step is given by

$$\begin{aligned} \bar{N}_i &= \sum_{n=1}^{N} f(i|y_n, \theta^{(k)}) \\ \bar{S}_i &= \sum_{n=1}^{N} Y_n f(i|y_n, \theta^{(k)}) \ , \end{aligned}$$

and the M-step is given by

$$
\begin{aligned}
\pi_i &= \frac{\bar{N}_i}{N} \\
\mu_i &= \frac{\bar{Y}_i}{N} \, .
\end{aligned}
$$

Problem 4. (25pt) Consider the positive random variable X with density

$$p(x) = \frac{1}{\sigma z(p)} \exp\left\{ -\frac{1}{p\sigma^p}|x|^p \right\} ,$$

with $p = 1.2$ and $\sigma = 1$. Consider the case of a Hastings-Metropolis simulation algorithm for sampling from the distribution of $p(x)$ with the proposals generated as $W \sim q(w)$ where

$$q(w) = \frac{1}{2\alpha} \exp\left\{ -\frac{1}{\alpha}|w| \right\} ,$$

where $\alpha = 1$.

a) Sketch the density function for $p(x)$ and the proposal density $q(w)$. How do they compare?

b) Derive an expression for the acceptance probability α.

c) Explain how the inverse CDF method can be used to generate random variables from the distribution of X? What are there advantages and disadvantages of this approach?

d) Propose a method which is both accurate and computationally efficient for generating samples from the distribution of X.

Solution:

a) They are similar but not the same. In general, $q(w)$ has heavier tails than $p(x)$.

b) In the general case, the acceptance probability is given by

$$\begin{aligned}
\alpha &= \min\left\{ 1, \frac{q(x^k|w)}{q(w|x^k)} \exp\left\{ -[u(w) - u(x^k)] \right\} \right\} \\
&= \min\left\{ 1, \frac{q(x^k)}{q(w)} \exp\left\{ -[u(w) - u(x^k)] \right\} \right\} \\
&= \min\left\{ 1, \frac{\exp\{-\frac{1}{\alpha}|x^k|\}}{\exp\{-\frac{1}{\alpha}|w|\}} \exp\left\{ -\frac{1}{p\sigma^p}(|w|^p - |x^k|^p) \right\} \right\} \\
&= \min\left\{ 1, \exp\left\{ -\frac{1}{\alpha}(|x^k| - |w|) - \frac{1}{p\sigma^p}(|w|^p - |x^k|^p) \right\} \right\}
\end{aligned}$$

c) Let $F(x)$ be the CDF of X. Then, we can generate a random variable with the desired distribution by first generating a uniformly distributed random variable U on the interval $[0, 1]$, and then computing $X = F^{-1}(U)$. The advantage of this method is that it is not iterative, but the disadvantage is that you must be able to compute a closed form expression for $F^{-1}(U)$. In practice, it will be necessary to compute a table that approximates the function $F^{-1}(u)$. This approximation might be quite significant for some applications.

d) Compute an approximation to the inverse CDF denoted by $\tilde{F}^{-1}(u)$. This approximation can be computed using a look up table for exmple. Then use the proposals generated by $W \sim \tilde{F}^{-1}(U)$ where U is a uniform random variable. Then use the Hastings-Metropolis algorithm to generate random variables X from the exact distribution.