EE 641 Final Exam

Fall 2016

Name: Solution

Instructions

- This exam contains 4 problems worth a total of 100 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)
Consider the problem

$$\hat{x} = \arg \min_{x \in \mathbb{R}^{+N}} f(x) ,$$

where $f : \mathbb{R}^N \to \mathbb{R}$ is a convex function. In order to remove the constraint, we may define the proper, closed, convex function

$$g(x) = \begin{cases} 0 & \text{if } x \in \mathbb{R}^{+N} \\ \infty & \text{if } x \notin \mathbb{R}^{+N} \end{cases} .$$

Then the minimum is given by the solution to the unconstrained optimization problem

$$\hat{x} = \arg \min_{x \in \mathbb{R}^N} \{f(x) + g(x)\} . \tag{1}$$

Using this formulation, do the following.
a) Use variable splitting to derive a constrained optimization problem that is equivalent to equation (1).
b) Formulate the augmented Lagrangian for this constrained optimization problem and give the iterative algorithm for solving the augmented Lagrangian problem.
c) Use the ADMM approach to formulate an iterative algorithm for solving the augmented Lagrangian.
d) Simplify the expressions for the ADMM updates and give the general simplified ADMM algorithm for implementing positivity constraints in convex optimization problems.

Solution

(a) $$\hat{x} = \arg\min_{\substack{x, z \in \mathbb{R}^N \\ x = z}} \{ f(x) + g(z) \}$$

(b) $$L(x, z; a, u) = f(x) + g(z) + \frac{a}{2} \| x - z + u \|^2$$

Initialize a, u
Repeat {

$$\hat{x}, \hat{z} \leftarrow \arg\min_{\substack{x \in \mathbb{R}^N \\ z \in \mathbb{R}^N}} \{ L(x, z; a, u) \}$$

$$u \leftarrow u + \hat{x} - \hat{z}$$

}

Name: ______________________________

(c) Initialize $a, u, \hat{z}$

Repeat {

$$\hat{x} \leftarrow \underset{x \in \mathbb{R}^N}{\operatorname{argmin}} \left\{ f(x) + \frac{a}{2} \| x - (\hat{z} - u) \|^2 \right\}$$

$$\hat{z} \leftarrow \underset{z \in \mathbb{R}^N}{\operatorname{argmin}} \left\{ \frac{1}{2} \| z - (\hat{x} + u) \|^2 + \frac{1}{a} g(z) \right\}$$

$$u \leftarrow u + \hat{x} - \hat{z}$$

}

Initialize $a, u, \hat{z}$

(d) Repeat {

$$\hat{x} \leftarrow \underset{x \in \mathbb{R}^N}{\operatorname{argmin}} \left\{ f(x) + \frac{a}{2} \| x - (\hat{z} - u) \|^2 \right\}.$$

$$\hat{z} \leftarrow \max \{ \hat{x} + u , 0 \}$$

$$u \leftarrow u + \hat{x} - \hat{z}$$

}

Problem 2. (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) Use the results of from class to derive the EM algorithm for computing the ML estimate of θ from Y.

Solution

$$p(y, x \mid \pi, \mu) = \prod_{n=1}^{N} \prod_{m=0}^{M-1} \left(\frac{\pi_m e^{-y_n/\mu_m}}{\mu_m} \right)^{\delta(x_n - m)} u(y_n).$$

$$\log p(y, x \mid \pi, \mu) = \sum_{m=0}^{M-1} \log\left(\frac{\pi_m}{\mu_m}\right) \cdot N_m - \frac{S_m}{\mu_m} + \log u(y_n)$$

where $N_m = \sum_{n=1}^{N} \delta(x_n - m)$

$$S_m = \sum_{n=1}^{N} y_n \delta(x_n - m).$$

We find ML estimate by forming a Lagrangian function

$$L(\pi, \mu, \lambda) = \sum_{m=0}^{M-1} \log\left(\frac{\pi_m}{\mu_m}\right) \cdot N_m - \frac{S_m}{\mu_m} + \lambda\left(\sum_{m=0}^{M-1} \pi_m - 1\right)$$

$$\frac{\partial}{\partial \pi_m} L(\pi, \mu, \lambda) = \frac{N_m}{\pi_m} + \lambda \quad \text{———} \quad (1)$$

Name: ______________________

$$\frac{\partial}{\partial \lambda} L(\pi, \mu, \lambda) = \sum_{m=0}^{M-1} \pi_m - 1 \quad \text{———} \quad (2)$$

$$\frac{\partial}{\partial \mu_m} L(\pi, \mu, \lambda) = -\frac{N_m}{\mu_m} + \frac{S_m}{\mu_m^2} \quad \text{———} \quad (3).$$

setting (1), (2) and (3) equal to 0 and solving them simultaneously gives

$$\boxed{\pi_m = \frac{N_m}{N}}$$

$$\boxed{\mu_m = \frac{S_m}{N_m}}$$

(b) The natural sufficient statistics are given by

$$\{N_m, S_m\}_{m=0}^{M-1}$$

(c) First we find posterior distribution using bayes rule

$$f(x=m \mid y, \pi, \mu) = \frac{\pi_m e^{-y/\mu_m} / \mu_m}{\sum_{i=0}^{M-1} \pi_i \, e^{-y/\mu_i} / \mu_i}$$

Name: ____________________

So then the E-step is given as

E-step

$$\overline{N_m} = \sum_{n=1}^{N} f(m \mid y_n, \pi^{(k)}, \mu^{(k)})$$

$$\overline{S_m} = \sum_{n=1}^{N} y_n \, f(m \mid y_n, \pi^{(k)}, \mu^{(k)})$$

And M-step is given as

M-step

$$\pi_m = \frac{\overline{N_m}}{N}$$

$$\mu_m = \frac{\overline{S_m}}{\overline{N_m}}$$

Problem 3. (25pt)

Let X_n be a Markov chain known as as a birth-death process that takes on values in the set $\{0, \cdots, M-1\}$. With each new time increment, the value of X_n is either incremented (i.e., birth occurs), decremented (i.e., death occurs), or the state remains unchanged. So for $i = 0, \cdots, M-2$, then $P_{i,i+1} = \lambda$ is the birth rate; and for $i = 1, \cdots, M-1$, then $P_{i,i-1} = \mu$ is the death rate where $\lambda + \mu < 1$.

a) Calculate all the values of $P_{i,j}$ for all $i, j \in \{0, \cdots, M-1\}$.

b) Prove or disprove that the Markov chain is stationary, irreducible, and aperiodic.

c) Prove or disprove that the Markov chain is reversible.

d) Calculate the density π_i that is the stationary distribution of the Markov chain.

Solution

This is a finite state birth-death process. because the Markov chain, X_n, takes values in the set $\{0, 1, \cdots, M-1\}$. With each new time increament, the value of X_n is either increamented (i.e, birth occurs), decreamented (i.e., death occurs), or the state does not change.

Mathematically, this can be expressed as

$$P_{i,j} = \begin{cases} \lambda & \text{if } j = i+1 \text{ and } i < M-1 \\ \mu & \text{if } j = i-1 \text{ and } i > 0 \\ 1-\lambda-\mu & \text{if } j = i \text{ and } 0 < i < M-1 \\ 1-\lambda & \text{if } j = i \text{ and } i = 0 \\ 1-\mu & \text{if } j = i \text{ and } i = M-1 \\ 0 & \text{otherwise} \end{cases}$$

where λ is probability of birth and μ is probability of death.

Name: ____________________

b) This markov chain has finite number of states. And all of the states communicate $[P^M]_{i,j} > 0$, so the Markov chain is <u>irreducible</u>.

The Markov chain is <u>aperiodic</u> because $[P^k]_{i,j} > 0$ for $k = M, \cdots, 2M-1$

Since this is a finite state Markov chain which is irreducible and aperiodic, therefore there exists a unique stationary distribution which is the solution to the full balance equations.

(c) In order to see if the Markov chain is reversible or not, we shall if its stationary distribution satisfies the detailed balance equations

$$\pi_i \lambda = \pi_{i+1} \mu \qquad \text{for } i \in \{0, 1, \ldots, M-2\}.$$

$$\Rightarrow \quad \pi_i = \pi_0 \left(\frac{\lambda}{\mu}\right)^i$$

Name: ______________________

where π_0 can be found as

$$\pi_0 = \frac{1}{\sum_{i=0}^{M-1} \left(\frac{\lambda}{\mu}\right)^i} = \frac{1 - \lambda/\mu}{1 - \left(\frac{\lambda}{\mu}\right)^M} \quad \text{for } \frac{\lambda}{\mu} \neq 1$$

and $\pi_0 = \frac{1}{M}$ if $\frac{\lambda}{\mu} = 1$.

Since we have a solution to the detailed balance equations, we know that this Markov chain is <u>reversible</u>.

(d). The solution to the full balance equations is the same as the solution to the detailed balance equations, therefore the stationary distribution is given as

$$\pi_i = \begin{cases} \dfrac{1 - \lambda/\mu}{1 - \left(\frac{\lambda}{\mu}\right)^M} \left(\dfrac{\lambda}{\mu}\right)^i & \text{when } \dfrac{\lambda}{\mu} \neq 1. \\[2ex] \dfrac{1}{M} & \text{when } \dfrac{\lambda}{\mu} = 1. \end{cases}$$

Problem 4. (25pt) Consider a stochastic simulation designed to generate samples from the Gibbs distribution given by

$$p(x) = \frac{1}{z}\exp\{-u(x)\}$$

using the Hastings-Metropolis algorithm. Also assume that the following proposal distribution is used.

$$q(w|x^k) = \frac{1}{z}\exp\{-u(w)\}$$

Then show that the acceptance probability is 1, and the algorithm converges to the Gibbs distribution in one step.

Solution

$$q(w|x^{(k)}) = \frac{1}{z}\exp\{-u(w)\}.$$

and $$q(x^{(k)}|w) = \frac{1}{z}\exp\{-u(x^{(k)})\}.$$

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

$$= \min\left\{1, \frac{\exp\{-u(x^{(k)})\}}{\exp\{-u(w)\}}\frac{\exp\{-u(w)\}}{\exp\{-u(x^{(k)})\}}\right\}$$

$$= \min\{1, 1\}$$

$$= 1.$$

Name: ______________________

The algorithm converges to the Gibbs distribution in one step because the very first sample has the distribution $\frac{1}{Z}\exp\{-U(x^{(0)})\}$ since that is the proposal density function. And because of this proposal density, acceptance probability is equal to 1.