EE 641 Final Exam

Fall 2017

Name: Key

Instructions

- This exam contains 4 problems worth a total of 100 points.
- You may have up to 120 minutes to take the exam.
- 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 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.

a) $P = \begin{bmatrix} 1/2 & 1/2 & 0 & 0 \\ 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 1/2 & 1/2 \\ 1/2 & 0 & 0 & 1/2 \end{bmatrix}.$

b) Irreducible.

Note: $P_{i,(i+1) \bmod M} = 1/2 > 0$ for all i

$\Rightarrow$ $i \leftrightarrow i+1$ (communicate) for all i ②

Also use the equivalence property of communication:

If $i \leftrightarrow j$ and $j \leftrightarrow k$ then $i \leftrightarrow k$ for all i, j, k ①

Use math induction on ① and ② to show.

$i \leftrightarrow j \quad \forall i, j$

Name: ______________________

c). Not periodic.

For all i, we have

$$P\{X_{n+1}=i \mid X_n=i\} = 1/2 > 0$$

d) It is ~~er~~godic.

Satisfies:

- Finite set of states (M)
- Irreducible.
- Aperiodic.

e). Since X is an ergodic MC,

then $P^{\infty} = \lim_{n\to\infty} P^n$ ~~exists~~. and we have unique π for the full balance equation

$$\pi \cdot P = \pi. \quad \text{s.t. } \sum_i \pi_i = 1. \quad ①$$

Note: $\pi_i = 1/M \; \forall i$ satisfies ①.

$\Rightarrow (P^{\infty})_{ij} = 1/M \quad \forall i, j$

f) Not reversible.

Does not satisfy detailed balance equation.

$$\pi_i P_{ij} = \pi_j P_{ji} \quad \forall i, j$$

count example: when $j = i+1$ $\frac{1}{M} \cdot \frac{1}{2} \neq \frac{1}{M} \cdot 0$

Problem 2. (25pt) Let $\rho(x)$ for $x \in \mathbb{R}$ be the Huber function for $T = 1$ given by

$$\rho(x) = \begin{cases} \frac{x^2}{2} & \text{for } |x| < 1 \\ |x| - \frac{1}{2} & \text{for } |x| \geq 1 \end{cases}.$$

The objective of this problem is to determine a surrogate function, $Q(x; x')$, for $\rho(x)$.
a) Use the symmetric bound method to determine $Q(x; x')$.
b) From $Q(x; x')$, determine the function $q(x; x')$, so that $q(x'; x') = \rho(x')$.
c) Plot the functions $\rho(x)$ and $q(x; x')$ for $x' = 4$ and indicate the location of the minimum

$$x'' = \arg\min_{x \in \mathbb{R}} q(x; 4)$$

d) What is the advantage of the symmetric bound methods?

a) Associated influence function for $\rho(x)$ is given by:

$$\rho'(x) = \text{clip}\{x, [-1, 1]\}$$

where $\text{clip}\{x, [a,b]\}$ clips value of x to interval $[a, b]$.

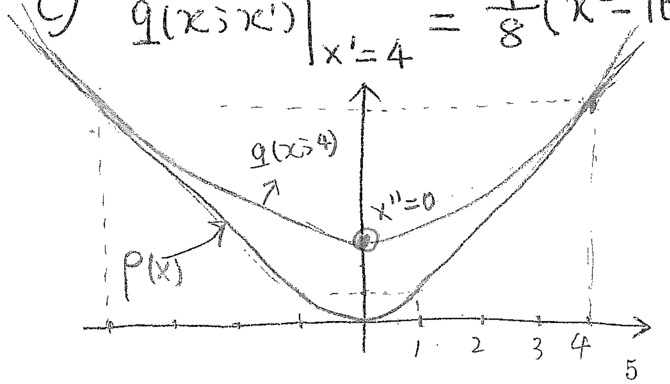
For the symmetric bound the surrogate function is given by

$$Q(x; x') = \frac{\text{clip}\{x', [-1,1]\}}{2x'} x^2$$

b) $$q(x; x') = Q(x; x') - Q(x'; x') + \rho(x')$$

$$= \frac{\text{clip}\{x', [-1,1]\}}{2x'}(x^2 - x'^2) + \rho(x').$$

c) $$q(x; x')\Big|_{x'=4} = \frac{1}{8}(x^2 - 16) + \frac{7}{2} = \frac{1}{8}x^2 + \frac{3}{2}$$



Name: ______________________________

d) Symmetric bound methods usually result in a much less Conservative upper bound so it converges much faster.

Problem 3. (25pt) Let $\{X_n\}_{n=1}^N$ be a i.i.d. random varibles with distribution

$$P\{X_n = m\} = \pi_m \ ,$$

where $\sum_{m=0}^{M-1} \pi_m = 1$. Also, let Y_n be conditionally independent random variables given X_n, with Poisson conditional distribution

$$p(y_n|x_n = m) = \frac{\lambda_m^{y_n} e^{-\lambda_m}}{y_n!} \ .$$

a) Write out the density function for the vector Y.
b) What are the natural sufficient statistics for the complete data (X, Y)?
c) Give an expression for the ML estimate of the parameter

$$\theta = (\pi_0, \lambda_0, \cdots, \pi_{M-1}, \lambda_{M-1}) \ ,$$

given the complete data (X, Y).
d) Give the EM update equations for computing the ML estimate of the parameter $\theta = (\pi_0, \lambda_0, \cdots, \pi_{M-1}, \lambda_{M-1})$ given the incomplete data Y.

a) $p(y_n) = \sum_{m=0}^{M-1} P(x_n=m)\, p(y_n|x_n=m) = \sum_{m=0}^{M-1} \pi_m\, p(y_n|x_n=m)$.

$p(Y) = \prod_{n=1}^{N} p(y_n) = \prod_{n=1}^{N} \left\{ \sum_{m=0}^{M-1} \pi_m \frac{\lambda_m^{y_n} e^{-\lambda_m}}{y_n!} \right\}$.

b) $p(X,Y) = \prod_{n=1}^{N} \prod_{m=0}^{M-1} \left(\frac{\lambda_m^{y_n} e^{-\lambda_m}}{y_n!} \pi_m \right)^{\delta(X_n-m)}$

$= \prod_{n=1}^{N} \prod_{m=0}^{M-1} \exp\left\{ \log \left(\frac{\lambda_m^{y_n} e^{-\lambda_m}}{y_n!} \pi_m \right)^{\delta(X_n-m)} \right\}$.

$= \exp\left\{ \sum_{n=1}^{N} \sum_{m=0}^{M-1} \log \left(\frac{\lambda_m^{y_n} e^{-\lambda_m}}{y_n!} \pi_m \right)^{\delta(X_n-m)} \right\}$.

$= \exp\left\{ \sum_{m=0}^{M-1} \sum_{n=1}^{M} \left[\delta(X_n-m) \left(y_n \log\lambda_m - \log(y_n!) - \lambda_m + \log\pi_m \right) \right] \right\}$

$\therefore$ Sufficient statistic:

$$\begin{cases} N_m = \sum_{n=1}^{N} \delta(X_n - m) \\ b_m = \sum_{n=1}^{N} y_n\, \delta(X_n - m) \end{cases}$$

Name: ______________________________

c) Since parameter $\pi_0 \cdots \pi_{M-1}$ must add to 1, we use Lagrange multiplier to compute ML Estimate: First define

$$L(\theta, \lambda) = \log P(X, Y \mid \theta) + \lambda \left(\sum_{m=0}^{M-1} \pi_m - 1 \right).$$

$$= \sum_{m=0}^{M-1} \sum_{n=1}^{N} \delta(x_n - m) \left(y_n \log \lambda_m - \log (y_n!) - \lambda_m + \log \pi_m \right) + \lambda \left(\sum_{m=0}^{M-1} \pi_m - 1 \right)$$

Maximize $L(\theta, \lambda)$ w.r.t. $[\theta, \lambda]$

$$\frac{\partial L(\theta, \lambda)}{\partial \pi_m} = \frac{\sum_{n=1}^{N} \delta(x_n - m)}{\pi_m} + \lambda = 0 \quad ①$$

$$\frac{\partial L(\theta, \lambda)}{\partial \lambda_m} = \frac{\sum_{n=1}^{N} y_n \delta(x_n - m)}{\lambda_m} - \sum_{n=1}^{N} \delta(x_n - m) = 0 \quad ②$$

$$\frac{\partial L(\theta, \lambda)}{\partial \lambda} = \sum_{m=0}^{M-1} \pi_m - 1 = 0 \quad ③$$

Combine ①②③ and solve them, we get

$$\begin{cases} \hat{\pi}_m = \dfrac{\sum_{n=1}^{N} \delta(x_n - m)}{N} \\ \hat{\lambda}_m = \dfrac{\sum_{n=1}^{N} y_n \delta(x_n - m)}{\sum_{n=1}^{N} \delta(x_n - m)} \end{cases}$$

Check the Hessian is P.d

so $\hat{\pi}_m$, $\hat{\lambda}_m$ is the ML Estimate.

Name: ______________________

d). First we find posterior distribution using bayes rule.

$$f(x_n=m \mid Y=y_n, \theta) = \frac{\pi_m \cdot \frac{\lambda_m^{y_n} e^{-\lambda_m}}{y_n!}}{\sum_{j=0}^{M-1} \frac{\lambda_j^{y_n} e^{-\lambda_j}}{y_n!} \pi_j}$$

then E step:

$$\bar{N}_m = E\left[\sum_{n=1}^{N} \delta(x_n - m) \mid Y=y, \theta^{(k)}\right]$$

$$= \sum_{n=1}^{N} f(x_n=m \mid Y=y_n, \pi_m^{(k)}, \lambda_m^{(k)})$$

$$= \sum_{n=1}^{N} \frac{\pi_m^{(k)} \cdot \frac{\lambda_m^{(k) y_n} \cdot e^{-\lambda_m^{(k)}}}{y_n!}}{\sum_{j=1}^{M-1} \frac{\lambda_j^{(k) \cdot y_n} e^{-\lambda_j^{(k)}}}{y_n!} \cdot \pi_j^{(k)}}$$

$$\bar{b}_m = E\left[\sum_{n=1}^{M} y_n \delta(x_n - m) \mid Y=y, \theta^{(k)}\right]$$

$$= \sum_{n=1}^{N} y_n \cdot f(x_n=m \mid Y=y_n, \theta^{(k)})$$

$$= \sum_{n=1}^{N} y_n \cdot \frac{\pi_m^{(k)} \frac{\lambda_m^{(k) y_n} e^{-\lambda_m^{(k)}}}{y_n!}}{\sum_{j=1}^{M-1} \frac{\lambda_j^{(k) y_n} e^{-\lambda_j^{(k)}}}{y_n!} \pi_j^{(k)}}.$$

M- step:

$$\pi_m^{(k+1)} = \frac{\bar{N}_m}{N}$$

$$\lambda_m^{(k+1)} = \frac{\bar{b}_m}{\bar{N}_m}$$

Problem 4. (25pt) Let X_s be a 2D random field with an Ising distribution given by

$$p(x) = \frac{1}{z(\beta)} \exp\left\{ -\beta \sum_{\{r,s\}\in C} \delta(x_r \neq x_s) \right\} ,$$

and let the function $v(x_s, x_{\partial s})$ be defined by

$$v(x_s, x_{\partial s}) \triangleq \sum_{r\in\partial s} \delta(x_s \neq x_r) .$$

Then do the following:

a) Derive an expression for the conditional distribution of a pixel given its neighbors.

b) Sketch the conditional probability $p(x_s = 1|x_{i\neq s})$ as a function of $v(1, x_{\partial s})$ with $\beta = 1$.

c) What would a "typical" sample of the X look like when $\beta = -2$?

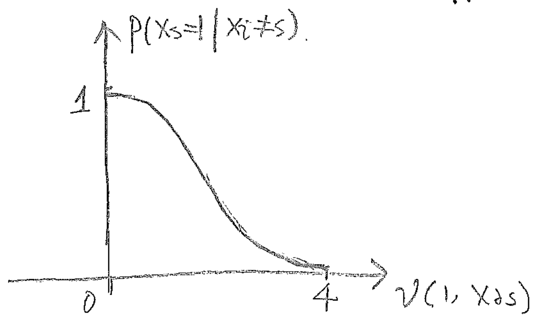
a) $p(x)$ is Gibbs distribution, so X_s is Markov Random Field.

$$P(X_s|X_{\partial s}) = P(X_s|X_{i\neq s}) = \frac{\exp\{-\beta \sum_{r\in\partial s} \delta(X_s \neq X_r)\}}{\exp\{-\beta\sum_{r\in\partial s}\delta(1\neq X_r)\} + \exp\{-\beta\sum_{r\in\partial s}\delta(-1\neq X_r)\}}$$

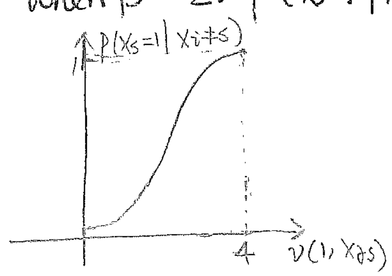
$$= \frac{\exp\{-\beta v(X_s, X_{\partial s})\}}{\exp\{-\beta v(1, X_{\partial s})\} + \exp\{-\beta v(-1, X_{\partial s})\}}$$

b) $$P(X_s=1|X_{i\neq s}) = \frac{1}{1+e^{2\beta[v(1,X_{\partial s})-2]}}$$

for $\beta=1$, $$P(X_s=1|X_{i\neq s}) = \frac{1}{1+e^{2[v(1,X_{\partial s})-2]}}$$



c) when $\beta=-2$, $$P(X_s=1|X_{i\neq s}) = \frac{1}{1+e^{-4[v(1,X_s)-2]}}$$



From the plot, we can see the pixel will mostly be the opposite with the neighbors, so typical sample of X would be like a check board pattern.