CSCI-567: Assignment #5

Due on Monday, November 16, 2015

Saket Choudhary
2170058637

Contents

Problem 1

Problem 1: (a)

To find $\nabla_{y_t} L$:

$$\nabla_{y_t} L = \frac{\partial}{\partial y_t} \frac{1}{2} \sum_{i=1}^{N} ()$$

Problem 3

Given:

$$p(x_i) = \begin{cases} \pi + (1-\pi)e^{-\lambda} & x_i = 0 \\ (1-\pi)\frac{\lambda^{x_i}e^{-\lambda}}{x_i!} & x_i > 0 \end{cases}$$

Alternatively:

$$X_i = \begin{cases} 0 & \text{probability} = \pi + (1-\pi)e^{-\lambda} \\ x_i & \text{probability} = (1-\pi)\frac{\lambda^{x_i}e^{-\lambda}}{x_i!} \end{cases}$$

We define a *latent* variable Z_i for all cases where $X_i = 0$. It is latent because when we observed $X_i = 0$ we do not know if it came out of the 'Poisson' distribution or it came out the 'degenerate' distribution(which has a probability of 1 at point 0.). we cannot observe the following. So X_i comes out of a mixture of a degenerate distribution as follows:

$$Z_i = \begin{cases} 1 & X_i \text{ is from the degenerate distrbution} \\ 0 & \text{otherwise} \end{cases}$$

$$p(X_i = 0, Z_i = 1) = p(Z_i = 1) \times p(X_i = 0|Z_i = 1) = \pi \times 1$$
$$P(X_i = 0, Z_i = 0) = p(Z_i = 0) \times p(X_i = 0|Z_i = 0) = (1-\pi)e^{-\lambda} \times 1$$

$$L(Complete) = \prod_{x_i=0} \pi^{Z_i} \times ((1-\pi)e^{-\lambda})^{1-Z_i} \times \prod_{x_i>0} (1-\pi)e^{\frac{\lambda_i^{x}e^{-\lambda}}{x_i!}} \quad (1)$$

$$\log L = \sum_{I(x_i=0)} z_i \log(\pi) + (1-z_i)\big(\log(1-\pi) - \lambda\big) + \sum_{I(x_i>0)} \big(\log(1-\pi) + x_i\log(\lambda_i) - \lambda - \log(x_i!)\big) \quad (2)$$

E step:

$$Q(\theta, \theta_0) = \sum_{I(x_i=0)} E_{P(Z|X)}[z_i]\log(\pi) + (1 - E_{P(Z|X)}[z_i])\big(\log(1-\pi) - \lambda\big) + \sum_{I(x_i>0)} \big(\log(1-\pi) + x_i\log(\lambda_i) - \lambda - \log(x_i!)\big)$$

$$\begin{aligned} E_{P(Z|X_i)}[z_i] &= 0 \times p(Z_i = 0|X) + 1 \times p(Z_i = 1|X_i = 0) \\ &= \frac{p(X_i = 0|Z_i = 1)p(Z_i = 1)}{p(X_i = 0|Z_i = 0)p(Z_i = 0) + p(X_i = 0|Z_i = 1)p(Z_i = 1)} \\ &= \frac{\pi_0}{\pi_0 + (1-\pi_0)e^{-\lambda_0}} \end{aligned}$$

Hence,

$$Q(\theta, \theta_0) = \sum_{I(x_i=0)} \frac{\pi_0}{\pi_0 + (1-\pi_0)e^{-\lambda_0}}\log(\pi) + \Big(\frac{(1-\pi_0)e^{-\lambda_0}}{\pi_0 + (1-\pi_0)e^{-\lambda_0}}\Big)\big(\log(1-\pi) - \lambda\big) + \sum_{I(x_i>0)} \big(\log(1-\pi) + x_i\log(\lambda) - \lambda - \log(x_i!)\big)$$

Problem 3 continued on next page...

M step:

$$
\begin{aligned}
\frac{\partial Q}{\partial \lambda} &= 0 \\
&= \sum_{I(x_i=0)} (1-E[z_i])(-1) + \sum_{I(x_i>0)} (\frac{x_i}{\lambda} - 1) = 0 \\
\implies \hat{\lambda} &= \frac{\sum_{I(x_i>0)} x_i}{n - \sum_{I(x_i=0)} E[z_i]} \\
\hat{\lambda} &= \frac{\sum_{I(x_i>0)} x_i}{n - \sum_{I(x_i=0)} \hat{z}_i} \\
\text{where } \hat{z} &= \frac{\pi_0}{\pi_0 + (1-\pi_0)e^{-\lambda_0}}
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial Q}{\partial \pi} &= 0 \\
&= \sum_{I(x_i=0)} (\frac{E[z_i]}{\pi} - \frac{1-E[z_i]}{1-\pi}) - \sum_{I(x_i>0)} \frac{1}{1-\pi} = 0 \\
&= \sum_{I(x_i=0)} (\frac{E[z_i]}{\pi} + \frac{E[z_i]}{1-\pi}) - \frac{n}{1-\pi} = 0 \\
\implies \hat{\pi} &= \sum_{I(x_i=0)} \frac{\hat{z}_i}{n}
\end{aligned}
$$