MTH210 (2024): Quiz 4

Name Solutions .... Roll. No. ....

Let $X_1, X_2, \dots, X_n \overset{iid}{\sim} F$, where F is a 2-component Poisson(λ) mixture distribution with mass function

$$\Pr(X = k|\lambda_1, \lambda_2, \pi_1, \pi_2) = \pi_1 p_1(k|\lambda_1) + \pi_2 p_1(k|\lambda_2),$$

where $p_1(k|\lambda_1)$ is the pmf of Poisson(λ_1), $p_2(k|\lambda_2)$ is the pmf of Poisson(λ_2), and $\pi_1 + \pi_2 = 1$. Recall that if $Y \sim$ Poisson(α), then

$$\Pr(Y = k) = \frac{\alpha^k e^{-\alpha}}{k!} \qquad k = 1, 2, 3, \dots.$$

Clearly, write down the steps of the EM Algorithm to obtain the maximum likelihood estimate of $\theta = (\pi_1, \pi_2, \lambda_1, \lambda_2)$.

Solution:
Assume iid latent variables Z_i such that $\Pr(Z_i = c) = \pi_c$ for $c = 1, 2$ and $X_i|Z_i \overset{iid}{\sim}$ Poisson(λ_c). With this, we can obtain the E and M steps of the EM algorithm. Given any iterate $\theta_{(k)}$, the E-step is similar to the Gaussian Mixture Model

E-step:

$$q(\theta|\theta_{(k)}) = \sum_{i=1}^{n} \sum_{c=1}^{2} \log\{p_c(x_i|\lambda_c)\pi_c\} \underbrace{\frac{p_c(x_i|\lambda_{c,(k)})\pi_{c,(k)}}{\sum_{j=1}^{2} p_j(x_i|\lambda_{j,(k)})\pi_{j,(k)}}}_{\gamma_{i,c,(k)}}.$$

Simplifying this,

since $\log\{p_c(x_i|\lambda_c)\pi_c\} = \log \pi_c + x_i \log \lambda_c - \lambda_c - \log(x_i!)$, we get

$$q(\theta|\theta_{(k)}) = \sum_{i=1}^{n} \sum_{c=1}^{2} \left[\log \pi_c + x_i \log \lambda_c - \lambda_c - \log(x_i!)\right] \gamma_{i,c,(k)}.$$

$$= \sum_{i=1}^{n} \left[\log \pi_1 + x_i \log \lambda_1 - \lambda_1 - \log(x_i!)\right]\gamma_{i,1,(k)} + \sum_{i=1}^{n}\left[\log(1-\pi_1) + x_i \log \lambda_2 - \lambda_2 - \log(x_i!)\right]\gamma_{i,1,(k)}.$$

4

M-step: Taking derivative since we need $\theta_{(k+1)} = \arg\max_{\theta \in \Theta} q(\theta|\theta_{(k)})$

~~M-step:~~

$$\frac{\partial q}{\partial \lambda_c} = \sum_{i=1}^{n}\left[\frac{x_i}{\lambda_c} - 1\right]\gamma_{i,c,(k)} \overset{set}{=} 0 \Rightarrow \frac{\sum_{i=1}^{n} x_i \gamma_{i,c,(k)}}{\lambda_c} = \sum_{i=1}^{n} \gamma_{i,c,(k)}$$

$$\Rightarrow \boxed{\lambda_{c,(k+1)} = \frac{\sum_{i=1}^{n} x_i \gamma_{i,c,(k)}}{\sum_{i=1}^{n} \gamma_{i,c,(k)}}} \quad \text{— } 1$$

1

6

Taking second derivative,

$$\frac{\partial^2 q}{\partial \lambda_c^2} = -\sum_{i=1}^{n} \frac{x_i}{\lambda_c^2} \cdot \gamma_{i,c,(k)} < 0$$

$\therefore$ we have a maxima

(−1) if maxima not checked

Next taking derivative w.r.t π_1

$$\frac{\partial q}{\partial \pi_1} = \sum_{i=1}^{n} \frac{\gamma_{i,1,(k)}}{\pi_1} - \sum_{i=1}^{n} \frac{\gamma_{i,2,(k)}}{1-\pi_1} \stackrel{\text{set}}{=} 0$$

$$\Rightarrow \quad (1-\pi_1) \sum \gamma_{i,1,(k)} - \pi_1 \sum \gamma_{i,2,(k)} = 0$$

$$\Rightarrow \quad \sum_{i=1}^{n} \gamma_{i,1,(k)} - \pi_1 \Big(\sum_{i=1}^{n} \underbrace{(\gamma_{i,1,(k)} + \gamma_{i,2,(k)})}_{=1} \Big) = 0$$

$$\Rightarrow \quad \boxed{\pi_{1,(k+1)} = \frac{\sum_{i=1}^{n} \gamma_{i,1,(k)}}{n}} \quad \text{--- (2)}$$

$$\text{so} \quad \boxed{\pi_{2,(k+1)} = 1 - \pi_{1,(k+1)}} \quad \text{--- (3)}$$

Checking second derivative,

$$\frac{\partial^2 q}{\partial \pi_1^2} = -\sum_{i=1}^{n} \frac{\gamma_{i,1,(k)}}{\pi_1^2} - \sum_{i=1}^{n} \frac{\gamma_{i,2,(k)}}{(1-\pi_1)^2} < 0$$

(-1) if maxima not checked

(8) so we have maxima.

Final Algo

(1) Set $\lambda_{1,(0)} = \frac{\sum x_i}{n} + 1$ and $\lambda_{2,(0)} = \max\left\{ \frac{\sum x_i}{n} - 1, 0.05 \right\}$.

$\pi_{1,(0)} = 0.5$ and $\pi_{2,(0)} = 0.5$

(2) Update $\theta_{(k)} = (\pi_{1,(k)}, \pi_{2,(k)}, \lambda_{1,(k)}, \lambda_{2,(k)})$

by setting $\theta_{(k+1)} = (\pi_{1,(k+1)}, \pi_{2,(k+1)}, \lambda_{1,(k+1)}, \lambda_{2,(k+1)})$

using eqns (1), (2) and (3)

(3) Stop when $\| \theta_{(k+1)} - \theta_{(k)} \| < \varepsilon$.

(-2) if algo not done

(10)