MTH210a (2023): Quiz 3

Name: DOOTIKA VATS   Roll. No.: Solutions.

Let $X_1, X_2, \dots, X_n \overset{iid}{\sim} F$, where F is a 2-component $\text{Exp}(\lambda)$ mixture distribution with density

$$f(x|\lambda_1, \lambda_2, \pi_1, \pi_2) = \pi_1 f_1(x|\lambda_1) + \pi_2 f_2(x|\lambda_2)\,, \qquad \lambda_1, \lambda_2 > 0 \text{ and } \pi_1, \pi_2 \in (0,1), \pi_1 + \pi_2 = 1.$$

Recall that: $f_c(x|\lambda_c) = \lambda_c e^{-\lambda_c x}$. 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 $X_i | Z_i \overset{iid}{\sim} \text{Exp}(\lambda_c)$. 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:

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

Simplifying this,

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

M-step Taking derivative w.r.t π_c, λ_c:

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

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

(−2) if calculation error.

~~M-step~~: Further $\dfrac{\partial^2 q}{\partial \lambda_c^2} = \sum_{i=1}^{n}\left(-\dfrac{\gamma_{i,c,(k)}}{\lambda_c^2}\right) < 0$ hence maxima.

(−1) if maxima not checked

Now, taking derivative w.r.t π_c, I need Lagrange multiplier

Define $\tilde{q}(\theta|\theta_{(k)}) = q(\theta|\theta_{(k)}) - \lambda\left(\sum_{c=1}^{2}\pi_c - 1\right)$

(−2) if Langrange multiplier not done

$$\frac{\partial \tilde{q}}{\partial \pi_c} = \sum_{i=1}^{n}\frac{\gamma_{i,c,(k)}}{\pi_c} - \lambda = 0 \Rightarrow \pi_c = \frac{\sum \gamma_{i,c,(k)}}{\lambda}$$

summing for all π_c: $\sum \pi_c = 1 = \dfrac{\sum_{i=1}^{n}\sum_{c=2}^{2}\gamma_{i,c,(k)}}{\lambda} = \dfrac{n}{\lambda} \Rightarrow \lambda = n$

So : $\boxed{\pi_{c,(k+1)} = \dfrac{\sum_{i=1}^{n} \gamma_{i,c,(k)}}{n}}$

Second derivative is similarly negative hence, we have maxima.

(-1) if maxima not checked

EM Algo

① Set $(\lambda_{1(0)}, \lambda_{2(0)}, \pi_{1(0)}, \pi_{2,(0)}) = \theta_{(0)}$

② Repeat:

$\cancel{\pi_{c,(k+1)}}$ calculate $\gamma_{i,c,(k)} = \dfrac{f_c(x_i|\lambda_{c,(k)})\pi_{c,(k)}}{\sum_{j=1}^{g} f_j(x_i|\lambda_{j,(k)})\pi_{j,(k)}}$

Set $\lambda_{c,(k+1)} = \dfrac{\sum_{i=1}^{n} \gamma_{i,c,(k)}}{\sum_{i=1}^{n} x_i \gamma_{i,c,(k)}}$ and $\pi_{c,(k+1)} = \dfrac{\sum_{i=1}^{n} \gamma_{i,c,(k)}}{n}$

Set $\theta_{(k+1)} = (\lambda_{c,(k+1)}, \pi_{c,(k+1)})$

③ Stop if

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

(-2) if final algorithm not written.