EM-Normal mixture

Saturday, 5 September 2020 11:25 am

$$\theta = (\pi_1, \pi_2, \cdots \pi_{K-1}, \mu_1, \cdots \mu_K, \sigma_1^2, \cdots \sigma_K^2)$$

eg $Y_i = \begin{cases} 1 & wp\ p_1 \\ 2 & wp\ p_2 \\ 3 & wp\ p_3 \end{cases}$

$$P(Y_i) = \prod_{k=1}^{3} [P(Y_i=k)]^{I(Y_i=k)} = \prod_{k=1}^{3} p_k^{I(Y_i=k)} = p_1^{I(Y_i=1)} p_2^{I(Y_i=2)} p_3^{I(Y_i=3)} = p_1+p_2+p_3$$

$$P(X_1, \cdots X_n, Z_1, \cdots Z_n \mid \theta) = \underbrace{\prod_{i=1}^{n} P(X_i \mid Z_i, \theta) P(Z_i \mid \theta)} = \prod_{i=1}^{n} \prod_{j=1}^{K} [P(X_i \mid Z_i=j, \theta) P(Z_i=j \mid \theta)]^{I(Z_i=j)}$$

$$\log P(X_1, \cdots X_n, Z_1, \cdots Z_n \mid \theta)$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{K} I(Z_i=j) \left[\log P(X_i \mid Z_i=j, \theta) + \log P(Z_i=j \mid \theta)\right]$$

$$= \sum_{i=1}^{n} \left[\sum_{j=1}^{K} I(Z_i=j) \left(-\log \sqrt{2\pi\sigma_j^2} - \frac{1}{2} \frac{(x_i-\mu_j)^2}{\sigma_j^2} + \log \pi_j\right)\right]$$

$$\pi_K = 1 - (\pi_1 + \pi_2 + \cdots \pi_{K-1})$$

$$Q(\theta, \theta^\circ) = E_{Z \mid X, \theta^\circ}\left[\log P(X_1, \cdots X_n, Z_1 \cdots Z_n \mid \theta)\right]$$

$$= \sum_{i=1}^{n} \left[\sum_{j=1}^{K} P(Z_i=j \mid X, \theta^\circ) \left(-\log \sqrt{2\pi\sigma_j^2} - \frac{1}{2} \frac{(x_i-\mu_j)^2}{\sigma_j^2} + \log \pi_j\right)\right]$$

E-step:

$j = 1, 2, \cdots K$ $\quad \theta^\circ = (\pi_1^\circ, \pi_2^\circ, \cdots \pi_{K-1}^\circ, \mu_1^\circ, \cdots \mu_K^\circ, \sigma_1^{2\circ}, \cdots \sigma_K^{2\circ})$

$$P(Z_i=j \mid X, \theta^\circ) = \frac{P(Z_i=j, X_i \mid \theta^\circ)}{P(X_i \mid \theta^\circ)}$$

$$= \frac{P(X_i \mid Z_i=j, \theta^\circ) P(Z_i=j \mid \theta^\circ)}{\sum_{k=1}^{K} P(X_i \mid Z_i=k, \theta^\circ) P(Z_i=k \mid \theta^\circ)}$$

$$= \frac{\frac{1}{\sqrt{2\pi\sigma_j^{2\circ}}} \exp\left(-\frac{(x_i-\mu_j^\circ)^2}{2\sigma_j^{2\circ}}\right) \pi_j^\circ}{\sum_{k=1}^{K} \frac{1}{\sqrt{2\pi\sigma_k^{2\circ}}} \exp\left(-\frac{(x_i-\mu_k^\circ)^2}{2\sigma_k^{2\circ}}\right) \pi_k^\circ}$$

$$\pi_K^\circ = 1 - (\pi_1^\circ + \pi_2^\circ \cdots + \pi_{K-1}^\circ)$$

M-step: $j = 1, 2, \cdots K-1$

$$\frac{\partial Q(\theta, \theta^\circ)}{\partial \pi_j} = \sum_{i=1}^{n} \left[\frac{P(Z_i=j \mid X, \theta^\circ)}{\pi_j} - \frac{P(Z_i=K \mid X, \theta^\circ)}{1-(\pi_1+\pi_2+\cdots \pi_{K-1})}\right]$$

$$= \frac{[1-(\pi_1+\pi_2+\cdots \pi_{K-1})]\left[\sum_{i=1}^{n} P(Z_i=j \mid X, \theta^\circ)\right] - \pi_j\left[\sum_{i=1}^{n} P(Z_i=K \mid X, \theta^\circ)\right]}{\pi_j[1-(\pi_1+\pi_2+\cdots \pi_{K-1})]} = 0$$

$j = 1, 2, \cdots K-1$

$$\underbrace{[1-(\pi_1+\pi_2+\cdots \pi_{K-1})]}_{\pi_K}\Big[\sum_{i=1}^{n} \underbrace{P(Z_i=j \mid X, \theta^\circ)}_{A_{ij}}\Big] - \pi_j\Big[\sum_{i=1}^{n} \underbrace{P(Z_i=K \mid X, \theta^\circ)}_{A_{iK}}\Big] = 0$$

$$A_{iK} = 1 - \sum_{j=1}^{K-1} A_{ij}$$

$j=1$ $\quad \pi_K \sum_{i=1}^{n} A_{i1} = \pi_1 \sum_{i=1}^{n} A_{iK}$

$j=2$ $\quad \pi_K \sum_{i=1}^{n} A_{i2} = \pi_2 \sum_{i=1}^{n} A_{iK}$

$j=K-1$ $\quad \pi_K \sum_{i=1}^{n} A_{iK-1} = \pi_{K-1} \sum_{i=1}^{n} A_{iK}$

sum

$$\pi_K \sum_{i=1}^{n} \sum_{j=1}^{K-1} A_{ij} = (1-\pi_K) \sum_{i=1}^{n} A_{iK}$$

$$\pi_K \sum_{i=1}^{n} 1 = \sum_{i=1}^{n} A_{iK} \qquad \pi_K = \frac{\sum_{i=1}^{n} A_{iK}}{n}$$

$$\hat{\pi}_j = \frac{1}{\sum_{i=1}^{n} A_{iK}} \pi_K \sum_{i=1}^{n} A_{ij} = \frac{1}{\cancel{\sum_{i=1}^{n} A_{iK}}} \frac{\cancel{\sum_{i=1}^{n} A_{iK}}}{n} \sum_{i=1}^{n} A_{ij} = \frac{1}{n} \sum_{i=1}^{n} P(Z_i=j \mid x_i, \theta)$$

$j = 1, \cdots, K$

$$\boxed{\begin{aligned} Q(\theta, \theta^\circ) &= E_{Z \mid X, \theta^\circ}\left[\log P(X_1, \cdots X_n, Z_1 \cdots Z_n \mid \theta)\right] \\ &= \sum_{i=1}^{n} \left[\sum_{j=1}^{K} P(Z_i=j \mid X, \theta^\circ) \left(-\log \sqrt{2\pi\sigma_j^2} - \frac{1}{2} \frac{(x_i-\mu_j)^2}{\sigma_j^2} + \log \pi_j\right)\right] \end{aligned}}$$

$$\frac{\partial Q(\theta, \theta^\circ)}{\partial \mu_j} = \sum_{i=1}^{n} P(Z_i=j \mid X, \theta^\circ)\left[-\frac{\cancel{2}(x_i-\mu_j)}{\cancel{2}\sigma_j^2} \times (-1)\right]$$

$$= \frac{1}{\sigma_j^2}\left[\sum_{i=1}^{n} P(Z_i=j \mid X, \theta^\circ) x_i - \mu_j \sum_{i=1}^{n} P(Z_i=j \mid X, \theta^\circ)\right] = 0$$

$$\hat{\mu}_j = \frac{\sum_{i=1}^{n} P(Z_i=j \mid X, \theta^\circ) x_i}{\sum_{i=1}^{n} P(Z_i=j \mid X, \theta^\circ)}$$

$j = 1, \cdots, K$ $\qquad -\log(\sigma_j^2)^{\frac{1}{2}} = -\frac{1}{2}\log \sigma_j^2$

$$\frac{\partial Q(\theta, \theta^\circ)}{\partial \sigma_j^2} = \sum_{i=1}^{n} P(Z_i=j \mid X, \theta^\circ)\left(-\frac{1}{2}\frac{1}{\sigma_j^2} - \frac{1}{2}(x_i-\mu_j)^2\left(-\frac{1}{(\sigma_j^2)^2}\right)\right)$$

$$= \sum_{i=1}^{n} P(Z_i=j \mid X, \theta^\circ) \cdot \frac{1}{2}\frac{1}{(\sigma_j^2)^2}\left(\sigma_j^2 - (x_i-\mu_j)^2\right) = 0$$

$$\hat{\sigma}_j^2 = \frac{\sum_{i=1}^{n} P(Z_i=j \mid X, \theta^\circ)(x_i-\mu_j)^2}{\sum_{i=1}^{n} P(Z_i=j \mid X, \theta^\circ)}$$