ML HALET 2 Nov. 27 2025

1-

$$D = \{(1,2),(2,3),(3,5),(4,4),(5,6)\} \rightarrow (3,4)$$

a)

$$X = \begin{bmatrix} -2 & -2 \\ -1 & -1 \\ 0 & 1 \\ 1 & 0 \\ 2 & 2 \end{bmatrix} \qquad \Sigma = \frac{1}{n-1} X^T X$$

$$X^T X = \begin{bmatrix} 10 & 9 \\ 9 & 10 \end{bmatrix}$$

$$\Sigma = \frac{1}{4} \begin{bmatrix} 10 & 9 \\ 9 & 10 \end{bmatrix}$$

$$\Sigma \gamma_1 = \lambda_1 \gamma_1 \rightarrow (\Sigma - \lambda_1 I)\gamma_1 = 0 \rightarrow \det(\Sigma - \lambda_1 I) = 0$$

$$\lambda_1 = \frac{19}{4} \quad \gamma_1 = (1,1) \leftarrow PC\ 1$$

$$\lambda_2 = \frac{1}{4} \quad \gamma_2 = (-1,1) \leftarrow PC\ 2$$

b)

$$\begin{bmatrix} -2 & -2 \\ -1 & -1 \\ 0 & 1 \\ 1 & 0 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} +1 & -1 \\ +1 & +1 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ -2 & 0 \\ +1 & +1 \\ +1 & -1 \\ +4 & 0 \end{bmatrix}$$

c)

$$\frac{19/4}{19/4 + 1/4} = 19/20 = 0.95$$

2-

$$X = \{x_1, x_2, \dots, x_n\} \qquad \gamma_{ij} = \mathbb{1}(x_i \in \text{cluster } j)$$

$$\mu_1, \mu_2, \dots, \mu_k$$

$$J = \sum_{j=1}^{k} \sum_{i=1}^{n} \gamma_{ij} \| x_i - \mu_j \|^2$$

a) each step: 1- assignment 2- update $\gamma \rightarrow k^n$

↳ decreases J

b) $T = \frac{1}{n}\sum_i \|x_i - \hat{x}\|^2 \qquad \hat{x} = \frac{1}{n}\sum_i x_i$ → واریانس کلی

→ $W_j(X) = \left(\sum_i \gamma_{ij}\|x_i - \mu_j\|^2\right) \Big/ \sum_i \gamma_{ij}$ → واریانس درون خوشه‌ای

$B(X) = \sum_j \left(\frac{\sum_i \gamma_{ij}}{n}\right)\|\mu_j - \hat{x}\|^2$ → cluster-level variance

$$T = \underbrace{\sum_j \left(\frac{\sum_i \gamma_{ij}}{n}\right) W_j(X)}_{\propto J} + B(X) \qquad J = \sum_j \sum_i \gamma_{ij}\|x_i - \mu_j\|^2$$

c) $\min J(k) \geq \min J(k+1) \qquad k = n \qquad J \geq 0$

$\gamma_{ij} = I_{n\times n} \rightarrow J = 0$

3 –

a) $P(x, z; \theta) = \underbrace{P(z;\theta)}\, P(x \mid z;\theta) = \underbrace{\pi^{z}(1-\pi)^{(1-z)}}\, P_r^{xz} P_b^{x(1-z)} (1-P_r)^{(1-x)z} (1-P_b)^{(1-x)(1-z)}$

$$P(z;\theta) = \begin{cases} P(1;\theta) = \pi & z = 1 \\ P(0;\theta) = (1-\pi) & z = 0 \end{cases} \qquad \ln a^b = b \ln a$$

b) $$L_c^{(\theta)} = \sum_{i=1}^{m} \Big[z_i \ln \pi + (1-z_i)\ln(1-\pi) + x_i z_i \ln P_r + x_i(1-z_i)\ln P_b + (1-x_i) z_i \ln(1-P_r) + (1-x_i)(1-z_i)\ln(1-P_b) \Big]$$

c) $\frac{\partial L_c^{(\theta)}}{\partial \pi}\Big|_{\pi^*} = 0 \rightarrow \pi^* = ? \quad \frac{\sum_i z_i}{m} \qquad \hat{P}_r = \frac{\sum_i x_i z_i}{\sum z_i} \qquad \hat{P}_b = \frac{\sum_i x_i - \sum_i x_i z_i}{m - \sum_i z_i}$

$\frac{\partial L_c}{\partial P_a}\Big|_{P_a^*} = 0 \nearrow$

$\frac{\partial L_c}{\partial P_b}\Big|_{P_b^*} = 0 \nearrow$

d) $$P(\underset{Z_i =}{z_i = 1} \mid \underset{X_i =}{x_i}; \theta^{(t)}) = \frac{P(x_i \mid z_i = 1; \theta^t)\, P(z_i = 1; \theta^t)}{\left[P(x_i \mid z_i = 1; \theta^t)\, P(z_i = 1; \theta^t) + P(x_i \mid z_i = 0; \theta^t)\, P(z_i = 0; \theta^t)\right]} \checkmark$$

E

$$= \frac{\left[P_r^{t\,x_i}(1-P_r^t)^{1-x_i}\right]\pi^t}{P_r^{t\,x_i}(1-P_r^t)^{1-x_i}\pi^t + P_b^{t\,x_i}(1-P_b^t)^{1-x_i}(1-\pi^t)} \checkmark$$

e) M $\quad \gamma_i^t = P(z_i = 1 \mid x_i ; \theta^t) \qquad Q = E_{z \mid x, \theta^t}\left[\ln P(x, z \mid \theta)\right]$

$$Q = \sum_{i=1}^{m} \left(\gamma_i^t \ln P(x_i ; z=1; \theta) + (1-\gamma_i^t) \ln P(x_i ; z=0; \theta) \right) =$$

$$= \sum_{i=1}^{m} \Big[\gamma_i^t \left(\ln \pi + x_i \ln P_r + (1-x_i) \ln(1-P_r) \right) + (1-\gamma_i^t)\left(\ln(1-\pi) + x_i \ln P_b + (1-x_i) \ln(1-P_b) \right) \Big]$$

$\partial / \partial \pi$:

$z_i \to \gamma_i^t$

$$\sum_{i=1}^{m} \left(\frac{\gamma_i^t}{\pi} - \frac{(1-\gamma_i^t)}{1-\pi} \right) \Big|_{\pi^{t+1}} = 0 \rightarrow \frac{1}{\pi} \sum_i \gamma_i^t = \frac{1}{1-\pi} \underbrace{\sum_i (1-\gamma_i^t)}_{m - \sum_i \gamma_i^t}$$

$$\rightarrow m \pi^{t+1} = \sum_i \gamma_i^t \rightarrow \pi^{t+1} = \frac{\sum_i \gamma_i^t}{m} \qquad P_r^{t+1} = \frac{\sum_i x_i \gamma_i^t}{\sum_i \gamma_i^t}$$

$$P_b^{t+1} = \frac{\sum_i x_i - \sum_i x_i \gamma_i^t}{m - \sum_i \gamma_i^t}$$

4-

$$P(x_i \mid \mu, \Sigma) = \frac{1}{(2\pi)^{d/2} |\Sigma|^{1/2}} \exp\left(-\frac{1}{2} (x_i - \mu)^T \Sigma^{-1} (x_i - \mu) \right)$$

$$L = P(X \mid \mu, \Sigma) = \prod_{i=1}^{N} P(x_i \mid \mu, \Sigma) \qquad \log(a/b) = \log a - \log b$$

$$\log L = \sum_i \log P(x_i \mid \mu, \Sigma) = \sum_i \left[-\frac{d}{2} \log(2\pi) - \frac{1}{2} \log |\Sigma| - \frac{1}{2} (x_i - \mu)^T \Sigma^{-1} (x_i - \mu) \right] =$$

$$= -Nd/2 \log(2\pi) - \frac{N}{2} \log |\Sigma| - \frac{1}{2} \sum_i (x_i - \mu)^T \Sigma^{-1} (x_i - \mu)$$

$$\frac{\partial \log L}{\partial \mu} = -\frac{1}{2} \sum_i \partial/\partial\mu \left[(x_i - \mu)^T \Sigma^{-1} (x_i - \mu) \right] = -1/2 \sum_i \left[-2 \Sigma^{-1} (x_i - \mu) \right] =$$

$$= \sum_i \Sigma^{-1} (x_i - \mu) = \Sigma^{-1} \sum_i (x_i - \mu) = 0 \rightarrow \sum_i x_i = n\mu$$

$$\mu = \frac{1}{n} \sum_i x_i$$

b) $\quad \pi_1 N(x \mid \mu, \Sigma) + \pi_2 N(x \mid \mu, \Sigma) = (\pi_1 + \pi_2) \cdots$

c)

$$P(z=1 \mid x) \propto \pi_1 \mathcal{N}(x \mid \mu_1, \sigma^2)$$

$$P(z=2 \mid x) \propto \pi_2 \mathcal{N}(x \mid \mu_2, \sigma^2)$$

$$\pi_1 \mathcal{N}(x \mid \mu_1, \sigma^2) = \pi_2 \mathcal{N}(x \mid \mu_2, \sigma^2) \rightarrow \pi_1 e^{-\frac{(x-\mu_1)^2}{2\sigma^2}} = \pi_2 e^{-\frac{(x-\mu_2)^2}{2\sigma^2}}$$

$$\rightarrow \ln \pi_1 - \frac{(x-\mu_1)^2}{2\sigma^2} = \ln \pi_2 - \frac{(x-\mu_2)^2}{2\sigma^2} \quad *$$

$$\ln \pi_1 - \ln \pi_2 = \frac{(x-\mu_1)^2 - (x-\mu_2)^2}{2\sigma^2} = -2\mu_1 x + 2\mu_2 x + \mu_1^2 - \mu_2^2 = 2\sigma^2 \ln\left(\frac{\pi_1}{\pi_2}\right)$$

$$\rightarrow 2x^* (\mu_2 - \mu_1) = 2\sigma^2 \ln(\pi_1/\pi_2) - (\mu_1 - \mu_2)(\mu_1 + \mu_2)$$

$$\rightarrow x^* = \frac{\mu_1 + \mu_2}{2} + \frac{\sigma^2}{\mu_2 - \mu_1} \ln(\pi_1/\pi_2)$$