Derivation $(G=2)$

① Incomplete likelyhood function:

$$p_1 N(x; \mu_1, \Sigma) + p_2 N(x; \mu_2, \Sigma)$$

where data $= [x_1, \dots x_n] = \mathbf{X} \in R^2$

② Complete likelyhood function (with latent variable Z)

$$P_\theta(X, Z) = \begin{cases} p_1 N(X; \mu_1, \Sigma) & \text{for } Z=1 \\ p_2 N(X, \mu_2 \Sigma) & \text{for } Z=2 \end{cases}$$

$$= \quad P_\theta(X, Z) = p_1 N(X; \mu_1 \Sigma)^{\pi_{Z=1}} \left[p_2 N(X; \mu_2 \Sigma)\right]^{\pi_{Z=2}} \quad \text{—①}$$

Take log on both sides of ①

$$\log P_\theta(X, Z) = \pi_{Z=1}\Big[\log p_1 - \frac{1}{2}\log|\Sigma| - \frac{1}{2}(x-\mu_1)^T \Sigma^{-1}(x-\mu_1)\Big]$$
$$+ \pi_{Z=2}\Big[\log p_2 - \frac{1}{2}\log|\Sigma| - 1/2 (x-\mu_2)^T \Sigma^{-1} (x-\mu_2)\Big]$$

log Joint Likelyhood with individual x_i

$$\log \prod_{i=1}^{n} p(x_i, z_i) = \sum_{i=1}^{n} \log p(x_i, z_i)$$

$$= \sum_{i=1}^{n} \Big\{ \pi_{z_i=1}\Big[\log p_1 - \frac{1}{2}\log|\Sigma| - \frac{1}{2}(x_i-\mu_1)^T\Sigma^{-1}(x_i-\mu_1)\Big]$$

$$+ \pi_{z=2}\Big(\log p_2 - \frac{1}{2}\log|\Sigma| - \frac{1}{2}(x_i-\mu_2)^T\Sigma^{-1}(x_i-\mu_2)\Big) \quad \text{—②}$$

③ E-step : $Z_1 \cdots Z_n \mid X, \theta^{(0)}$; $Z_i \mid X_i, \theta^{(0)} \sim \text{Multi}(P_{i1}, P_{i2})$

(distribution of Z_i)

$$\Rightarrow P_{\theta^{(0)}}(Z_i = 1 \mid X_i) = \frac{P(Z_i=1, x_i)}{P(x_i)} = \frac{P(Z_i=1, X_i)}{P(Z_i=1, X_i) + P(Z_i=2, X_i)}$$

$$\boxed{= p_1^{(0)} \cdot N(X_i, \mu_1^{(0)}, \Sigma^{(0)})}$$

and similarly

$$P_{\theta^{(0)}}(Z_i = 2 \mid X_i) = \boxed{P_2^{(0)} \cdot N(X_i, \mu_2^{(0)}, \Sigma^{(0)})}$$

④ M-step (objective function)

Taking Expectation of ② wrt Z

$$E\left(\log \prod_i^{n} p(x_i, z_i)\right) = E\Big[\sum_{i=1}^{n} \Big\{ \prod_{z=1} \Big[\log p_1 - \frac{1}{2}\log|\Sigma| - \frac{1}{2}(x_i-\mu_1)^T \Sigma^{-1}(x_i-\mu_1)\Big] + \prod_{z=2}\Big[\log p_2 - \frac{1}{2}\log|\Sigma| - \frac{1}{2}(x_i-\mu_2)^T \Sigma^{-1}(x_i-\mu_2)\Big]\Big\}\Big]$$

$$= p_{i1}\Big[\log p_1 - \frac{1}{2}\log|\Sigma| - \frac{1}{2}(x_i-\mu_1)^T \Sigma^{-1}(x_i-\mu_2)\Big] + p_{i2}\Big[\log p_2 - \frac{1}{2}\log|\Sigma| - \frac{1}{2}(x_i-\mu_2)^T \Sigma^{-1}(x_i-\mu_2)\Big]$$

$$\boxed{= J(p_1, p_2, \mu_1, \mu_2, \Sigma) \text{ — objective function}}$$

Update Formula

$$p_1 = \frac{1}{N}\sum_{i=1}^{N} p_{i1} \quad ; \quad p_2 = \frac{1}{N}\sum_{i=1}^{N} p_{i2}$$

$$\mu_1 = \frac{\sum_{i=1}^{n} p_{i1} \cdot x_i}{\sum_{i=1}^{n} p_{i1}} \quad ; \quad \mu_2 = \frac{\sum_{i=1}^{n} p_{i2} \cdot x_i}{\sum_{i=1}^{n} p_{i2}}$$

$$\Sigma = \frac{\sum_{i=1}^{n} p_{i1}(x_i - \hat{\mu}_1)(x_i - \hat{\mu}_1)^t + \sum p_{i2}(x_i - \hat{\mu}_2)(x_i - \hat{\mu}_2)^t}{2}$$