Q1. a

$$p(z_n = k \mid x_n) = \frac{p(x_n, z_n = k)}{p(x_n)} \quad (\because \text{Bayes Rule})$$

$$\Leftrightarrow p(z_n = k \mid x_n) = \frac{p(x \mid z_n = k)\, p(z_n = k)}{\sum_{j=1}^{K} p(x_n, z_n = j)}$$

$$\Leftrightarrow p(z_n = k \mid x_n) = \frac{\mathcal{N}(x_n \mid \mu_k, \Sigma_k)\, \pi_k}{\sum_{j=1}^{K} \mathcal{N}(x_n \mid \mu_j, \Sigma_j)\, \pi_j}$$

(Q.E.D)

Q1.b

$$\underset{\theta}{\operatorname{argmax}} \sum_{n=1}^{N} \sum_{k=1}^{K} r_{nk} \log P(x_n, z_n = k \mid \theta)$$

$$\underset{\theta}{\operatorname{argmax}} \sum_{n=1}^{N} \sum_{k=1}^{K} r_{nk} \left\{ \log \mathcal{N}(x_n \mid \mu_k, \Sigma_k) + \log \pi_k \right\} = \alpha$$

1) find μ_k

$$\frac{\partial \alpha}{\partial \mu_k} = \sum_{n=1}^{N} \frac{r_{nk}}{\mathcal{N}(x_n \mid \mu_k, \Sigma_k)} \times \mathcal{N}(x_n \mid \mu_k, \Sigma_k) \cdot \Sigma_k^{-1}(x_n - \mu_k)$$

$$= \sum_{n=1}^{N} r_{nk} \cdot \Sigma_k^{-1}(x_n - \mu_k)$$

$$= \Sigma_k^{-1} \sum_{n=1}^{N} r_{nk}(x_n - \mu_k) = 0 \qquad \Leftrightarrow \mu_k = \frac{1}{N_k} \sum_{n=1}^{N} r_{nk} x_n \ //$$

2) find Σ_k

$$\frac{\partial \alpha}{\partial \Sigma_k} = \sum_{n=1}^{N} r_{nk} \left(\frac{1}{2} \Sigma_k^{-1} (x_n - \mu_k)(x_n - \mu_k)^T \Sigma_k^{-1} - \frac{1}{2} \Sigma_k^{-1} \right) = [0]$$

$$\Leftrightarrow \sum_{n=1}^{N} r_{nk} \left\{ \Sigma_k^{-1} (x_n - \mu_k)(x_n - \mu_k)^T \Sigma_k^{-1} - \Sigma_k^{-1} \right\} = [0]$$

$$\Leftrightarrow \sum_{n=1}^{N} r_{nk} (x_n - \mu_k)(x_n - \mu_k)^T - \Sigma_k \sum_{n=1}^{N} r_{nk} = [0] \Leftrightarrow \Sigma_k = \frac{1}{N_k} \sum_{n=1}^{N} r_{nk} (x_n - \mu_k)(x_n - \mu_k)^T \ //$$

3) find π_k, since $\sum_{k=1}^{K} \pi_k = 1$, instead of maximizing α, we will maximize

$$\alpha' = \alpha + \lambda \left(\sum_{k=1}^{K} \pi_k - 1 \right)$$

$$\frac{\partial \alpha'}{\partial \pi_k} = \sum_{n=1}^{N} \frac{r_{nk}}{\pi_k} + \lambda = 0 \Leftrightarrow \pi_k = -\frac{N_k}{\lambda}$$

Enforcing the constraint $\sum_{k=1}^{K} \pi_k = 1$, $\lambda = -N$

Therefore,

$$\pi_k = \frac{N_k}{N} \ //$$