Q2

$L =$ log likelihood of EM for GMM

$$L = \sum_{n=1}^{N} \ln\left(\sum_{k=1}^{K} \pi_k N(x_n \mid \mu_k \Sigma_k)\right)$$

Multivariate Gaussian Distribution

$$N(x \mid \mu, \Sigma) = \frac{1}{(2\pi |\Sigma|)^{1/2}} \cdot \exp\left(\frac{-1}{2}(x-\mu)^T \Sigma^{-1} \cdot (x-\mu)\right) \rightarrow (1)$$

$$\boxed{\frac{\partial}{\partial A}(x-A)^T \cdot W \cdot (x-A) = -2W(x-A)}$$

From the above

Differentiating w.r.t μ_k

$$\frac{\partial L}{\partial \mu_k} = \sum_{i=1}^{N} \frac{\pi_k N_k}{\sum \pi_k \cdot N_k} \cdot \Sigma_k^{-1} (x - \mu_k)(-1) = 0$$

From Bayes rule.

$$V_k(x) = p(k \mid x) = \frac{p(k) \cdot p(x \mid k)}{p(x)}$$

$$= \frac{\pi_k N(x \mid \mu_k, \Sigma_k)}{\sum_{j=1}^{K} \pi_j N(x \mid \mu_j, \Sigma_j)} \quad \text{where } \pi_k = \frac{N_k}{N}$$

Multiply both sides by Σ_k to remove Σ_k^{-1}

$$\sum_{i=1}^{N} r(z_{mk})(x_m - \mu_k)$$

$$\mu_k = \frac{\sum_{1}^{N} r(z_{mk})\cdot x_m}{\sum_{1}^{N} r(z_{mk})}$$

$$\boxed{\mu_k = \frac{1}{N_k}\cdot \frac{\sum r(z_{mk})\cdot x_m}{}} \qquad N_k = \sum_{1}^{N} r(z_{mk})$$

To compute Σ

$$\frac{\partial L}{\partial \Sigma} = \frac{\partial}{\partial \Sigma} \sum_{m=1}^{N} \ln\left(\sum_{k=1}^{K} \pi_k N(x_m \mid \mu_k \Sigma_k)\right)$$

$$\frac{d}{dx}\log(f(x)) = \frac{1}{f(x)} \cdot f'(x)$$

$$\frac{\partial L}{\partial \Sigma} = \sum_{i=}^{N} \frac{1}{\sum \pi_k \cdot N(x_m \mid \mu_k \Sigma_k)} \cdot \frac{\partial}{\partial \Sigma}\left(\sum_{k=1}^{K} \cdot \pi_k \cdot N(x_m \mid \mu_k \Sigma_k\right)$$

From Eq ①

$$N(x_m \mid \mu_k \Sigma_k) = \frac{1}{(2\pi|\Sigma|)^{1/2}} \cdot \exp\left(-1/2\,(x-\mu)^T \Sigma^{-1} (x-\mu)\right)$$

Using product rule of derivatives

$$\boxed{\frac{d}{dx}(u(x)\cdot v(x)) = u(x)\cdot v'(x) + u'(x)\cdot v(x)}$$

$$\frac{\partial L}{\partial \Sigma} = \frac{1}{(2\pi)^{1/2}} \cdot |\Sigma|^{-1/2} \cdot \frac{\partial |\Sigma|}{\partial \Sigma} \cdot e^{\left(-1/2 (x-\mu)^T \cdot \Sigma^{-1} \cdot (x-\mu)\right)}$$

$$+ e^{\left(-1/2 (x-\mu)^T \cdot \Sigma^{-1} \cdot (x-\mu)\right)} \cdot \frac{1}{(2\pi |\Sigma|)^{1/2}} \cdot (x-\mu)^T \cdot (x-\mu) \cdot (-1)$$

Reading through literature

$$\frac{\partial |\Sigma|}{\partial \Sigma} = |\Sigma| \cdot \Sigma^{-1} \rightarrow$$ (Struggling to prove this)

$$\frac{\partial L}{\partial \Sigma_k} = 0 = \sum_{m=1}^{N} \left(\frac{\pi_k e^{(-1/2}}{\sum_{k=1}^{K} \pi_k \cdot N_{\mu,\Sigma}} \cdot \left[\frac{1}{(2\pi)^{1/2}} \frac{\Sigma}{|\Sigma|^{1/2}} - \frac{1}{2\pi |\Sigma|^{1/2}} (x_m - \mu_m)(x_m - \mu_m)^T \right] \right.$$

$$0 = \sum_{m=1}^{N} r(Z_{m,k}) \left[\Sigma_k - (x_m - \mu_m)(x_m - \mu_m)^T \right]$$

$$\Sigma_k \cdot r(Z_{m,k}) = r(Z_{mk}) (x - \mu_k)(x - \mu_k)^T$$

$$\boxed{\Sigma_k = \frac{1}{N_k} \sum_{m=1}^{N} r(Z_{mk}) (x - \mu_k)(x - \mu_k)^T}$$

3) For π_k

The maximization of the log likelihood is constrained by $\sum \pi_k = 1$. Hence taking Lagrange multipliers

$$L(\pi_k) = L + \lambda \left(\sum_{k=1}^{K} \pi_k - 1 \right) = 0$$

$$\frac{\partial L(\pi)}{\partial \pi_j} = \sum_{m=1}^{N} \frac{N(x_m | \mu_k \Sigma_k)}{\sum_j \pi_k N(x_m |} + \lambda$$

$$\sum_{m=1}^{N} \frac{N(x_m \mid \mu_k, \Sigma_k)}{\sum_{j=1}^{K} \pi_j \cdot N(x_m \mid \mu_j \Sigma_j)} + \lambda = 0$$

$$\sum_{m=1}^{N} \frac{N(x_m \mid \mu_k \Sigma_k)}{\sum_{j=1}^{K} \pi_j \cdot N(x_m \mid \mu_j \Sigma_j)} = -\lambda \quad - (4)$$

Multiplying by π_k on both sides and summing from 1 to J

$$\Rightarrow -\lambda \sum_{1}^{J} \pi_j = \sum \frac{N(x_m \mid \mu_k \Sigma_k)}{\sum \pi_j \cdot N(x_m \mid \mu_j \Sigma_j)}$$

$$= \lambda = -N$$

Substituting in (4)

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