Q2(d) The expectation of the log likelihood function with regularization looks like the following.

$$\mathcal{L} = \sum_{t=1}^{N} \sum_{i=1}^{K} \gamma(z_i^t) \left(\log \pi_i + \log \frac{1}{(2\pi)^{d/2} |\Sigma_i|^{\frac{1}{2}}} \cdot e^{1/2 (x^t - \mu_i) \Sigma^{-1} (x^t - \mu_i)'} \right)$$

$$- \frac{\lambda}{2} \sum_{i=1}^{K} \sum_{j=1}^{d} (\Sigma_i^{-1})_{jj}$$

On taking the partial derivative of $\mathcal{L}$ w.r.t to Σ^{-1} and equating that to 0 we have

$$\frac{\partial \mathcal{L}}{\partial \Sigma^{-1}} = \sum_{t=1}^{N} (z_i^t) \left(\frac{\Sigma_i}{2} - \frac{1}{2} (x^t - \mu_i) (x^t - \mu_i)^T \right) - \frac{\lambda I}{2}$$

$$\frac{\partial L}{\partial \Sigma^{-1}} = 0$$

$$\Rightarrow \sum_{t=1}^{N} (z_i^t) \left(\frac{\Sigma_i}{2} - \frac{1}{2} (x^t - \mu_i) (x^t - \mu_i)^T \right) - \frac{\lambda I}{2}$$

$$\frac{\Sigma}{2} N_i = \frac{1}{2} \sum_{i=1}^{N} \gamma(z_i^t) \left(\frac{\Sigma_i}{2} - \frac{1}{2} (x^t - \mu_i) (x^t - \mu_i)^T \right) + \frac{\lambda I}{2}$$

$$\cancel{\Sigma = \frac{1}{N} \sum_{i=1}^{N} \gamma(z_i^t) (x_i^t - \mu_i) (x_i^t - \mu}$$

$$\Sigma = \frac{1}{N} \sum_{t=1}^{N} \gamma(z_i^t) (x^t - \mu_i) (x^t - \mu_i)^T + \frac{\lambda I_i}{N_i}$$

Therefore we have

$$\Sigma = \frac{1}{N} \sum_{i=1}^{N} \gamma(z_i^t) (x^t - \mu_i) (x^t - \mu_i)^T + \frac{\lambda I}{N_i}$$

Scanned with CamScanner