HW #6

1. The data log likelihood is :

$$l(\mu_k, \Sigma_k) = \sum_k \sum_i r_{ik} \log P(x_i | \theta_k)$$

$$= -\frac{1}{2} \sum_i r_{ik} \left(\log |\Sigma_k| + (x_i - \mu_k)^T \Sigma_k^{-1} (x_i - \mu_k) \right)$$

Differentiating wrt μ_k :

$$\frac{\partial l}{\partial \mu_k} = \sum_i r_{ik} \Sigma_k^{-1} (x_i - \mu_k)$$

$$= \Sigma_k^{-1} \sum_i r_{ik} (x_i - \mu_k) = 0$$

$$\sum_i r_{ik} x_i = \mu_k \sum_i r_{ik}$$

(Had to refer to solution)

Differentiating wrt Σ_k,

$$\frac{\partial l}{\partial \Sigma_k} = -\frac{1}{2} \sum_i r_{ik} \left(\Sigma_k^{-1} - \Sigma_k^{-1} (x_i - \mu_k)(x_i - \mu_k)^T \Sigma_k^{-1} \right) = 0$$

Multiplying by Σ_k and dividing by $r_k = \sum_i r_{ik}$ to get required results.

2. Refer to attached plot and pictures.