consulted solution

1. Show that the M step for ML estimation of a Gaussians is given by

$$\mu_k = \frac{\sum_i r_{ik} x_i}{r_k} \qquad \Sigma_k = \frac{1}{r_k} \sum_i r_{ik} (x_i - \mu_k)(x_i - \mu_k)^T$$

$$= \frac{1}{r_k} \sum_i r_{ik} x_i x_i^T - r_k \mu_k \mu_k^T.$$

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

By differentiating w.r.t. μ_k we obtain

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

so at optimality we have $\sum_i r_{ik} x_i = \mu_k \sum_i r_{ik}$, which gives us the desired result.

By differentiating w.r.t. Σ_k we obtain

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

the optimality condition that $\sum_i r_{ik} I = \left(\sum_i r_{ik} (x_i - \mu_k)(x_i - \mu_k)^T\right) \Sigma_k^{-1}$

By multiplying the right side by Σ_k and dividing by $r_k = \sum_i r_{ik}$, we obtain the desired result.