Alex Fay HW6

Note: Question 2 is posted on GitHub w/ code separately

① Show that the M step for ML estimation of Mixture of Guassians given by:

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

— — Proof — —

ⓐ I will first show that $\mu_k = \sum_i r_{ik} x_i / r_k$

We can show this by first finding the rate of change in the log likelihood

$$l(\mu_k, \Sigma_k) = \sum_k \left(\sum_i r_{ik} \log P(x, \theta_k) \right)$$

Substituting Σ_k from first half given above,

$$= \left(\frac{1}{r_k} \sum_i r_{ik}(x-\mu_k)(x-\mu_k)^T \right) \left(\sum_i r_{ik} \log P(x_i \mid \theta_k) \right)$$

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

For the rate of change,

$$\frac{\partial l}{\partial \mu_k} = \frac{\partial l}{\partial \mu_k} \left[\text{function above} \right]$$

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

To optimize the above so that we find μ_k

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

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

$$\boxed{\mu_k = \sum_i (r_{ik} x_i) / r_{ik}}$$

(b) To find Σ_k, we find the change w/ respect to Σ_k

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

Optimizing this similar to part a,

$$0 = \frac{\partial l}{\partial \Sigma_k}$$

$$\frac{\sum_i r_{ik} I (\Sigma_k)}{\left(\sum r_{ik}\right)} = \frac{\left(\sum_i r_k (x_i-\mu_k)(x_i-\mu_k)^T\right)\cancel{\Sigma_k^{-1}}\,(\cancel{\Sigma_k})}{\sum_i r_{ik}}$$

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