Problem 1 : KMeans Proof

objective = $\min \sum_i \sum_K \pi_{iK} \cdot \|x_i - \mu_K\|^2$

→ E step → updating π_{iK}

As objective function is linear function of π_{iK} we can choose π_{iK} to be 1 for whichever value gives the minimum $\|x_i - \mu_K\|^2$. Which essentially means we are assigning j^{th} data point to the closest cluster. Thus it can be written as

$$\pi_{iK} = \begin{cases} 1 & \text{if } \min \|x_i - \mu_K\|^2 \\ 0 \end{cases}$$

→ M step → updating μ_K

This can be minimized by taking derivative of objective which gives

$$2 \sum_i \pi_{iK} (x_i - \mu_K) = 0$$

$$\Rightarrow \quad \mu_K = \frac{\sum \pi_{iK} x_i}{\sum_i \pi_{iK}}$$

Which is essentially the centeroid of a particular cluster.

→ Convergence : For each iteration of algorithm (Estep & Mstep) we produce a new cluster based on previous cluster. Then:

- If previous cluster is same as new cluster, then we stop as next cluster will again be the same
- If new cluster is different from previous, then the new one has a lower objective value (as Estep, Mstep will find a minimum)

Since the algorithm iterates over the above two conditions it is expected to converge at a local minimum. Global minima isn't guaranteed as initiallization is random