Problem 4.1

$$J = \sum_{n=1}^{k} \sum_{k=1}^{k} r_{nk} \|\vec{x_n} - \vec{\mu_k}\|^2$$

n = index at data

k = index of clusters

$r_{nk} \in \{0,1\}$ represents which of the K clusters x_n is assigned to. cluster k

μ_k centroid for cluster k

Since J is a function of r_{nk} and $\|\vec{x_n} - \vec{\mu_k}\|^2$, J is non-negative. We have to find values of r_{nk} and μ_k such that J is minimized. Closed form solution for each iteration can be obtained by setting either $\frac{\partial J}{\partial r_{nk}} = 0$ or $\frac{\partial J}{\partial \mu_k} = 0$.

$$r_{nk} = \begin{cases} 1 & \text{if } k = \arg\min \|\bar{x}_n - \bar{\mu}_k\|^2 \\ 0 & \text{otherwise} \end{cases} \quad + \quad \mu_k = \frac{\sum_n r_{nk} \vec{x}_n}{\sum_n r_{nk}}$$

Since closed form solution exists local minima for J exists. As soon as either r_{nk} or μ_k converges, the other parameter will converge and assignment enters a cycle. Since there are finite number of possible assignments convergence is ensured. as J is a monotonically dec function

Problem 4.2

$$p(x) = \sum_{z} p(\bar{x}|\bar{z})\, p(z)$$

$$= \sum_{z} \prod_{k=1}^{k} N(x|\mu_k, \Sigma_k)^{z_k} \pi_k^{z_k}$$

$$= \sum_{z} \prod_{k=1}^{k} \left(N(x|\mu_k, \Sigma_k)\pi_k\right)^{z_k}$$

$$p(x) = \sum N(\vec{x}|\vec{\mu_k}, \Sigma_k)\pi_k$$

Problem 4.3

PCA is focused on finding the orthogonal projections of the data set with highest variance possible. If there is noise and we plot the data. If the data is not linearly correlated (i.e in spiral) PC doesnot work. It always assumes that the data is a Gaussian destribution