1

Since the distortion function:
$J = \sum_{n=1}^{N} \sum_{k=1}^{K} r_{nk} ||x_n - \mu_k||^2$
and the dissimilarity function:
$||x_n - \mu_k||^2$

if we want to find the values for $\{r_{nk}\}$ and $\{\mu_k\}$ to minimize J. We can perform a iteration of two steps: first,minimize J with respect to the $\{r_{nk}\}$, keeping the $\{\mu_k\}$ fixed. Second, minimize J with respect to the $\{\mu_k\}$, keeping $\{r_{nk}\}$ fixed.

when minimize J respect to μ_k. Take derivative with respect to μ_k.

$\frac{dJ}{d\mu_k} = 2\sum_{n=1}^{N} r_{nk}(x_n - \mu_k) = 0$

$\sum_{n=1}^{N} r_{nk}(x_n - \mu_k) = 0$

$\sum_{n=1}^{N} r_{nk}x_n = \sum_{n=1}^{N} r_{nk}\mu_k$

$\mu_k = \frac{\sum_{n=1}^{N} r_{nk}x_n}{\sum_{n=1}^{N} r_{nk}}$

Since μ_k is the mean of all of the data points x_n assigned to cluster k, it also means that μ_k is the center of k-th cluster.

2

The steps of calculating the cluster means and reassign data points to clusters are repeated until no further change in the clusters or maximal number of iteration is exceeded.Since each iteration reduces the value of the distortion function, convergence is guaranteed. Since we are fixing some of the variables such as r_{nk} and μ_k in each step, this is a local search procedure which does not guarantee optimality.

3

Since the K-mean clustering calculates the center of cluster by averaging all the data points in the same cluster. The average linkage would most likely result in clusters most similar to those given by K-means.

4

The single linkage would successfully separate the two moons. Because for bottom-up hierarchical clustering, we start by assgin each data points to their own cluster, then two closest clusters are merged till we have just one cluster at the top. As we can see from the figure, the distance between any pairs of neighbouring points in the same moon is shorter than the distance between the closest pairs of points from the two moons. Thus by using the single linkage metrics, all the data points in the same moon will be clustered together first before any point from the other moon can be included in the same cluster.