HW-3

Q 1 (1) Center of first cluster:

Center of first cluster after first Iteration is

X	Y
5.171	3.171

(2) Center of Second cluster

Center of second cluster after two iteration is

X	Y
5.300	4.000

(3) Center of third cluster

Center of third cluster when clustering converges is

X	Y
6.200	3.025

(4) It required 2 iterations for clusters to converge.

Q2

$$L = \sum_{j=1}^{k} \sum_{x_i \in S_j} ||x_i - z_j||^2$$

$k \rightarrow$ clusters

$x_1, x_2, \ldots, x_n \rightarrow$ sample points

$z_1, z_2, \ldots, z_k \rightarrow$ Centers

$S_j \rightarrow$ Set of points that are close to z_j

(1) L function for z_1 will be

$$L = \sum_{x \in S_1} ||x_i - z_1||^2$$

No taking differiation of L with respect to z_1.

$$\frac{\partial L}{\partial z_1} = \frac{\partial}{\partial z_1} \sum_{x_i \in S_1} ||x_i - z_1||^2$$

$$= \partial / \partial z_1 \sum_{x_i \in S_1} 2(x_i - z_1)$$

$$= \sum_{x_i \in S_1} 2(-x_i + z_1)$$

And hence the updation rule will be

$$z_1 \leftarrow z_1 + 2\varepsilon \sum_{x_i \in S_1} (x_i - z_1) \quad \text{-- [1]}$$

(2)

$$u_1 = \begin{cases} u_1 + 2\varepsilon(x_i - u_1) & \text{if } x_i \in S_1 \\ \text{otherwise} & \text{change nothing} \end{cases}$$

Note:- u_1 will be updated after each ~~iteration~~ point.

(3) we can derive value of ε by comparing ~~[illegible]~~ [1] and updation in standard algoritham.

In general K-means we assign for u_1

$$u_1 \leftarrow \sum_{x_i \in S_1} {}^1/_{|S_1|} \, x_i \qquad \text{- - - [2]}$$

Now lets compere [1] and [2]

$$u_1 + 2\varepsilon \sum_{x_i \in S_1} (x_i - u_1) = \sum_{x \in S_1} \frac{1}{|S_1|} x_i$$

$$2\varepsilon \sum_{x_i \in S_1} (x_i - u_1) = \sum_{x \in S_1} \frac{1}{|S_1|} x_i - \sum_{x \in S_1} \frac{1}{|S_1|} u_1$$

$$2\varepsilon \sum_{x \in S_1} (x_i - u_1) = \sum_{x \in S_1} \frac{1}{|S_1|} (x_i - u_1)$$

$$2\varepsilon = \sum_{x \in S_1} \frac{1}{|S_1|} (x_i - u_1) \Big/ \sum_{x \in S_1} (x_i - u_1)$$

$$2\varepsilon = {}^1/_{|S_1|}$$

$$\boxed{\varepsilon = \frac{1}{2|S_1|}}$$

Q-3 (1) Write down the compact form of $p(z)$ and $p(x|z)$

The compact form of $p(z)$

~~p(z)~~

$$p(z) = \prod_{k=1}^{K} \pi_k^{z_k}$$

The compact form of $p(x|z)$

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

(2) Show that the marginal distribution $p(x)$ has the following form.

$$p(x) = \sum_{k=1}^{K} \pi_k N(x|\mu_k, \Sigma_k)$$

~~It is clear~~

It is clear that $p(x) = \sum_z p(z) p(x|z)$.

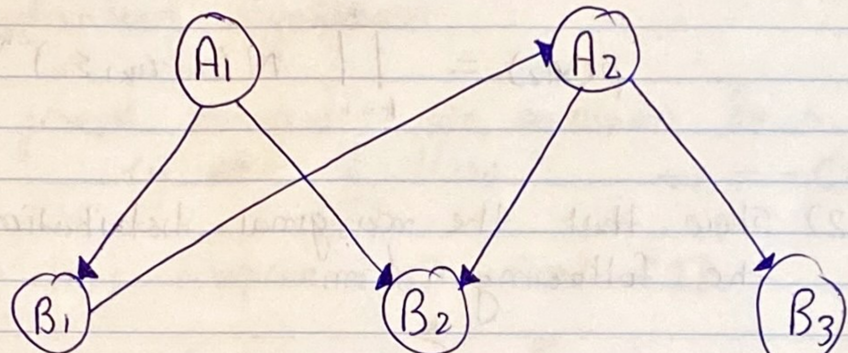
From answers of (1)

$$p(x) = \sum_{k=1}^{K} \pi_k N(x|\mu_k, \Sigma_k)$$

(3) Here it is clear that parameters can not be estimated in closed form. as there is N samples and each sample contains d features denoted by x_i. Expectation Maximization algorithm looks better fit here. It is an iterative way to find MLE of model parameters data has missing values, or hidden variables. EM chooses some random values for missing values and estimates next set of data.

Then recursively these new generated values are used to measure better data by filling up missing points until they are fixed.

Q 4 (1) Draw the corresponding bayesian network.



(2) Joint distribution $P(A_1, A_2, B_1, B_2, B_3)$

$$= P(A_1)\,P(A_2|B_1)\,P(B_1|A_1)\,P(B_2|A_1A_2)\,P(B_3|A_2)$$
$$= P(A_1)\,P(A_2|B_1)\,P(B_1|A_1)\,P(B_3|A_2)\,P(B_2|A_1, A_2)$$

(3) How many independent parameters are required.
$$= 1 + 2 + 2 + 2 + 4$$
$$= 11$$

(4) One such factorization can be
$$= P(A_1)\,P(A_2|A_1)\,P(B_1|A_2A_1)\,P(B_2|B_1A_2A_1)\,P(B_3|B_2B_1A_2A_1)$$
And independent parameters will be
$$= 1 + 2 + 4 + 8 + 16$$
$$= 31$$