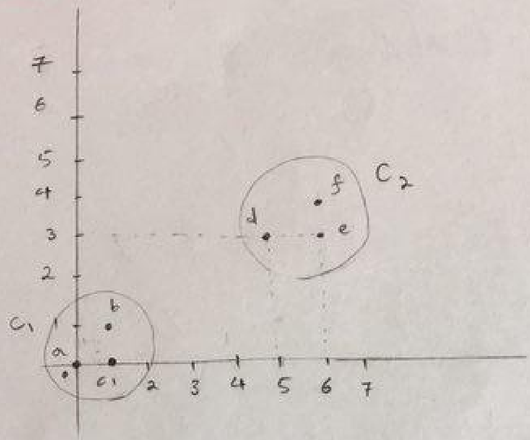
Mahsa
Daneshmand
V00892119
CSC 578D

Q2



points in C_1 : a, b, c where $a=(0,0)$, $b=(1,1)$, $c=(1,0)$

points in C_2 : d, e, f where $d=(5,3)$, $e=(6,3)$, $f=(6,4)$

(I) Distance of clusters C_1, C_2 by method (i): calculating the cluster centroids (mean point of all points in a cluster)

mean point of cluster C_1:

$$\underline{m_1} = \frac{a+b+c}{3} = \frac{(0,0)+(1,1)+(1,0)}{3} = \frac{(2,1)}{3} = (2/3, 1/3)$$

mean point of cluster C_2:

$$\underline{m_2} = \frac{d+e+f}{3} = \frac{(5,3)+(6,3)+(6,4)}{3} = \frac{(17,10)}{3}$$

$$= (17/3, 10/3)$$

$$d(c_1, c_2) = d(m_1, m_2) = \sqrt{(x_2-x_1)^2 + (y_2-y_1)^2}$$

$$= \sqrt{(17/3 - 1/3)^2 + (10/3 - 1/3)^2}$$

$$= \sqrt{37.44} \cong \boxed{6.11} \rightarrow \text{distance by method (i)}$$

(II) Distance of clusters C_1 and C_2 by method(ii):

Calculating the average distance between all points across pairs of clusters.

$$d(C_1, C_2) = \sum_{u \in \{a,b,c\}} \sum_{v(e,f,g)} d(u,v) \Big/ 9$$

$$= \frac{d(a,d) + d(a,e) + d(a,f) + d(b,d) + d(b,e) + d(b,f) + d(c,d) + d(c,e) + d(c,f)}{9}$$

$$= \frac{\sqrt{25+9} + \sqrt{36+9} + \sqrt{64+16} + \sqrt{16+4} + \sqrt{25+4} + \sqrt{25+9} + \sqrt{16+9} + \sqrt{25+9} + \sqrt{25+16}}{9}$$

$$= \sqrt{34} + \sqrt{45} + \sqrt{80} + \sqrt{20} + \sqrt{29} + \sqrt{34} + \sqrt{25} + \sqrt{34} + \sqrt{41} \Big/ 9$$

$= \boxed{6.045} \rightarrow$ distance of C_1, C_2 by method (ii)

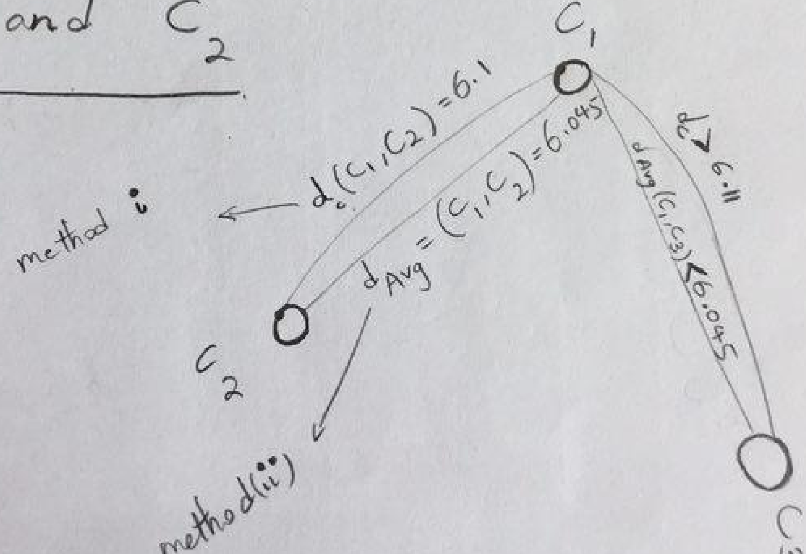
Therefore, these two methods don't give the same distance between clusters!

$(6.11$ and $6.045)$

If a cluster C_3 of point exists such that $dist(C_3, C_2)$ by method (i) be greater than $dist(C_1, C_2) = 6.11$, then by calculating the distance using method (i), clusters C_1, C_2 are selected to be merged.

If the $dist(C_3, C_2)$ by using method (ii) be less than $dist(C_1, C_2)$, then calculating distance by using method (ii) will lead to merging two different clusters: C_3 and C_2.

(Such a cluster (C_3) can be found but is very time consuming)



Q3

Most probable path Q given observations O

$$Q = q_1 \cdots q_t \quad , \quad O = o_1 \cdots o_t$$

recall : $\underset{Q}{\arg\max}\ P(Q|O) = \arg\max\ P(O|Q)P(Q)$

most likely path ending in S_i that emits $o_1 \cdots o_t$:

$$\delta_t(i) = \max_{q_1 \cdots q_{t-1}} P(q_1 \cdots q_{t-1} \wedge q_t = S_i \wedge o_1 \cdots o_t)$$

$$\delta_{t+1}(i) = \underset{q_1 \cdots q_t}{\arg\max} \cdot P(q_1 \cdots q_t \wedge q_{t+1} = S_i \wedge o_1 \cdots o_{t+1})$$

$$= \max \underbrace{\delta_t(j)}_{\substack{\downarrow \\ \text{previous calculation}}} \times \underbrace{P(q_{t+1} = S_i \mid q_t = S_j)}_{\substack{\downarrow \\ \text{transition probability}}} \times \underbrace{P(O_{t+1} | q_{t+1} = S_i)}_{\substack{\downarrow \\ \text{emission prob}}}$$

Once we have $\delta_t(i)$ for all

we can solve $\underset{Q}{\arg\max}\ P(Q|O) = \underset{j}{\arg\max}\ O_t(j)$

①

part a : Consider Sequence 3, 1

I will show that the most probable path for this sequence is HOT, COLD that part a asks. Consider the following table :

State \ t	$O_1 = 3$	$O_2 = 1$
HOT	0.28	0.028
COLD	0.03	0.056

most probable path

max of 2nd col

$$\delta_1(\text{HOT}) = b_{\text{HOT}}(3) \times \underset{\text{start prob}}{\pi_{\text{HOT}}} = 0.4 \times 0.7 = \underbrace{0.28}_{\text{more probable}}$$

$$\delta_1(\text{COLD}) = b_{\text{COLD}}(3) \times \pi_{\text{COLD}} = 0.1 \times 0.3 = 0.03$$

$$\delta_2(\text{HOT}) = \max_j \left[\delta_1(j) \times a_{\text{HOT},j} \times b_{\text{HOT}}(1) \right]$$

$$= \max \left[\delta_1(\text{HOT}) \times a_{\text{HOT, HOT}} \times b_{\text{HOT}}(1), \ \delta_1(\text{COLD}) \times a_{\text{HOT,COLD}} \times b_{\text{HOT}}(1) \right]$$

$$= \max \left[\underbrace{0.28 \times 0.5 \times 0.2}_{= 0.028}, \ \underbrace{0.03 \times 0.3 \times 0.2}_{= 0.0018} \right] = 0.028$$

part (b) Consider Sequence 1, 3

I will show the most probable path for this Sequence starts at COLD & ends in HOT ;

State \ t	$O_1 = 1$	$O_2 = 3$
HOT	0.14	0.028
COLD	0.15	0.009

most probable path : COLD, HOT

$$\delta_1(\text{HOT}) = b_{\text{HOT}}(1) \times \pi_{\text{HOT}} = 0.2 \times 0.7 = 0.14$$

$$\delta_1(\text{COLD}) = b_{\text{COLD}}(1) \times \pi_{\text{COLD}} = 0.5 \times 0.3 = \underline{0.15}$$

more probable

$$\delta_2(\text{HOT}) = \max_j \left[\delta_1(j) \times a_{\text{HOT}, j} \times b_{\text{HOT}}(3) \right]$$

$$= \max \left[0.14 \times 0.5 \times 0.4 ,\ 0.15 \times 0.3 \times 0.4 \right]$$

$$= \max \left[0.028 ,\ 0.018 \right] = 0.028$$

$$\delta_2(\text{COLD}) = \max_j \left[\delta_1(j) \times a_{\text{COLD}, J} , b_{\text{COLD}}(3) \right]$$

$$= \max \left[0.14 \times 0.4 \times 0.1 ,\ 0.15 \times 0.6 \times 0.1 \right] = \max \left[0.0056 , 0.009 \right]$$

$$= 0.009$$