Assignment 9

Clustering

① Given points,

1, 4, 9, 16, 25, 36, 49, 64, 81, 100

There can be 9 different clusters

$\{1\}$, $\{4, 9, \ldots 100\}$

$\{1, 4\}$, $\{9, 16, \ldots 100\}$

⋮

$\{1, 4, \ldots 81\}$, $\{100\}$

Only one point has to be shifted b/w clusters if we change the centroid.

Let the initial values be 36 and 100

mean of 36 & 100 = $\frac{36 + 100}{2} = 68$

So the clusters will be

$\{1, 4, 9, 16, 25, 36, 49, 64\}$, $\{81, 100\}$

Centroids of these clusters are 25.5 and 90.5

mean = $\frac{25.5 + 90.5}{2} = 58$

Now the clusters are

$\{1, 4, 9, 16, 25, 36, 49\}$, $\{64, 81, 100\}$

Only one element is shifted b/w clusters, So true.

(2) Given centroids (0,0) (100, 40)

cluster with centroid (0,0) when L_1 is used

cluster with centroid (100, 40) when L_2 is used

(a) (57, 5)

L_1 norm $\Rightarrow$ (0,0) $= 57 + 5 = 62$

(100, 40) $= 43 + 35 = 78$

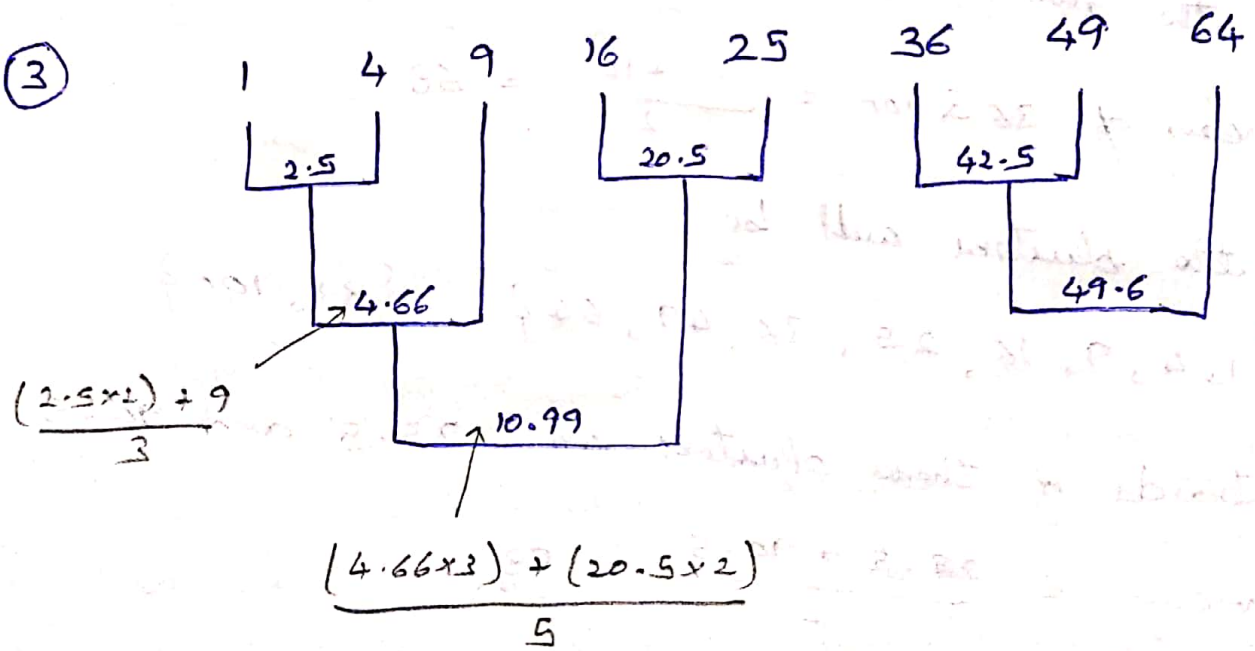
L_2 norm $\Rightarrow$ (0,0) $= \sqrt{57^2 + 5^2} = 57.21$

(100, 40) $= \sqrt{43^2 + 35^2} = 55.44$

∴ when L_1 norm is applied the centroid is (0,0)

L_2 norm is applied the centroid is (100, 40)

(3)



(4) Given there are three clusters,

	A	B	C
Origin	(-10,0)	(0,0)	(10,0)
Points	1000	8000	1000

Given centroids are x, y, z

We can assign each of x, y, z to A, B, C in 27 possible ways

The chance of being in $A = \frac{1000}{10000} = 0.1$

The chance of being in $B = 0.8$

The chance of being in $C = 0.1$

There are five different cases to interchange x, y, z in A, B, C which will total to 27

Finally, we conclude that

A being correct 24%.

C being correct 24%.

A&C together 4.8%.

5 Given

D (33, 33)

E (6, 27)

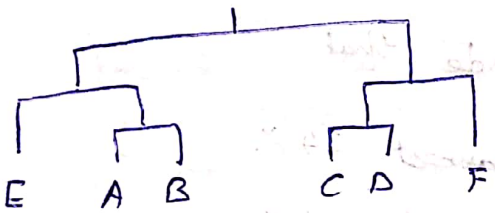
C (21, 21)

B (10, 10)

F (28, 6)

A (0, 0)

	A	B	C	D	E	F
A	0	14.1	29.6	46.6	27.4	28.6
B		0	19.5	32.5	17.7	18.4
C			0	16.9	17	16.5
D				0	28.6	27.4
E					0	31.1
F						0



A & B is done so cluster

AB	C	→	AC (29.6)	BC	19.5
AB	D	→	AD (46.6) ✗	BD	32.5
AB	E	→	AE (27.4) ✓	BE	17.7
AB	F	→	AF (28.6)	BF	18.4

Select the max from each. Then select the min of those maximum

---

CD	(16.9) ✓
DE	28.6
DF	27.4

CD	E	→	CE	17	DE (28.6)
CD	F	→	CF	16.5	DF (27.4) ✓