

\* Suppose that we have four observations, for which we compute a dissimilarity matrix, given by

$$\begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \end{matrix} & \begin{bmatrix} & 0.3 & 0.4 & 0.7 \\ 0.3 & & 0.5 & 0.8 \\ 0.4 & 0.5 & & 0.45 \\ 0.7 & 0.8 & 0.45 & \end{bmatrix} \end{matrix}$$

For instance, the dissimilarity between the second and fourth observations is 0.8.

(a) On the basis of dissimilarity matrix, sketch the dendrogram that results from hierarchically clustering these four observations using complete linkage.

→ ~~with~~ 0.3 is the minimum dissimilarity, so we combine obs 1 and 2 to form cluster (1,2) at height 0.3.

We now have:-

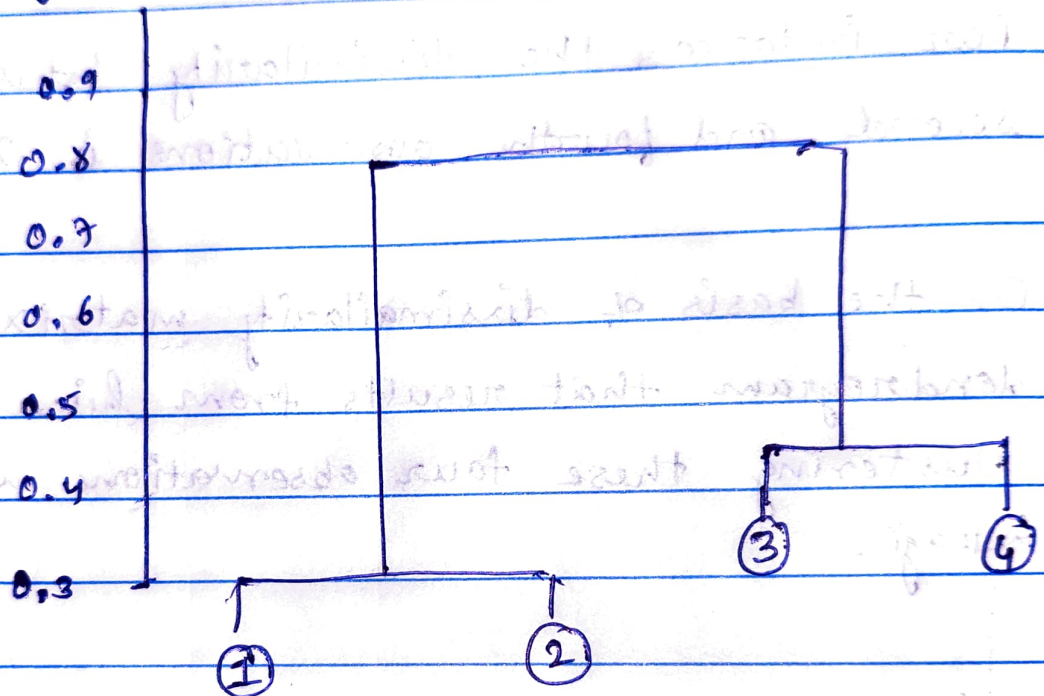
$$\begin{pmatrix} & 0.5 & 0.8 \\ 0.5 & & 0.45 \\ 0.8 & 0.45 & \end{pmatrix}$$

→ minimum dissimilarity 0.45, so we fuse obs 3 and 4 to form a cluster (3,4) at height 0.45

Now we have

$$\begin{pmatrix} & 0.8 \\ 0.8 & \end{pmatrix}$$

→ Now, we fuse clusters (1,2) and (3,4) to form cluster ((1,2), (3,4)) at height 0.8.



⑤ Repeat (a), this time using single linkage clustering.

→ We see 0.3 is minimum dissimilarity, so we fuse obs 1 & 2 at height 0.3

We have now:-  
new dissimilarity matrix

$$\begin{pmatrix} & 0.4 & 0.7 \\ 0.4 & & \\ 0.7 & 0.45 & \end{pmatrix}$$



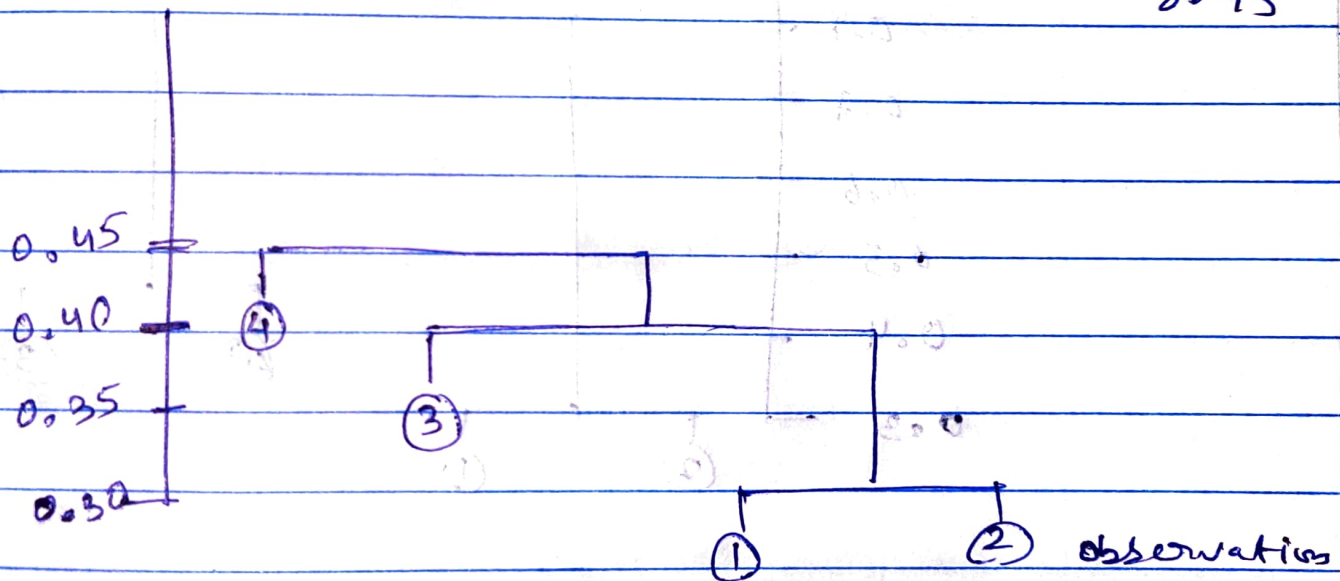
We now see that minimum dissimilarity is 0.4, so we fuse cluster (1, 2) and obs 3 to form  $((1, 2), 3)$  at height 0.4

Now we have :-

$$\begin{pmatrix} & & 0.45 \\ 0.45 & & \end{pmatrix}$$

→ Fuse clusters  $((1, 2), 3)$  and obs 4 to form

Cluster Dendrogram  $((((1, 2), 3), 4))$   
at height 0.45



(c) Suppose that we cut the dendrogram obtained in (a) such that two clusters results. which observations are in each cluster?

→ We will have clusters (1, 2) and (3, 4).

(d) Suppose that we cut the dendrogram obtained in (b) such that two clusters result, which observations are in each cluster?

→ In this case, we have clusters  $(1, 2, 3)$  and  $(4)$ .

(e)

