- $\hat{\Sigma} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 3 \end{bmatrix}$. If one original data is $x = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$. What is the new representation using only the first two principal components?

reconstruction

$\lambda_1 = 5,\ PC_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$

$\lambda_2 = 3,\ PC_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{P} \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & \boxed{3} \end{bmatrix}}_{D} \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{P^{-1}}$$

$\lambda_1 \quad \lambda_2 \quad \lambda_3$

$$\begin{pmatrix} \mathrm{proj}_{PC_1} x \\ \mathrm{proj}_{PC_2} x \end{pmatrix}$$

new representation $= \begin{pmatrix} PC_1^T x \\ PC_2^T x \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$

reconstruction

$$x \approx \underbrace{x_1^{new} PC_1 + x_2^{new} PC_2}\ \cancel{\left(+\, x_3^{new} PC_3\right)} = x$$

$$2\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + 3\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ 3 \end{pmatrix} \approx \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

• [4 points] Perform hierarchical clustering with complete linkage in one-dimensional space on the following points: [0], [−9], [−10], [4], [−4], [7]. Break ties in distances by first combining the instances with the smallest index (appears earliest in the list). Draw the cluster tree.

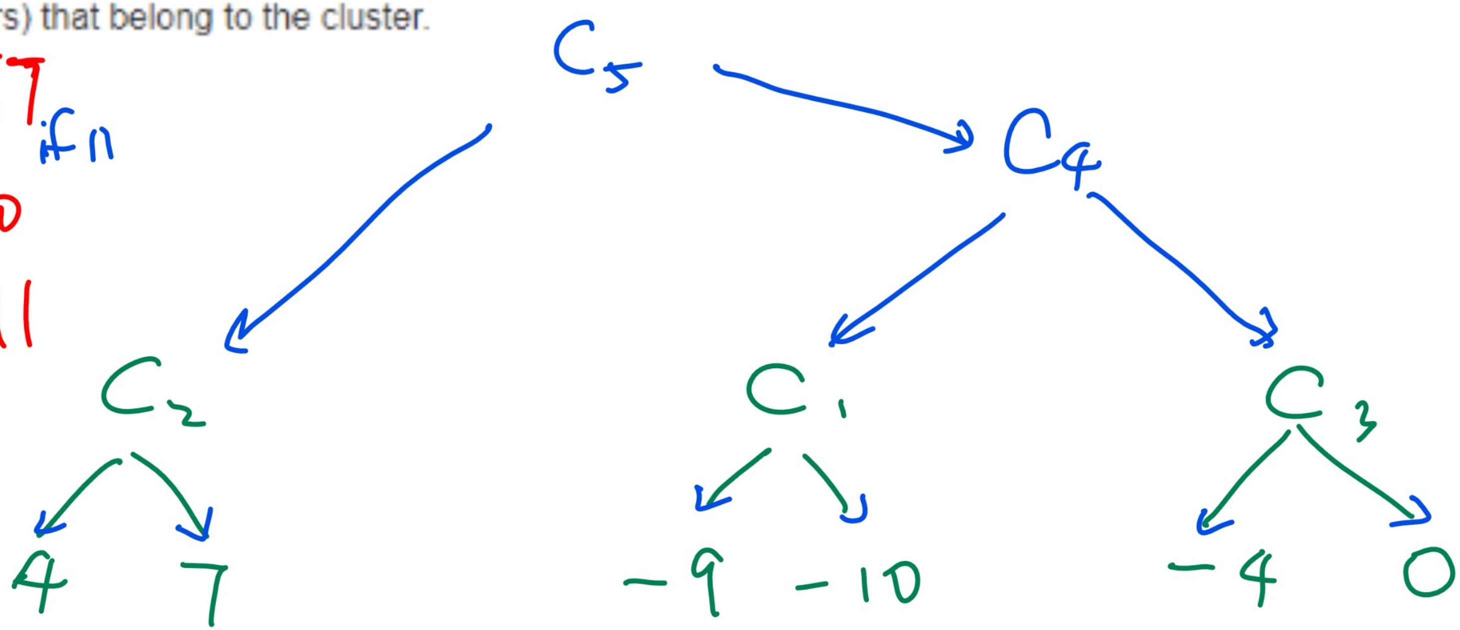
• Note: the node C_1 should be the first cluster formed, C_2 should be the second and so on. All edges to point to the instances (or other clusters) that belong to the cluster.

complete

$d(C_1, C_2) = 17$ if 11

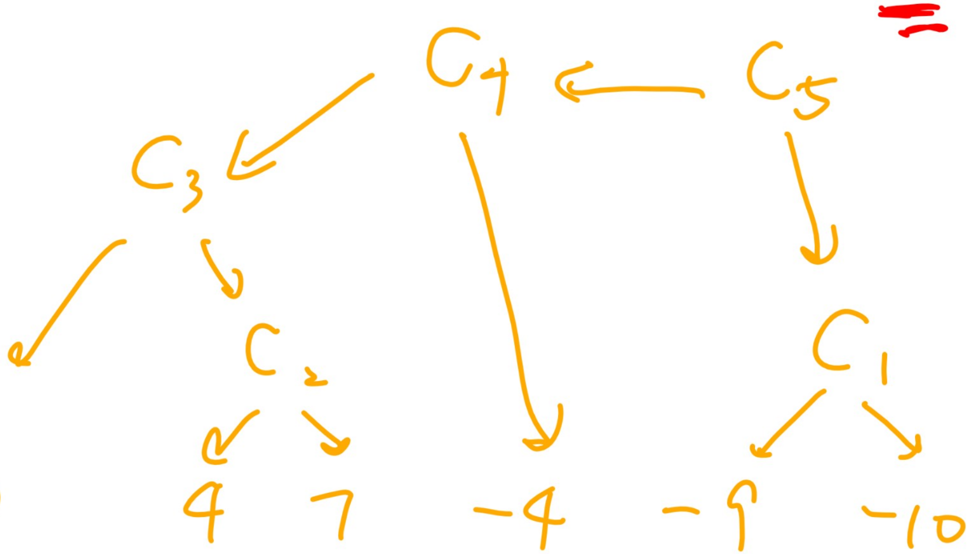
✓ $d(C_1, C_3) = 10$

$d(C_2, C_3) = 11$



0, −9, −10, 4, −4, 7

single



W8 (final) at bottom M8Q6.

☆ $d_{single}(0, (4, 7)) = 4 \rightarrow$ 0 closer to 4

$d_{complete}(0, (4, 7)) = 7 \rightarrow$ 0 closer to -4

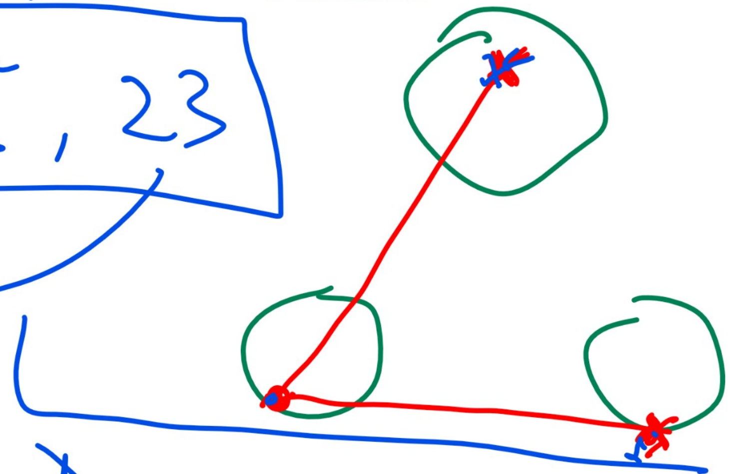
Question 7

• [3 points] Consider the 1D data set: $x_i = i$ for i = 5 to 27. To select good initial centers for k-means where k = 6, let's set $c_1 = 11$. Then select c_j from the unused points in the data set, so that it is farthest from any already-selected centers $c_1, \ldots, c_{j-1}$ (i.e. $c_j = \arg\max_{x_i} \min\{d(c_1, x_i), d(c_2, x_i), \ldots, d(c_{j-1}, x_i)\}$). Enter the initial centers (including c_1) in increasing order (from the smallest to the largest). In case of ties, select the smaller number.

$\boxed{11, 27, 19, 5, 15, 23}$

$5 \rightarrow d(5, (11, 27)) = 6$

$19 \rightarrow d(19, (11, 27)) = 8$



$d(\cancel{26}, (11, 27)) = 1$

M828 in lecture video on exam

Question 9

- [4 points] Suppose K-Means with $K = 2$ is used to cluster the data set $[-4 \quad -3 \quad -2 \quad 8 \quad 10]$ and initial cluster centers are $c_1 = 13$ and $c_2 = x$. What is the smallest value of x if cluster 1 has $n = 2$ points initially (before updating the cluster centers). Break ties by assigning the point to cluster 2.

C_2

$-4 \quad -3 \quad -2 \qquad 8 \quad 10$

$C_1 = 13$

belong to C_2 by tie breaking

$$-2 = \frac{\overset{13}{\cancel{C_1}} + C_2}{2}$$

$C_2 = -17$

$d(-2, \overset{-17}{C_2}) = 15$

$d(-2, \overset{13}{C_1}) = 15$

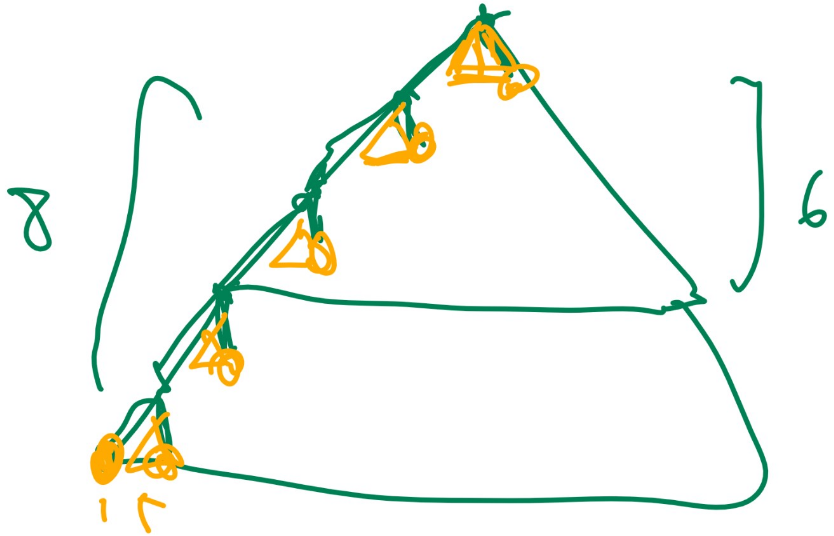
$$\boxed{C_2 \in [-17, 3)}$$

2.9999~

Question 9

• [2 points] Consider a search tree where the root is at depth 0, each internal node has 6 children, and all leaves are at depth 8. There is a single goal state at depth 6. How much stack space (in number of states including the root and the goal) is sufficient so DFS always succeeds? Select all that applies.

BFS IDS



$b = 6$

$d = 6$

$D = 8$

$5 \cdot 8 + 1 = 41$

$(b-1)D + 1$