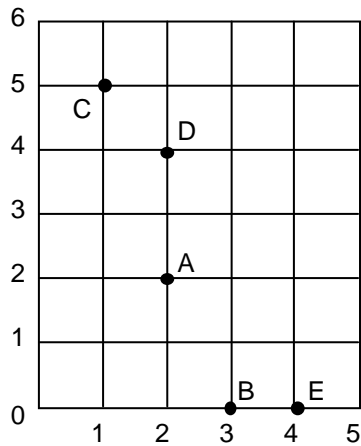


Classification ascendante hiérarchique : exemple numérique simple

Considérons les 5 points suivants dans \mathbf{R}^2 : $A = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$, $B = \begin{bmatrix} 3 \\ 0 \end{bmatrix}$, $C = \begin{bmatrix} 1 \\ 5 \end{bmatrix}$, $D = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$ et $E = \begin{bmatrix} 4 \\ 0 \end{bmatrix}$.

- i) Effectuez une classification ascendante hiérarchique (CAH) en utilisant le carré de la distance euclidienne comme mesure de dissimilarité entre les points, et la stratégie d'agrégation du saut maximum.
- ii) Sans refaire tous les calculs correspondants, indiquez si les regroupements seraient différents avec la stratégie du saut minimum. Justifiez de façon succincte.

i)



1° agrégation $h6 = \{B, E\}$, l'indice d'agrégation est 1, nouvelle matrice :

	A	C	D	h6
A	0	10	4	8
C		0	2	34
D			0	20
h6				0

2° agrégation $h7 = \{C, D\}$, l'indice d'agrégation est 2, nouvelle matrice :

	A	h6	h7
A	0	8	10
h6		0	34
h7			0

Matrice de dissimilarité initiale :

	A	B	C	D	E
A	0	5	10	4	8
B		0	29	17	1
C			0	2	34
D				0	20
E					0

3° agrégation $h8 = \{A, h6\}$, l'indice d'agrégation est 8, nouvelle matrice :

	h7	h8
h7	0	34
h8		0

4° agrégation $h9 = \{h7, h8\}$, l'indice d'agrégation est 34.

ii) Avec la stratégie du saut minimum, A serait regroupé avec h7 au lieu d'être regroupé avec h6. La distance entre A et C est supérieure à celle entre A et E, mais la distance entre A et D est inférieure à celle entre A et B.