Faculté des Sciences exactes
Département de Maths
2 ème Année STID

(Novembre2020)

Exercice 1

Soit le tableau de données correspondant aux réponses (0=non et 1=oui) de ' individus à 4 questions A, B, C et D

	A	B	C	D
i1	0	1	1	1
i2	0	0	1	1
i3	1	0	0	1
i4	1	1	0	1

1. Donner le tableau des dissimilarités en utilisant l'indice Jaccard.
2. Construire une partition en utilisant la méthode de liaison moyenne avec un seuil de 0,4
3. Décrire l'algorithme d"une hierarchie par regroupement progressif en utilisant l'indice d'agrégation du lien minimum.
4. Construire la hierarchie et la représentée.
5. Extraire une partition à 2 classes.

Exercice 2

On considère les 6 individus suivants définis par leurs coordonnées respectifs dans un espace de dimension 2.

E1 = (1,2)
E2 = (2,2)
E3 = (2,4)
E4 = (3,3)
E5 = (3,4)
E6 = (4,4)

1. Réaliser une classification hiérarchique par passage à l'ultramétrique inférieur maximale
de ces individus, en utilisant le tableau des distances L1.
2. Représenter par un dendrogramme l'hiérarchie obtenue.

corrigé de l'EMD : $2^{\text{ème}}$ Année STID ANAD.

Exercice 1

1) Tableau des dissimilarité de Jaccard : $\frac{a}{a+b+c}$

$a_{12} = 2$ $b_{12} = 1$ $c_{12} = 0$ $s(1,2) = s(2,1) = 2/3$

$a_{13} = 1$ $b_{13} = 2$ $c_{12} = 1$ $s(1,3) = s(3,1) = 1/4$

$a_{14} = 2$ $b_{14} = 1$ $c_{14} = 1$ $s(1,4) = s(4,1) = 2/4$

$a_{23} = 1$ $b_{23} = 1$ $c_{23} = 1$ $s(2,3) = s(3,2) = 1/3$

$a_{24} = 1$ $b_{24} = 1$ $c_{24} = 2$ $s(2,4) = s(4,2) = 1/4$

$a_{34} = 2$ $b_{34} = 0$ $c_{34} = 1$ $s(3,4) = s(4,3) = 2/3$

similarités

$d = (1 - s)$

1,5

Tableau des dissimilarités

d	i_1	i_2	i_3	i_4
i_1	0	1/3	3/4	1/2
i_2	1/3	0	2/3	3/4
i_3	3/4	2/3	0	1/3
i_4	1/2	3/4	1/3	0

0,5

2) Partition avec la méthode de laison moyenne $\delta = 0,4$

La plus petite dissimilarité est $1/3 < \delta$.

$d(i_1, i_2) = 1/3$ on forme la partie $\{i_1, i_2\}$

0,25

	$\{i_1, i_2\}$	i_3	i_4
$\{i_1, i_2\}$	0	$\frac{17}{24}$	5/8
$\{i_3\}$		0	1/3
i_4			0

0,5

0,25

La plus petite dissimilarité est $\delta = 1/3 < 0,4$

$\delta(i_3, i_4) = 1/3$, on forme la partie $\{i_3, i_4\}$

	$\{i_1, i_2\}$	$\{i_3, i_4\}$
$\{i_1, i_2\}$	0	$\frac{32}{48} = \frac{2}{3}$
$\{i_3, i_4\}$	$\frac{32}{48} = \frac{2}{3}$	0

0,5

Toutes les dissimilarités sont $> \delta$, on arrête le regroupement 0,5

$P = \{\{i_1, i_2\}; \{i_3, i_4\}\}$ 1

3/ Algorithme de construction d'une hiérarchie par regroupement progressif : indice d'agrégation du lien max.

$$d(A, B) = \max_{\substack{x \in A \\ y \in B}} d(x, y)$$

La plus petite dissimilarité est $\delta = \frac{1}{3} = d(\{i_1 ; i_2) ; \{i_1, i_2\}$ 0,5

	$\{i_1, i_2\}$	i_3	i_4
$\{i_1, i_2\}$	0	$1/2$	$1/2$
i_3		0	$1/3$
i_4			0

1

La + petite dissimilarité est $d(i_3, i_4) = 1/3$

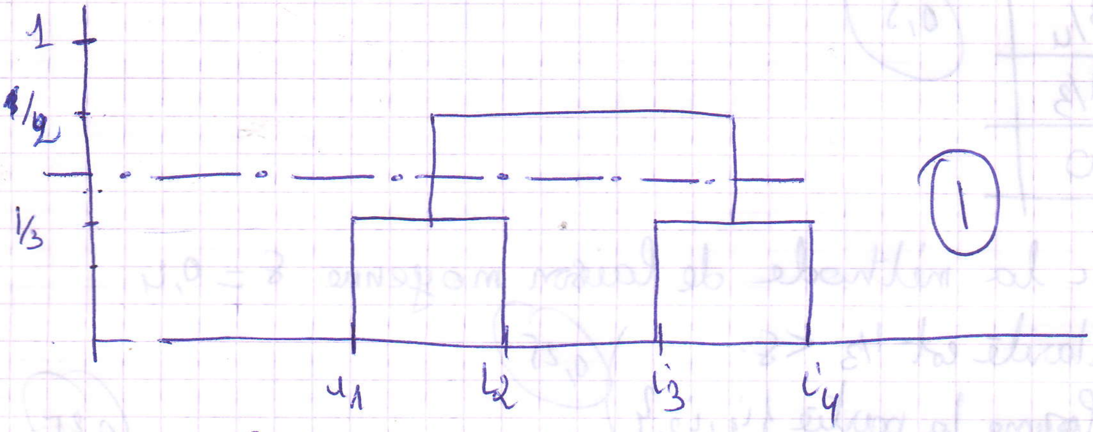
On forme la partie $\{i_3, i_4\}$

d	$\{i_1, i_2\}$	$\{i_3, i_4\}$
$\{i_1, i_2\}$	0	$\frac{1}{2}$
$\{i_3, i_4\}$	$\frac{1}{2}$	0

0,5

On regroupe $\{i_1, i_2, i_3, i_4\}$

$H = \{ \{i_1\} ; \{i_2\} \{i_3\} ; \{i_4\} ; \{i_1, i_2\} \{i_3, i_4\} \{i_1, i_2, i_3, i_4\} \}$. 1



1

Algorithme

(1) Repérer la plus petite valeur du tableau des dissimilarité $d(w_1, w_2)$,

(2)- Former la partie $\{w_1, w_2\}$

(3)- Calculer les dissimilarités entre $A = \{w_1, w_2\}$ et tous les autres inds en utilisant l'indice d'agrégation du lien maximum

$$d(A, w_i) = \max_{x \in A} d(x, w_i)$$

(4)- Reprendre l'étape (1) jusqu'à regroupement complet.

2

5) Partition à deux classes, on coupe l'arbre par une droite horizontale passant entre 1/3 et 1/2

La partition est $P = \{\{i_1, i_2\} ; \{i_3, i_4\}\}$ 1

Exercice 2

1) Hierarchie par passage à l'ultra-métrique inferieure max

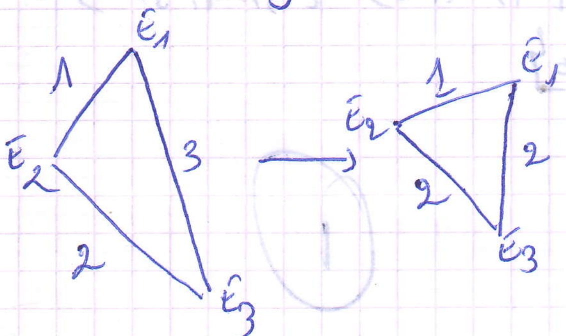
Tableau des distances L1. 2

d_{L1}	E_1	E_2	E_3	E_4	E_5	E_6
E_1	0	1	$\cancel{3}$ 2	$\cancel{3}$ 2	$\cancel{4}$ 2	$\cancel{5}$ 2
E_2		0	2	2	$\cancel{3}$ 2	$\cancel{4}$ 2
E_3			0	$\cancel{2}$ 1	1	$\cancel{2}$ 1
E_4				0	1	$\cancel{2}$ 1
E_5					0	1
E_6						0

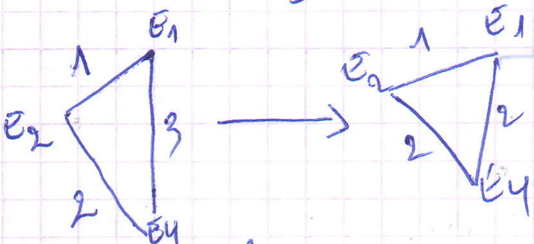
La méthode consiste à rendre tous les triangles isoceles avec une base inferieure au côtes 1

On a $C_6^3 = 20$ triangles.

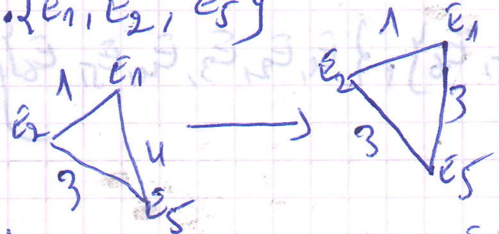
- $\{E_1, E_2, E_3\}$



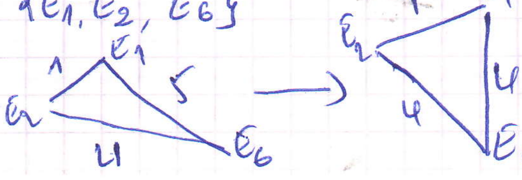
- $\{E_1, E_2, E_4\}$



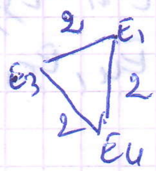
- $\{E_1, E_2, E_5\}$



- $\{E_1, E_2, E_6\}$

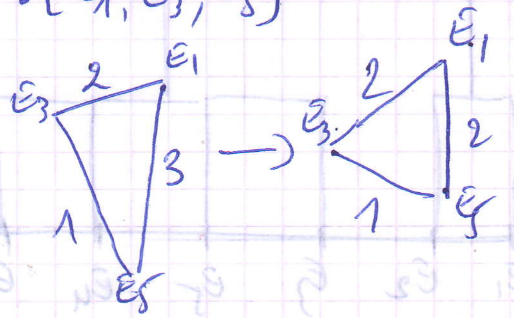


- $\{E_1, E_3, E_4\}$

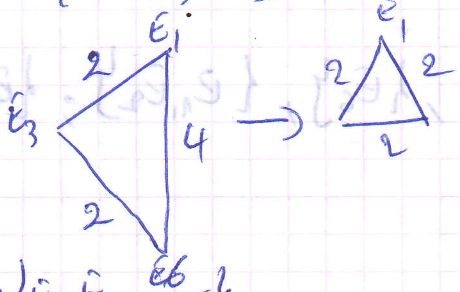


→ pas de changement

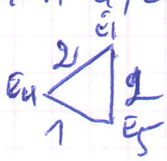
- $\{E_1, E_3, E_5\}$



- $\{E_1, E_3, E_6\}$

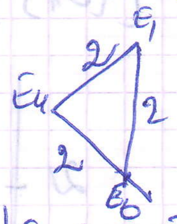


- $\{E_1, E_4, E_5\}$



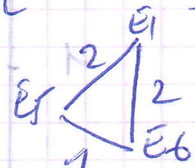
→ Pas de changement

- $\{E_1, E_4, E_6\}$



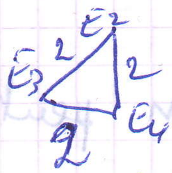
Pas de changem

- $\{E_1, E_5, E_6\}$



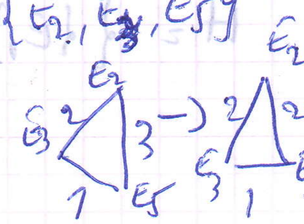
Pas cha

- $\{E_2, E_3, E_4\}$

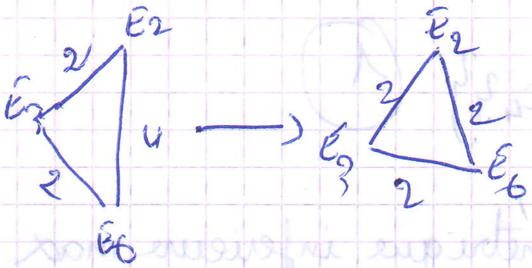


Pas d cha

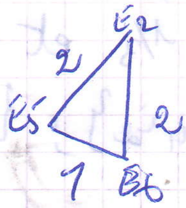
- $\{E_2, E_3, E_5\}$



• $\{E_2, E_3, E_6\}$

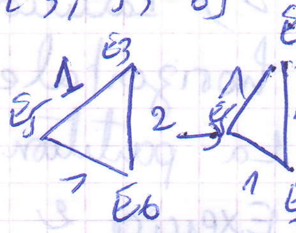


$\{E_2, E_5, E_6\}$

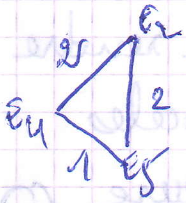


Pas chang

• $\{E_3, E_5, E_6\}$

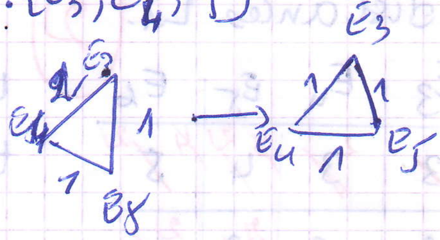


• $\{E_2, E_4, E_5\}$

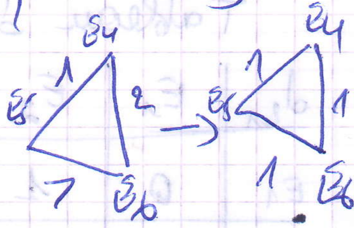


pas de changement

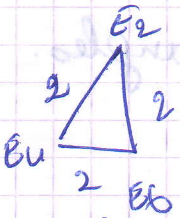
• $\{E_3, E_4, E_5\}$



$\{E_4, E_5, E_6\}$

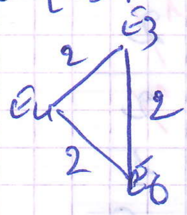


• $\{E_2, E_4, E_6\}$



pas de changement

• $\{E_3, E_4, E_6\}$



Pas change

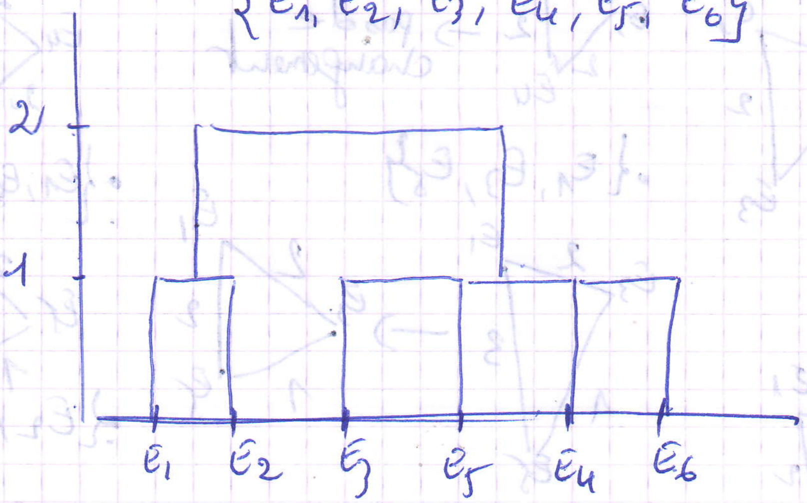
3

1

La hierarchie

$H = \{\{E_1\}; \{E_2\}; \ldots, \{E_6\}; \{E_1, E_2\}, \{E_3, E_5\}, \{E_3, E_4, E_5\}\ \{E_3, E_4, E_5, E_6\}$

$\{E_1, E_2, E_3, E_4, E_5, E_6\}$



1

En peut aussi prendre

$H = \{\{E_1\}, \ldots, \{E_6\}, \{E_1, E_2\}; \{E_3, E_4, E_5, E_6\}; \{E_1, E_2, E_3, E_4, E_5, E_6\}\}$