

Faculté des Sciences exactes
Département de Maths
2 ème Année STID

(2020)

Exercice 1

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. Pour les différentes distances vues en cours calculer les distances entes les 6 individus E1 à E6
2. Pour chaque cas , construire une partition rn utilisant les différentes méthodes vues en cours.
3. Discuter

Exercice 2

Soit le tableau de données correspondant aux réponses (0=non et 1=oui) des 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 similarités en utilisantles différents indices donnés dans la cours. En déduire le tableau des dissimilarités.
2. Pour chaque cas , construire une partition rn utilisant les différentes méthodes vues en cours.
3. Discuter

Exercice 3

Soi le tableau des données quantitatives décrivant 8 eaux minérales sur 5 variables.

	saveur.amère	saveur.sucrée
St Yorre (S ^t Y)	3.4	3.1
Badoit (B)	3.8	2.6
Vichy (V)	2.9	2.9
Quézac (Q)	3.9	2.6
Arvie (A)	3.1	3.2
Chateau Neuf (C.N)	3.7	2.8
Salvetat (S)	4.0	2.8
Perrier (P)	4.4	2.2

1/Donner le tableau des distances L_1 .

2/ Appliquer l'algorithme de classification hiérarchique ascendante par regroupement progressif en utilisant un indice d'agrégation du lien maximale aux données précédentes puis du lien minimal, en précisant à chaque étape le nouveau tableau de distances.

3/ Dessiner le dendrogramme associé.

4/ Extraire une partition à deux classes et une partition à trois classes.

5/ Calculer l'inertie inter classes dans les cas des deux partitions. Conclure.

Exercice 4

On dispose d'un tableau de données avec 7 individus repérés par 3 variables.

Tab.1 : tableau des données

	V1	V2	V3
I1	1	2	2
I2	2	0	1
I3	2	2	1
I4	8	7	0
I5	7	8	1
I6	6	9	0
I7	2	0	2

1/ Donner la matrice des variances covariances.

2/ Considérons la partition $\{\{I1, I2, I3, I7\}, \{I4, I5, I6\}\}$

3/ Donner les matrices des variances-covariances intra et inter classes.

Soi le tableau des données quantitatives décrivant 8 eaux minérales sur 5 variables.

	saveur.amère	saveur.sucrée	saveur.acide	saveur.salée	saveur.alcaline
St Yorre	3.4	3.1	2.9	6.4	4.8
Badoit	3.8	2.6	2.7	4.7	4.5
Vichy	2.9	2.9	2.1	6.0	5.0
Quézac	3.9	2.6	3.8	4.7	4.5
Arvie	3.1	3.2	3.0	5.2	5.0
Chateaufeuf	3.7	2.8	3.0	5.2	4.6
Salvetat	4.0	2.8	3.0	4.1	4.5
Perrier	4.4	2.2	4.0	4.9	3.9

1/ Donner le tableau des distances L_1 .

Exercice 5

On considère un tableau de données croisant 5 individus repérés par 3 variables.

	X	Y	Z
a	2	3	5
b	3	5	3
c	2	4	8
i	3	2	4
j	5	6	5

1. Calculer la matrice Σ de variance-covariance.

2. Si on considère les deux classes $\{a, b, c\}$ et $\{i, j\}$ donner la décomposition de la variance totale.
3. Construire une partition en utilisant la méthode de liaison moyenne pour $\alpha = 5.5$.

Exercice 3

On considère les 4 points de \mathbb{R}^2 $M_1 = (-2, 3)$ $M_2 = (-2, 1)$ $M_3 = (-2, 6 - 1)$ et $M_4 = (2, -1)$

1. Par passage à l'ultramétrie supérieur, construire une hiérarchie indicée et la représentée.

Exercice 4

Considérons un ensemble d'individus divisé en deux classe A et B , l'effectif de A est deux fois inférieur à l'effectif de B . Chacun des individu est décrit par deux variables numériques. Les centres de gravité des classes sont

$$G_A = \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \quad G_B = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$$

et les matrices des variances Intraclasse sont :

$$S_A = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}, \quad S_B = \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix}$$

- 1/ Déterminer la matrice des variances totale.
- 2/ Déterminer l'équation de la frontière de discrimination entre les classes suivant l'approche de Sebestyen suivante :

$$[\det(S_A)]^{\frac{1}{2}} (x - G_A) S_A^{-1} (x - G_A) = [\det(S_B)]^{\frac{1}{2}} (x - G_B) S_B^{-1} (x - G_B)$$

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$