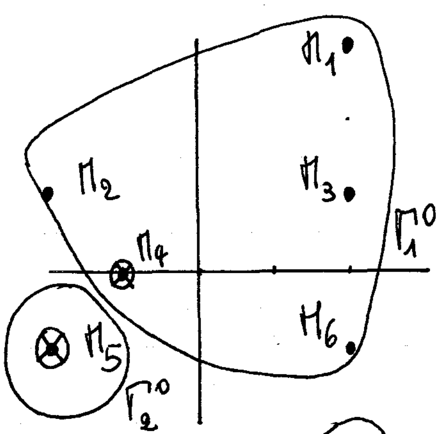
NOM : Corrigé
PRENOM :

Date :
Groupe :

Mathématiques pour la Biologie (semestre 2) : Feuille-réponses du TD 8
Classification par la méthode des centres mobiles

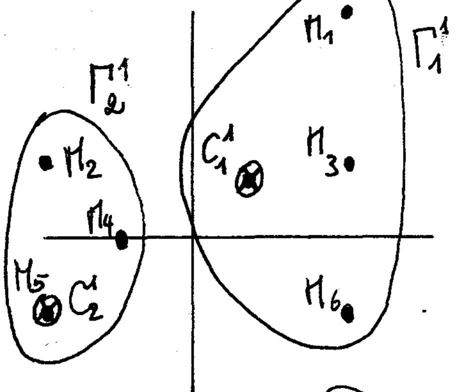
Exercice 1 : On considère les 6 points $M_1 = (2,3)$, $M_2 = (-2,1)$, $M_3 = (2,1)$, $M_4 = (-1,0)$, $M_5 = (-2,-1)$ et $M_6 = (2,-1)$. En supposant que les deux points M_4 et M_5 sont les centres initiaux, décrire par une succession de dessins, les étapes de l'algorithme des centres mobiles en représentant à chaque itération de l'algorithme les centres ainsi que les classes qu'on entourera chacune d'un rond.



Première itération :

$C_1^0 = M_4 \qquad C_2^0 = M_5$

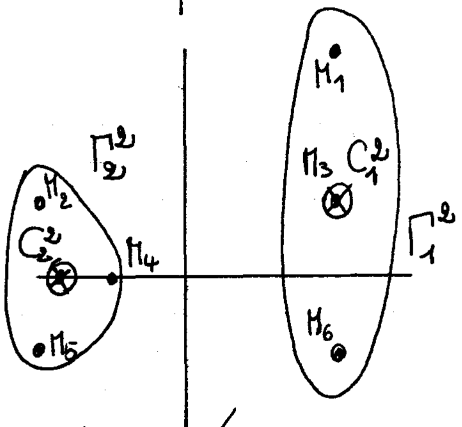
$\Gamma_1^0 = \{M_1, M_2, M_3, M_4, M_6\} \qquad \Gamma_2^0 = \{M_5\}$



Deuxième itération :

$$C_1^1 = \begin{pmatrix} \frac{2-2+2-1+2}{5} \\ \frac{3+1+1+0-1}{5} \end{pmatrix} = \begin{pmatrix} 0,6 \\ 0,8 \end{pmatrix} \qquad C_2^1 = M_5$$

$\Gamma_1^1 = \{M_1, M_3, M_6\} \qquad \Gamma_2^1 = \{M_2, M_4, M_5\}$



Troisième itération :

$$C_1^2 = \begin{pmatrix} \frac{2+2+2}{3} \\ \frac{3+1-1}{3} \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} = M_3 \qquad C_2^2 = \begin{pmatrix} \frac{-2-1-2}{3} \\ \frac{1+0-1}{3} \end{pmatrix} = \begin{pmatrix} -1,66 \\ 0 \end{pmatrix}$$

$\Gamma_1^2 = \{M_1, M_3, M_6\} \qquad \Gamma_2^2 = \{M_2, M_4, M_5\}$

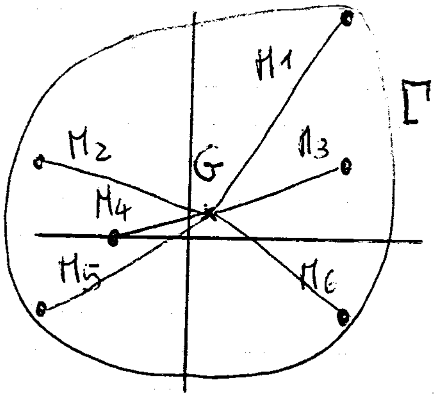
La procédure est stoppée puisque la partition est restée inchangée entre la deuxième et la troisième itération.

Exercice 2 : Les trois dessins de la page précédente représentent trois partitions différentes du même ensemble. Calculer l'inertie totale du nuage puis, pour chacune des partitions, l'inertie intra classe et vérifier qu'elle est bien décroissante au cours du processus. En calculant l'inertie inter de l'une des partition, vérifier sur l'exemple le théorème de Huygens.

- Le centre de gravité du nuage Γ est $G = \begin{pmatrix} \frac{2-2+2-1-2+2}{6} \\ \frac{3+1+1+0-1-1}{6} \end{pmatrix} = \begin{pmatrix} 1/6 \\ 1/2 \end{pmatrix}$

L'inertie totale est donc :

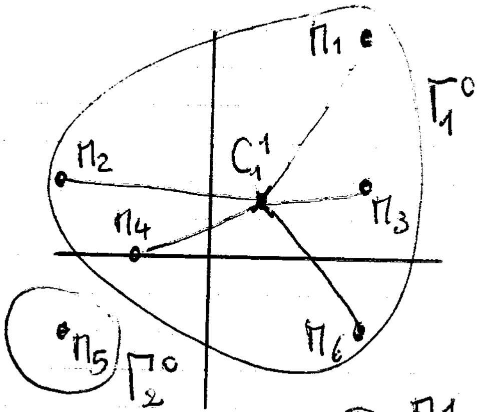
$$J(\Gamma) = \frac{1}{6}\left[d(M_1,G)^2 + d(M_2,G)^2 + d(M_3,G)^2 + d(M_4,G)^2 + d(M_5,G)^2 + d(M_6,G)^2\right]$$
$$= \frac{1}{6}\left[9{,}61 + 4{,}94 + 3{,}61 + 1{,}61 + 6{,}94 + 5{,}61\right] = \underline{5{,}39}$$



- Pour la partition de Γ en Γ_1^0 et Γ_2^0, on a :

$$J_{intra} = J(\Gamma_1^0) + J(\Gamma_2^0)$$
$$= \frac{1}{6}\left[d(M_1,C_1^1)^2 + d(M_2,C_1^1)^2 + d(M_3,C_1^1)^2 + d(M_4,C_1^1)^2 + d(M_5,C_1^1)^2\right] + 0$$
$$= \frac{1}{6}\left[6{,}8 + 6{,}8 + 2 + 3{,}2 + 5{,}2\right] = \underline{\underline{4}}$$

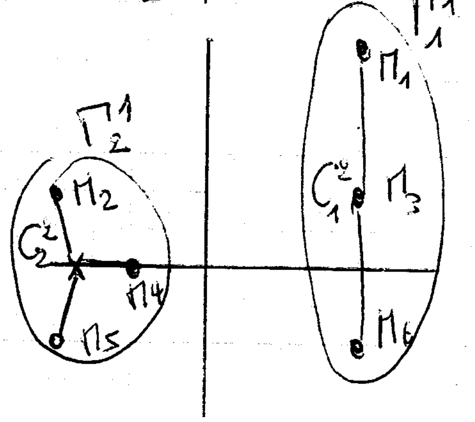
$$J_{inter} = \frac{5}{6} d(C_1^1, G)^2 + \frac{1}{6} d(M_5, G)^2 = \underline{\underline{1{,}39}}$$



- Pour la partition de Γ en Γ_1^1 et Γ_2^1, on a :

$$J_{intra} = J(\Gamma_1^1) + J(\Gamma_2^1)$$
$$= \frac{1}{6}\left[d(M_1,M_3)^2 + 0 + d(M_6,M_3)^2\right] + \frac{1}{6}\left[d(M_2,C_2^2)^2 + d(M_4,C_2^2)^2 + d(M_5,C_2^2)^2\right]$$
$$= \frac{1}{6}(4+4) + \frac{1}{6}\left[\frac{10}{9} + \frac{4}{9} + \frac{10}{9}\right] \simeq \underline{\underline{1{,}78}}$$

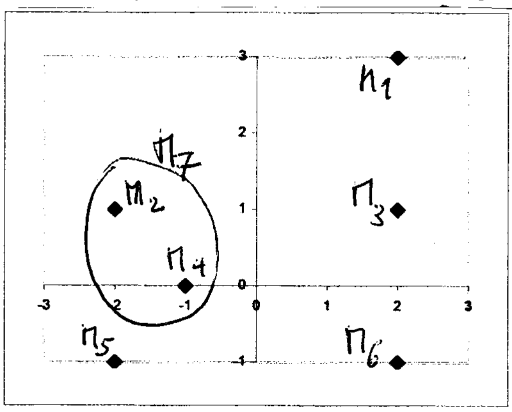
$$J_{inter} = \frac{3}{6} d(C_1^2, G)^2 + \frac{3}{6} d(C_2^2, G)^2$$
$$= \frac{1}{2}(3{,}61) + \frac{1}{2}(3{,}61) = \underline{\underline{3{,}61}}$$



On observe qu'à chaque itération, l'inertie intra diminue (elle passe de 5,39 à 4 puis 1,78) et l'inertie inter augmente (elle passe de 0 à 1,39 puis 3,61). Mais la somme des deux reste toujours égale à l'inertie totale ($4+1{,}39 = 5{,}39$ $\quad 1{,}78+3{,}61 = 5{,}39$).

Exercice 3 : Classifier les points du nuage précédent par une classification hiérarchique ascendante et représenter le dendrogramme (à noter que lorsqu'on doit regrouper les deux points les plus proches et qu'il existe deux couples de points satisfaisant cette condition, on convient de choisir les deux points dont les numéros sont les plus petits).

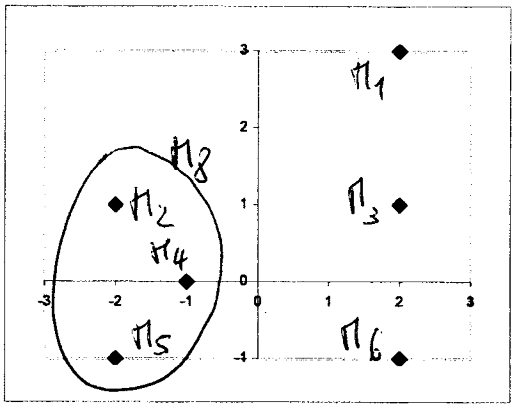
On calcule la "distance" entre les points en prenant le carré de la distance euclidienne et la distance "au plus proche voisins" entre les classes.



	M1	M2	M3	M4	M5	M6
M1	0	20	4	18	32	16
M2		0	16	2	4	20
M3			0	10	20	4
M4				0	2	10
M5					0	16
M6						0

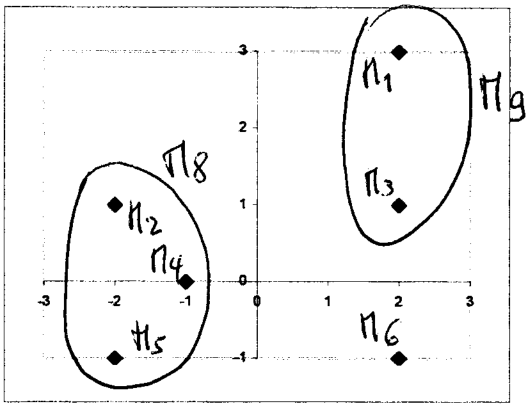
plus petite distance

On pose $M_7 = \{M_2, M_4\}$



	M1	M7	M3	M5	M6
M1	0	18	4	32	16
M7		0	10	2	10
M3			0	20	4
M5				0	16
M6					0

On pose $M_8 = \{M_7, M_5\}$ plus petite distance



	M1	M8	M3	M6
M1	0	18	4	16
M8		0	10	10
M3			0	4
M6				0

On pose $M_9 = \{M_1, M_3\}$ plus petite distance

	M9	M8	M6
M9	0	10	4
M8		0	10
M6			0

plus petite distance

On pose $M_{10} = \{M_8, M_6\}$

	M9	M10
M9	0	10
M10		0