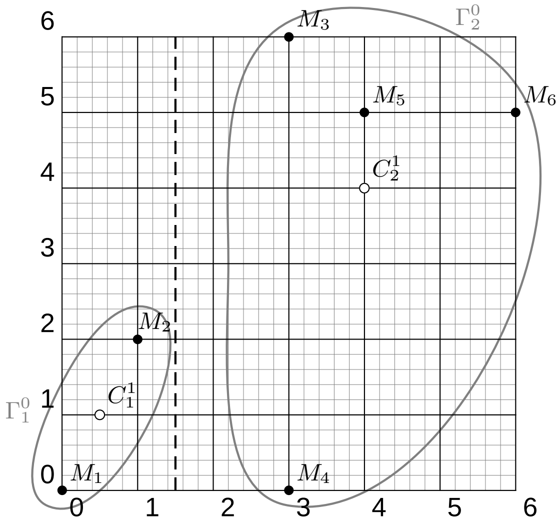
Université de Nice — L1SV, année 2016-2017
Mathématiques pour la Biologie (semestre 2)

CORRIGÉ

Mathématiques pour la Biologie : Feuille-réponses du TD 8
Classification automatique de données par la méthode des centres mobiles.

Exercice 1 : On considère les 6 points $M_1(0,0)$, $M_2(1,2)$, $M_3(3,6)$, $M_4(3,0)$, $M_5(4,5)$ et $M_6(6,5)$. En supposant que les deux points M_1 et M_4 sont les centres initiaux, décrire par une succession de dessins *commentés*, les étapes de l'algorithme des centres mobiles en représentant à chaque itération

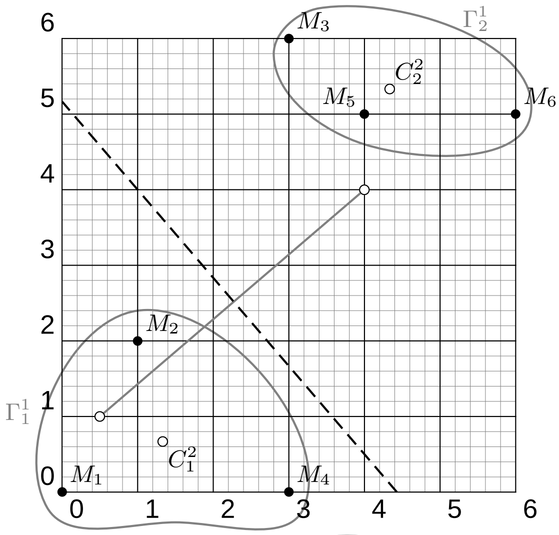
- — les classes *dont on donnera les éléments* et qu'on entourera chacune d'un arrondi,
- — ainsi que les centres de ces classes *qu'on calculera*.



1. Première étape
On trace la médiatrice du segment M_1M_4.
Les points sont alors séparés en deux classes :

la classe $\Gamma_1^0 = \{M_1, M_2\}$
dont le centre $C_1^1 = \left(\frac{0+1}{2}, \frac{0+2}{2}\right) = \left(\frac{1}{2}, 1\right)$

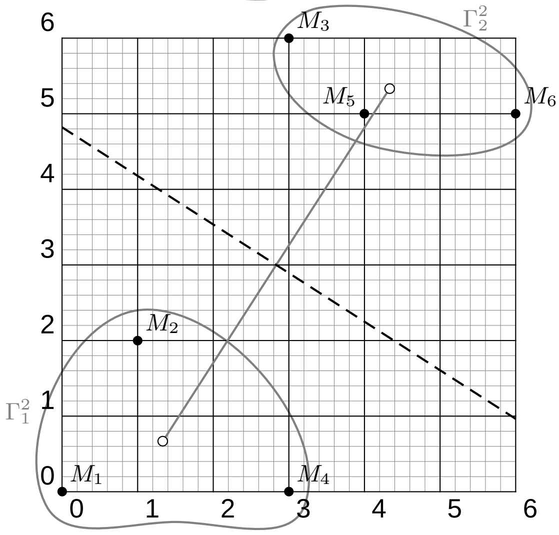
et la classe $\Gamma_2^0 = \{M_3, M_4, M_5, M_6\}$
dont le centre $C_2^1 = \left(\frac{3+3+4+6}{4}, \frac{6+0+5+5}{4}\right) = (4,4)$.



2. Deuxième étape
On trace la médiatrice du segment $C_1^1C_2^1$.
Les points sont alors séparés en deux classes :

la classe $\Gamma_1^1 = \{M_1, M_2, M_4\}$
dont le centre $C_1^2 = \left(\frac{0+1+3}{3}, \frac{0+2+0}{3}\right) = \left(\frac{4}{3}, \frac{2}{3}\right)$

et la classe $\Gamma_2^1 = \{M_3, M_5, M_6\}$
dont le centre $C_2^2 = \left(\frac{3+4+6}{3}, \frac{6+5+5}{3}\right) = \left(\frac{13}{3}, \frac{16}{3}\right)$.



3. Troisième étape
On trace la médiatrice du segment $C_1^2C_2^2$.
On constate alors que

la classe $\Gamma_1^2 = \{M_1, M_2, M_4\} = \Gamma_1^1$
donc le centre $C_1^3 = C_1^2$

et la classe $\Gamma_2^2 = \{M_3, M_5, M_6\} = \Gamma_2^1$
donc le centre $C_2^3 = C_2^2$.

L'algorithme est terminé.

Exercice 2 : Les dessins de la page précédente représentent des partitions différentes du même ensemble. Calculer l'inertie totale du nuage, puis pour chacune des partitions, l'inertie intraclasse et vérifier qu'elle est bien décroissante au cours du processus de classification.

0. Inertie totale du nuage :

On commence par calculer le centre de gravité G du nuage Γ :
$G = \left(\frac{0+1+3+3+4+6}{6}, \frac{0+2+6+0+5+5}{6}\right) = \left(\frac{17}{6}, 3\right) \approx (2{,}83\ , 3)$.
Puis l'inertie totale du nuage Γ :
$\mathcal{I}(\Gamma) = \frac{1}{6}\left(d_2(M_1,G)^2 + d_2(M_2,G)^2 + \cdots + d_2(M_6,G)^2\right) = \frac{1}{6}\frac{(613+157+325+325+193+505)}{36} = \frac{353}{36} \approx 9{,}806$

1. Première partition :

On commence par calculer l'inertie de chaque classe :
$\mathcal{I}(\Gamma_1^0) = \frac{1}{2}\left(d_2(M_1,C_1^1)^2 + d_2(M_2,C_1^1)^2\right) = \frac{1}{2}\left(\frac{5}{4} + \frac{5}{4}\right) = \frac{5}{4}$,
$\mathcal{I}(\Gamma_2^0) = \frac{1}{4}\left(d_2(M_3,C_2^1)^2 + d_2(M_4,C_2^1)^2 + d_2(M_5,C_2^1)^2 + d_2(M_6,C_2^1)^2\right) = \frac{5+17+1+5}{4} = 7$,
et pour finir l'inertie intraclasse $\mathcal{I}_{intra1} = \frac{2}{6}\frac{5}{4} + \frac{4}{6}7 = \frac{61}{12} \approx 5{,}083$.

2. Deuxième partition :

Comme précédemment on calcule l'inertie de chaque classe :
$\mathcal{I}(\Gamma_1^1) = \frac{1}{3}\left(d_2(M_1,C_1^2)^2 + d_2(M_2,C_1^2)^2 + d_2(M_4,C_1^2)^2\right) = \frac{1}{3}\frac{(20+17+29)}{9} = \frac{22}{9} \approx 2{,}444$,
$\mathcal{I}(\Gamma_2^1) = \frac{1}{3}\left(d_2(M_3,C_2^2)^2 + d_2(M_5,C_2^2)^2 + d_2(M_6,C_2^2)^2\right) = \frac{1}{3}\frac{(20+2+26)}{9} = \frac{16}{9} \approx 1{,}778$,
et pour finir l'inertie intraclasse $\mathcal{I}_{intra2} = \frac{3}{6}\frac{22}{9} + \frac{3}{6}\frac{16}{9} = \frac{19}{9} \approx 2{,}111$.
On constate que l'inertie intraclasse a nettement baissé.

3. Troisième partition :

Les classes et leurs centres ne changeant pas, on a donc :
$\mathcal{I}(\Gamma_1^2) = \mathcal{I}(\Gamma_1^1) = \frac{22}{9} \approx 2{,}444$ et $\mathcal{I}(\Gamma_2^2) = \mathcal{I}(\Gamma_2^1) = \frac{16}{9} \approx 1{,}778$,
et donc l'inertie intraclasse est la même $\mathcal{I}_{intra3} = \mathcal{I}_{intra2} = \frac{3}{6}\frac{22}{9} + \frac{3}{6}\frac{16}{9} = \frac{19}{9} \approx 2{,}111$.

4. En calculant l'inertie interclasse de la deuxième partition, vérifier sur cet exemple le théorème de Huygens.

On calcule l'inertie interclasse de la deuxième partition :
$\mathcal{I}_{inter2} = \frac{3}{6}d_2(C_1^2,G)^2 + \frac{3}{6}d_2(C_2^2,G)^2 = \frac{3}{6}\left(\frac{81}{36} + \frac{49}{9}\right) + \frac{3}{6}\left(\frac{81}{36} + \frac{49}{9}\right) = \frac{277}{36} \approx 7{,}694$,

puis on vérifie la *décomposition de Huygens* :

$\mathcal{I}(\Gamma) = \mathcal{I}_{inter2} + \mathcal{I}_{intra2} = \frac{277}{36} + \frac{19}{9} = \frac{353}{36}$

ou $\mathcal{I}(\Gamma) = \mathcal{I}_{inter2} + \mathcal{I}_{intra2} \approx 7{,}694 + 2{,}111 \approx 9{,}805$.