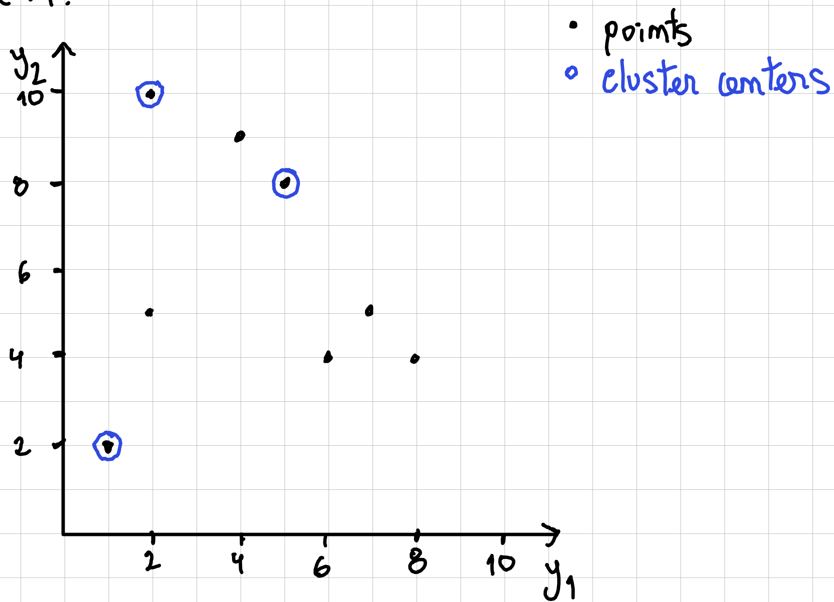
Pedro Nunes 109368
Andrei Barb 109762

# Part A.

a)




b) $\mu_1 = (2,10)$ $\mu_2 = (5,8)$ $\mu_3 = (1,2)$

$d(x_1, \mu_1) = \sqrt{(2-2)^2 + (10-10)^2} = 0 \leftarrow$

$d(x_1, \mu_2) = \sqrt{13}$

$d(x_1, \mu_3) = \sqrt{65}$

$d(x_2, \mu_1) = 5$

$d(x_2, \mu_2) = 3\sqrt{2}$

$d(x_2, \mu_3) = \sqrt{10} \leftarrow$

$d(x_3, \mu_1) = 6\sqrt{2}$

$d(x_3, \mu_2) = 5 \leftarrow$

$d(x_3, \mu_3) = \sqrt{53}$

$d(x_4, \mu_1) = \sqrt{13}$
$d(x_4, \mu_2) = 0 \leftarrow$
$d(x_4, \mu_3) = 2\sqrt{13}$

$d(x_5, \mu_1) = 5\sqrt{2}$
$d(x_5, \mu_2) = \sqrt{13} \leftarrow$
$d(x_5, \mu_3) = 3\sqrt{5}$

$d(x_6, \mu_1) = 2\sqrt{13}$
$d(x_6, \mu_2) = \sqrt{17} \leftarrow$
$d(x_6, \mu_3) = \sqrt{29}$

$d(x_7, \mu_1) = \sqrt{65}$
$d(x_7, \mu_2) = 2\sqrt{13}$
$d(x_7, \mu_3) = 0 \leftarrow$

$d(x_8, \mu_1) = \sqrt{5}$
$d(x_8, \mu_2) = \sqrt{2} \leftarrow$
$d(x_8, \mu_3) = \sqrt{58}$

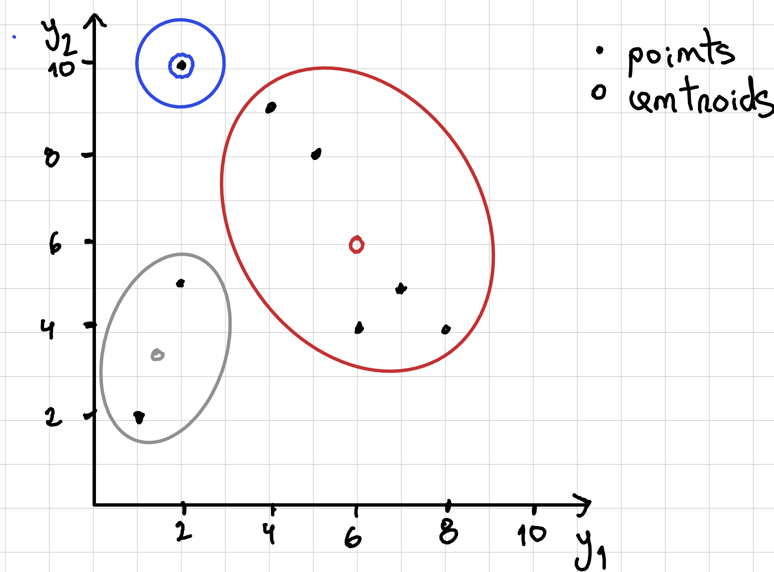
$C_1 = \{x_1\}$ $\quad C_2 = \{x_3, x_4, x_5, x_6, x_8\}$ $\quad C_3 = \{x_2, x_7\}$

$\mu_1 = (2, 10)$

$$\mu_2 = \left( \frac{8+5+7+6+4}{5}, \frac{4+8+5+4+9}{5} \right) = (6, 6)$$

$$\mu_3 = \left( \frac{2+1}{2}, \frac{5+2}{2} \right) = (1{,}5 \ ; \ 3{,}5)$$

c) $\mu_1 = (2, 10)$ $\mu_2 = (6, 6)$ $\mu_3 = (1{,}5 ; 3{,}5)$




d) The initialization of centroids in the k-means algorithm has a significant impact on both the convergence process, performance and the final results.

Because the k-means tries to minimize an objective functi[on], different initializations can lead the algorithm to converge to different local minimums.

If the initial centroids are chosen well, the algorithm quickly converges to clusters that accurately represent the data structure resolting in lower within-cluster variance. However, poor initial centroids can slow convergence, resulting in a worst performance, produce empty or unbalanced clusters, and lead to suboptimal partitions that do not reflect the true relationships between data points. In some cases, two different initializations can yield complety different final cluster assignements.

Part B

a) $\mu = \frac{1}{3}\left(\begin{bmatrix}5\\0\\1\end{bmatrix} + \begin{bmatrix}0\\5\\0\end{bmatrix} + \begin{bmatrix}1\\0\\-1\end{bmatrix}\right) = \begin{bmatrix}2\\5/3\\0\end{bmatrix}$

$$x_1 - \mu = \begin{bmatrix}5\\0\\1\end{bmatrix} - \begin{bmatrix}2\\5/3\\0\end{bmatrix} = \begin{bmatrix}3\\-5/3\\1\end{bmatrix}$$

$$x_2 - \mu = \begin{bmatrix}0\\5\\0\end{bmatrix} - \begin{bmatrix}2\\5/3\\0\end{bmatrix} = \begin{bmatrix}-2\\10/3\\0\end{bmatrix}$$

$$x_3 - \mu = \begin{bmatrix}1\\0\\-1\end{bmatrix} - \begin{bmatrix}2\\5/3\\0\end{bmatrix} = \begin{bmatrix}-1\\-5/3\\-1\end{bmatrix}$$

$$\tilde{X} = \begin{bmatrix}3 & -5/3 & 1\\-2 & 10/3 & 0\\-1 & -5/3 & -1\end{bmatrix}$$

$$C = \frac{1}{m}\tilde{X}^T\tilde{X} = \frac{1}{3}\begin{bmatrix}3 & -2 & -1\\-5/3 & 10/3 & -5/3\\1 & 0 & -1\end{bmatrix}\begin{bmatrix}3 & -5/3 & 1\\-2 & 10/3 & 0\\-1 & -5/3 & -1\end{bmatrix}$$

$$= \begin{bmatrix}14/3 & -10/3 & 4/3\\-10/3 & 50/9 & 0\\4/3 & 0 & 2/3\end{bmatrix}$$

b) $\det(C - \lambda I) = 0$

$$\Leftrightarrow \begin{vmatrix} 14/3 - \lambda & -10/3 & 4/3 \\ -10/3 & 50/9 - \lambda & 0 \\ 4/3 & 0 & 2/3 - \lambda \end{vmatrix} = 0$$

$$\Leftrightarrow \left(\frac{14}{3} - \lambda\right)\left(\frac{50}{9} - \lambda\right)\left(\frac{2}{3} - \lambda\right) + 0 + 0 - \left(\frac{4}{3} \times \frac{4}{3} \times \left(\frac{50}{9} - \lambda\right) + \left(\frac{-10}{3} \times \frac{-10}{3} \times \left(\frac{2}{3} - \lambda\right)\right)\right) = 0$$

$$\Leftrightarrow \left(\frac{14}{3} - \lambda\right)\left(\frac{50}{9} - \lambda\right)\left(\frac{2}{3} - \lambda\right) - \left(\frac{800}{81} - \frac{16}{9}\lambda\right) - \left(\frac{200}{27} - \frac{100}{9}\lambda\right) = 0$$

$$\Leftrightarrow \left(\frac{700}{27} - \frac{14\lambda}{3} - \frac{50}{9}\lambda + \lambda^2\right)\left(\frac{2}{3} - \lambda\right) - \left(\frac{800}{81} - \frac{16}{9}\lambda\right) - \left(\frac{200}{27} - \frac{100\lambda}{9}\right) = 0$$

$$\Leftrightarrow \frac{1400}{81} - \frac{700\lambda}{27} - \frac{28}{9}\lambda + \frac{14}{3}\lambda^2 - \frac{100}{27}\lambda + \frac{50}{9}\lambda^2 + \frac{2}{3}\lambda^2 - \lambda^3 - \frac{800}{81} + \frac{16}{9}\lambda - \frac{200}{27} + \frac{100\lambda}{9} = 0$$

$$\Leftrightarrow -\lambda^3 + \frac{98}{9}\lambda^2 - \frac{536}{27}\lambda = 0 \Leftrightarrow \lambda\left(-\lambda^2 + \frac{98}{9} - \frac{536}{27}\right) = 0$$

$$\Leftrightarrow \lambda_3 = 0 \vee \lambda_1 = 8{,}5733 \vee \lambda_2 = 2{,}3155$$

$\lambda_1: \quad (C - \lambda_1 I)\,\mu_1 = 0$

$$\begin{bmatrix} 14/3 - 8{,}5733 & -10/3 & 4/3 \\ -10/3 & 50/9 - 8{,}5733 & 0 \\ 4/3 & 0 & 2/3 - 8{,}5733 \end{bmatrix} \begin{bmatrix} \mu_{11} \\ \mu_{12} \\ \mu_{13} \end{bmatrix} = 0$$

$$\Leftrightarrow \begin{cases} -3{,}906\,\mu_{11} - 10/3\,\mu_{12} + 4/3\,\mu_{13} = 0 \\ -10/3\,\mu_{11} - 3{,}0177\,\mu_{12} = 0 \\ 4/3\,\mu_{11} - 7{,}9066\,\mu_{13} = 0 \end{cases}$$

$$\mu_{12} = -1{,}104594\,\mu_{11}$$

$$\mu_{13} = 0{,}16863548\,\mu_{11}$$

if $\mu_{11} = 1$ them: $\mu_1 = \begin{bmatrix} 1 \\ -1{,}104594 \\ 0{,}16863548 \end{bmatrix}$

Normalizing: $\mu_1 = \begin{bmatrix} 0{,}6669 \\ -0{,}7366 \\ 0{,}1125 \end{bmatrix}$

$\lambda_2 : (C - \lambda_2 I)\, u_2 = 0$

$$\begin{bmatrix} 14/3 - 2{,}3155 & -10/3 & 4/3 \\ -10/3 & 50/9 - 2{,}3155 & 0 \\ 4/3 & 0 & 2/3 - 2{,}3155 \end{bmatrix} \begin{bmatrix} u_{21} \\ u_{22} \\ u_{23} \end{bmatrix} = 0$$

$$\Leftrightarrow \begin{cases} 2{,}35117\, u_{21} & -10/3\, u_{22} & -4/3\, u_{23} \\ -10/3\, u_{21} & +3{,}24\, u_{22} & = 0 \\ 4/3\, u_{21} & -1{,}6488\, u_{23} = 0 \end{cases}$$

$u_{22} = 1{,}0288\, u_{21}$

$u_{23} = 0{,}80867\, u_{21}$

if $u_{21} = 1$ then $u_2 = \begin{bmatrix} 1 \\ 1{,}0288 \\ 0{,}80867 \end{bmatrix}$

Normalizing : $u_2 = \begin{bmatrix} 0{,}6072 \\ 0{,}6247 \\ 0{,}4910 \end{bmatrix}$

Plano de projeção = Span $\{u_1, u_2\}$

where $u_1 = \begin{bmatrix} 0{,}6669 \\ -0{,}7366 \\ 0{,}1125 \end{bmatrix}$ and $u_2 = \begin{bmatrix} 0{,}6072 \\ 0{,}6247 \\ 0{,}4910 \end{bmatrix}$

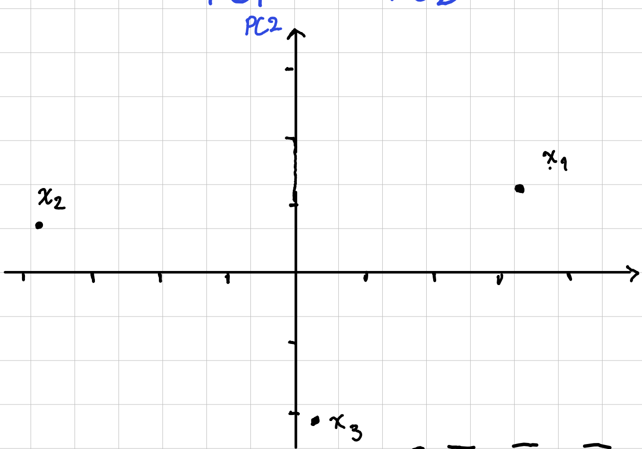
c) $X' = \tilde{X} U$

$$U = [u_1 \quad u_2] = \begin{bmatrix} 0{,}6669 & 0{,}6072 \\ -0{,}7366 & 0{,}6247 \\ 0{,}1125 & 0{,}4910 \end{bmatrix}$$

$$\tilde{X} = \begin{bmatrix} 3 & -5/3 & 1 \\ -2 & 10/3 & 0 \\ -1 & -5/3 & -1 \end{bmatrix}$$

$$X' = \begin{bmatrix} 3 & -5/3 & 1 \\ -2 & 10/3 & 0 \\ -1 & -5/3 & -1 \end{bmatrix} \begin{bmatrix} 0{,}6669 & 0{,}6072 \\ -0{,}7366 & 0{,}6247 \\ 0{,}1125 & 0{,}4910 \end{bmatrix}$$

$$= \begin{array}{c} \begin{bmatrix} 3{,}341 & 1{,}2714 \\ -3{,}789 & 0{,}8679 \\ 0{,}448 & -2{,}139 \end{bmatrix} \\ \text{PC1} \qquad \text{PC2} \end{array} \begin{array}{l} \leftarrow x_1\,(3{,}341\,;\,1{,}2714) \quad + \\ \leftarrow x_2\,(-3{,}789\,;\,0{,}8679) \quad + \\ \leftarrow x_3\,(0{,}448\,;\,-2{,}139) \quad - \\ \\ \end{array}$$




This projection plane helps discriminate between the two classes.
Considering the positioning of the points in the graph we can conclude that: when the PC2 is positive the class is "+", an when is negative the class is "-". This shows that PC2, alone, can discriminate the classes