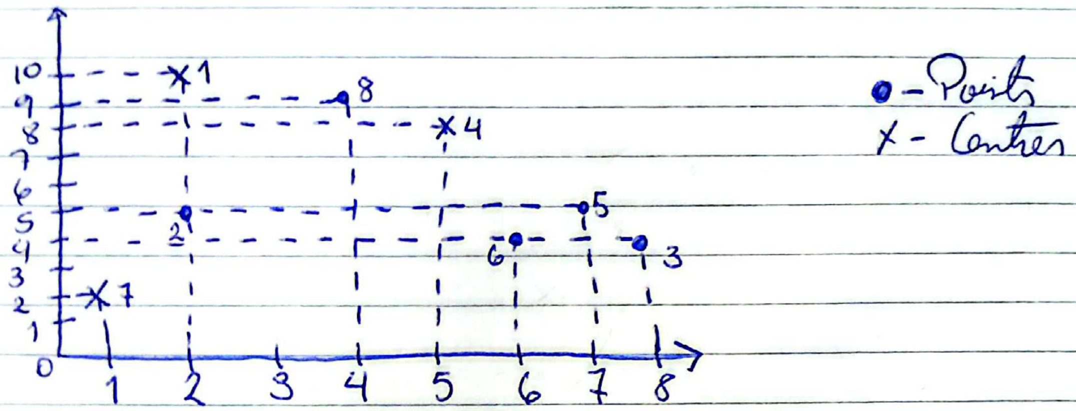
HW-4

Part A

1. (2,10) 2. (2,5) 3. (8,4) 4. (5,8)
5. (7,5) 6. (6,4) 7. (1,2) 8. (4,9)

a)



b) $C_1 = (2,10)$ $C_2 = (5,8)$ $C_3 = (1,2)$

$x_2 \rightarrow d(x_2, c_1) = \sqrt{(2-2)^2 + (5-10)^2} = 5$
$d(x_2, c_2) = \sqrt{(2-5)^2 + (5-8)^2} = \sqrt{18} \approx 4{,}24$
$d(x_2, c_3) = \sqrt{(2-1)^2 + (5-2)^2} = \sqrt{10} \approx 3{,}163$
} C_3

$x_3 \rightarrow d(x_3, c_1) = \sqrt{(8-2)^2 + (4-10)^2} = \sqrt{52} \approx 7{,}21$
$d(x_3, c_2) = \sqrt{(8-5)^2 + (4-8)^2} = \sqrt{25} = 5$
$d(x_3, c_3) = \sqrt{(8-1)^2 + (4-2)^2} = \sqrt{53} \approx 7{,}28$
} C_2

$x_5 \rightarrow d(x_5, c_1) = \sqrt{(7-2)^2 + (5-10)^2} = \sqrt{50} = 7{,}071$
$d(x_5, c_2) = \sqrt{(7-5)^2 + (5-8)^2} = \sqrt{13} \approx 3{,}61$
$d(x_5, c_3) = \sqrt{(7-1)^2 + (5-2)^2} = \sqrt{45} \approx 6{,}71$
} C_2

$x_6 \rightarrow d(x_6, c_1) = \sqrt{(6-2)^2 + (4-10)^2} = \sqrt{40} \approx 6{,}324$
$d(x_6, c_2) = \sqrt{(6-5)^2 + (4-8)^2} = \sqrt{15} \approx 3{,}87$
$d(x_6, c_3) = \sqrt{(6-1)^2 + (4-2)^2} = \sqrt{29} \approx 5{,}385$
} C_2

$x_8 \rightarrow d(x_8, c_1) = \sqrt{(4-2)^2 + (9-10)^2} = \sqrt{5} \approx 2.236$
$d(x_8, c_2) = \sqrt{(4-5)^2 + (9-8)^2} = \sqrt{2} \approx 1{,}41$
$d(x_8, c_3) = \sqrt{(4-1)^2 + (9-2)^2} = \sqrt{58} \approx 7{,}62$
} C_2

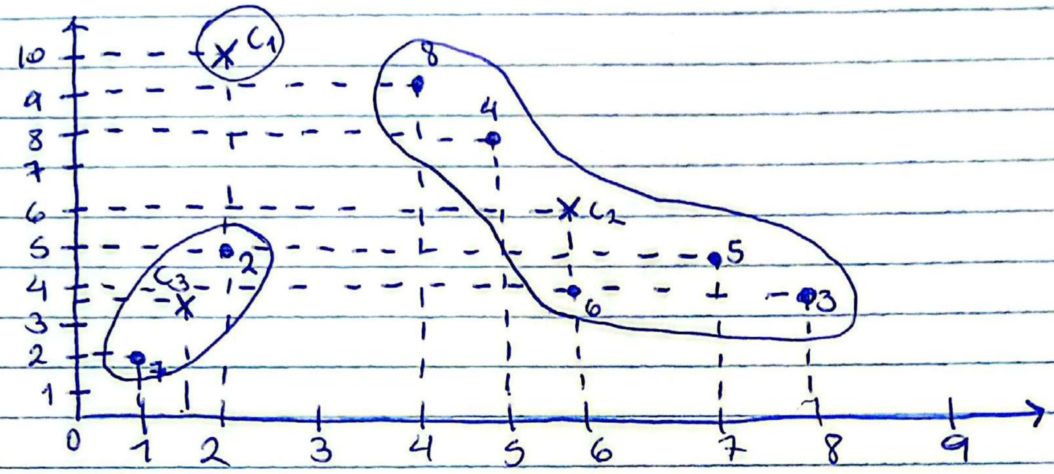
CamScanner

c) $C_1\{x_1\} = (2,10)$

$$C_2\{x_3, x_4, x_5, x_6, x_8\} = \left(\frac{8+5+7+6+4}{5}, \frac{4+8+5+4+9}{5}\right)$$

$= (6,6)$

$$C_3\{x_2, x_7\} = \left(\frac{1+2}{2}, \frac{5+2}{2}\right) = (1.5; 3.5)$$



d) Depending on how centroids are initialized, they can alter the result to concluding in no convergence. If centroids are too distant from the actual groups, the algorithm can be forced to execute several iteration, or no convergence at all due to local minimums.

Bad initialization can lead to the adoption of adding square errors of higher magnitude.

Part B

a) $$\bar{x} = \left(\frac{5+0+1}{3}, \frac{0+5+0}{3}, \frac{1+0-1}{3}\right) = (2; 1{,}67; 0)$$

$$x_1 - \bar{x} = \begin{pmatrix} 3 \\ -1{,}67 \\ 1 \end{pmatrix} \quad x_2 - \bar{x} = \begin{pmatrix} -2 \\ 3.33 \\ 0 \end{pmatrix} \quad x_3 - \bar{x} = \begin{pmatrix} -1 \\ -1{,}67 \\ -1 \end{pmatrix}$$

$$\bar{x} = \begin{pmatrix} 3 & -1{,}67 & 1 \\ -2 & 3.33 & 0 \\ -1 & -1{,}67 & -1 \end{pmatrix}$$

$$C = \frac{1}{3}\begin{pmatrix} 3 & -2 & -1 \\ -1{,}67 & 3.33 & -1{,}67 \\ 1 & 0 & -1 \end{pmatrix}\begin{pmatrix} 3 & -1{,}67 & 1 \\ -2 & 3.33 & 0 \\ -1 & -1{,}67 & -1 \end{pmatrix}$$

$$= \frac{1}{3}\begin{pmatrix} 14 & -10 & 4 \\ -10 & 16{,}67 & 0 \\ 4 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 14/3 & -10/3 & 4/3 \\ -10/3 & 5.557 & 0 \\ 4/3 & 0 & 2/3 \end{pmatrix}$$

b) $\det(C - \lambda I) = \begin{vmatrix} 14/3-\lambda & -10/3 & 4/3 \\ -10/3 & 5.557-\lambda & 0 \\ 4/3 & 0 & 2/3-\lambda \end{vmatrix} =$

$$= \left(\frac{14}{3}-\lambda\right)\left(\frac{50}{9}-\lambda\right)\left(\frac{2}{3}-\lambda\right) - \frac{16}{9}\left(\frac{50}{9}-\lambda\right) - \left(\frac{100}{9}\left(\frac{2}{3}-\lambda\right)\right)$$

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

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

$$\Leftrightarrow -\lambda^3 + \frac{98}{9}\lambda^2 - \frac{536}{27}\lambda = 0 \Leftrightarrow \lambda = 0 \vee -\lambda^2 + \frac{98}{9}\lambda - \frac{536}{27} = 0$$

$$\Leftrightarrow \lambda = 0 \vee \lambda = 2.315 \vee \lambda = 8.575$$

$$\lambda = 0 \rightarrow \begin{pmatrix} -1/2 \\ -3/10 \\ 1 \end{pmatrix} : V_1$$

$$\lambda = 2.315 \rightarrow \begin{pmatrix} 1.2366 \\ 1.2723 \\ 1 \end{pmatrix} : V_2$$

Projection plane : Span $\{V_2, V_3\}$

$$\lambda = 8.575 \Rightarrow \begin{pmatrix} 5.93 \\ -6.55 \\ 1 \end{pmatrix} : V_3$$

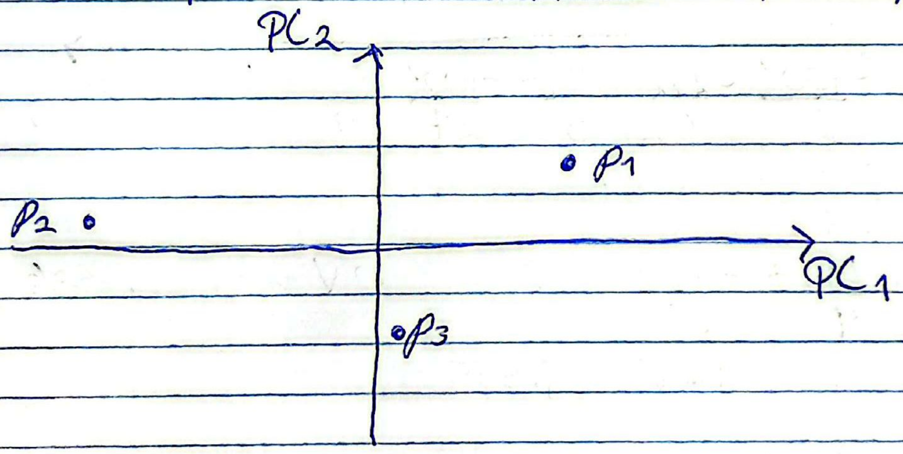
$$||V_2|| = \sqrt{(1.2366)^2 + (1.2723)^2 + 1} \approx 2.037$$

V_2 normalized: $\begin{pmatrix} 0.6071 \\ 0.6246 \\ 0.491 \end{pmatrix}$

$$||V_3|| = \sqrt{(5.93)^2 + (-6.55)^2 + 1} \approx 8.892$$

V_3 normalized: $\begin{pmatrix} 0.667 \\ -0.7367 \\ 0.1125 \end{pmatrix}$

$$\tilde{x} = \begin{pmatrix} 3 & -2 & -1 \\ -1{,}67 & 3.33 & -1{,}67 \\ 1 & 0 & -1 \end{pmatrix} \Rightarrow \tilde{x} \begin{pmatrix} V_2, V_3 \end{pmatrix} = \begin{array}{c|l} \begin{pmatrix} 3{,}341 & 1{,}2714 \\ -3{,}789 & 0{,}8679 \\ 0{,}448 & -2{,}139 \end{pmatrix} & \begin{array}{l} p_1 + \\ p_2 + \\ p_3 - \end{array} \end{array}$$



For a positive PC2, all plotted values correspond to the binary value ("+" or positive), which are the same as the y4 expected result.
The same behaviour is exhibited by the p_3 value, although for its corresponding negative value.