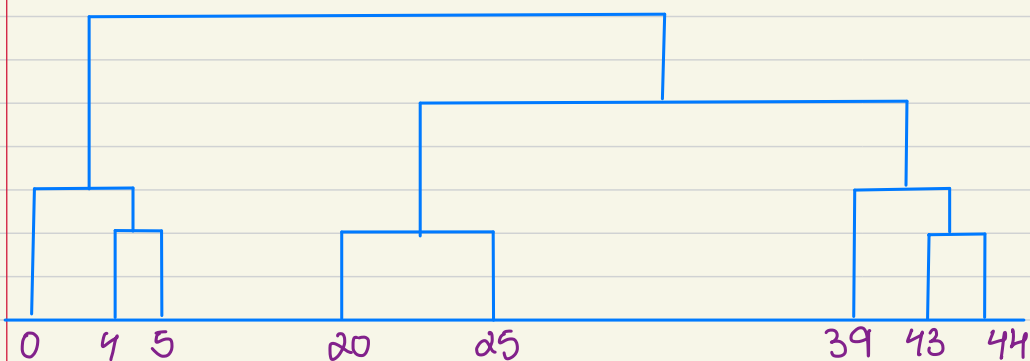
Homework 4

Due: 04/14/2022

Devika Chandnani
A13405666

Q1 Code and Readme in the ZIP file

Q2 a) Dendogram: {0, 4, 5, 20, 25, 39, 43, 44}



b) Cluster 1 = {0, 4, 5}
Cluster 2 = {20, 25, 39, 43, 44}

Q3 $m_1 = \frac{1}{3}\left[\begin{pmatrix}1\\1\end{pmatrix} + \begin{pmatrix}2\\2\end{pmatrix} + \begin{pmatrix}3\\3\end{pmatrix}\right] = 1/3\begin{bmatrix}6\\6\end{bmatrix}$

a) $m_1 = (2, 2)$

$m_2 = \frac{1}{5}\left[\begin{pmatrix}5\\2\end{pmatrix} + \begin{pmatrix}6\\2\end{pmatrix} + \begin{pmatrix}7\\2\end{pmatrix} + \begin{pmatrix}8\\2\end{pmatrix} + \begin{pmatrix}9\\2\end{pmatrix}\right]$

$= 1/5\left[\begin{pmatrix}35\\10\end{pmatrix}\right]$

$m_2 = (7, 2)$

b) Total Mean $(m) = 1/8\left[\begin{pmatrix}6\\6\end{pmatrix} + \begin{pmatrix}35\\10\end{pmatrix}\right]$

$m = 1/8\begin{bmatrix}41\\16\end{bmatrix}$

$m = \begin{pmatrix}5.125\\2\end{pmatrix}$

c) Scatter Matrices

$S_1 = \left[\begin{pmatrix}1\\1\end{pmatrix} - \begin{pmatrix}2\\2\end{pmatrix}\right]\left[\begin{pmatrix}1\\1\end{pmatrix} - \begin{pmatrix}2\\2\end{pmatrix}\right]^T + \left[\begin{pmatrix}2\\2\end{pmatrix} - \begin{pmatrix}2\\2\end{pmatrix}\right]\left[\begin{pmatrix}2\\2\end{pmatrix} - \begin{pmatrix}2\\2\end{pmatrix}\right]^T$

$+ \left[\begin{pmatrix}3\\3\end{pmatrix} - \begin{pmatrix}2\\2\end{pmatrix}\right]\left[\begin{pmatrix}3\\3\end{pmatrix} - \begin{pmatrix}2\\2\end{pmatrix}\right]^T + \begin{bmatrix}1\\1\end{bmatrix}\begin{bmatrix}1\\1\end{bmatrix}^T$

$= \begin{bmatrix}1&1\\1&1\end{bmatrix} + \begin{bmatrix}0&0\\0&0\end{bmatrix} + \begin{bmatrix}1&1\\1&1\end{bmatrix}$

$S_1 = \begin{bmatrix}2&2\\2&2\end{bmatrix}$

$$S_2 = \left[\begin{pmatrix}5\\2\end{pmatrix}-\begin{pmatrix}7\\2\end{pmatrix}\right]\left[\begin{pmatrix}5\\2\end{pmatrix}-\begin{pmatrix}7\\2\end{pmatrix}\right]^T + \left[\begin{pmatrix}6\\2\end{pmatrix}-\begin{pmatrix}7\\2\end{pmatrix}\right]\left[\begin{pmatrix}6\\2\end{pmatrix}-\begin{pmatrix}7\\2\end{pmatrix}\right]^T$$

$$+ \left[\begin{pmatrix}7\\2\end{pmatrix}-\begin{pmatrix}7\\2\end{pmatrix}\right]\left[\begin{pmatrix}7\\2\end{pmatrix}-\begin{pmatrix}7\\2\end{pmatrix}\right]^T + \left[\begin{pmatrix}8\\2\end{pmatrix}-\begin{pmatrix}7\\2\end{pmatrix}\right]\left[\begin{pmatrix}8\\2\end{pmatrix}-\begin{pmatrix}7\\2\end{pmatrix}\right]^T$$

$$+ \left[\begin{pmatrix}9\\2\end{pmatrix}-\begin{pmatrix}7\\2\end{pmatrix}\right]\left[\begin{pmatrix}9\\2\end{pmatrix}-\begin{pmatrix}7\\2\end{pmatrix}\right]^T$$

$$= \begin{bmatrix}-2\\0\end{bmatrix}\begin{bmatrix}-2\\0\end{bmatrix}^T + \begin{bmatrix}-1\\0\end{bmatrix}\begin{bmatrix}-1\\0\end{bmatrix}^T + \begin{bmatrix}0\\0\end{bmatrix}\begin{bmatrix}0\\0\end{bmatrix}^T + \begin{bmatrix}1\\0\end{bmatrix}\begin{bmatrix}1\\0\end{bmatrix}^T + \begin{bmatrix}2\\0\end{bmatrix}\begin{bmatrix}2\\0\end{bmatrix}^T$$

$$S_2 = \begin{bmatrix}4&0\\0&0\end{bmatrix} + \begin{bmatrix}1&0\\0&0\end{bmatrix} + \begin{bmatrix}0&0\\0&0\end{bmatrix} + \begin{bmatrix}1&0\\0&0\end{bmatrix} + \begin{bmatrix}4&0\\0&0\end{bmatrix}$$

$$S_2 = \begin{bmatrix}10&0\\0&0\end{bmatrix}$$

d) $S_w = S_1 + S_2 = \begin{bmatrix}2&2\\2&2\end{bmatrix} + \begin{bmatrix}10&0\\0&0\end{bmatrix}$

$$S_w = \begin{bmatrix}12&2\\2&2\end{bmatrix}$$

e) $$S_B = 3\left[\begin{pmatrix}2\\2\end{pmatrix}-\begin{pmatrix}5.125\\2\end{pmatrix}\right]\left[\begin{pmatrix}2\\2\end{pmatrix}-\begin{pmatrix}5.125\\2\end{pmatrix}\right]^T + 5\left[\begin{pmatrix}7\\2\end{pmatrix}-\begin{pmatrix}5.125\\2\end{pmatrix}\right]\left[\begin{pmatrix}7\\2\end{pmatrix}-\begin{pmatrix}5.125\\2\end{pmatrix}\right]^T$$

$$S_B = \begin{bmatrix}29.3&0\\0&0\end{bmatrix} + \begin{bmatrix}17.6&0\\0&0\end{bmatrix} = \begin{bmatrix}46.9&0\\0&0\end{bmatrix}$$

f) $$\frac{tr(S_B)}{tr(S_w)} = \frac{46.9+0}{12+2} = \frac{46.9}{14}$$

Q4

Instance #	True Class Label	Predicted Probability of Positive Class		
1	N	0.90	P	False
2	P	0.82	P	True
3	P	0.78	P	True
4	P	0.66	P	True
5	P	0.60	P	True
6	P	0.52	P	True
7	N	0.43	N	True
8	N	0.42	N	True
9	P	0.41	N	False
10	P	0.4	N	False

1. Confusion Matrix :

	Predicted Positive	Predicted Negative
Actual Positive	5	2
Actual Negative	1	2

2. Accuracy Precision $= (TP + TN) \div (TP + TN + FP + FN)$
$= (5+2) \div (10)$
$= 7/10$
$= 0.7$

3. Recall $= TP / (TP + FN)$
$= 5 / (5 + 2) = 5/7$

4. F1 Score $= 2PR(P+R) = 2(0.7)(5/7) / (0.7)(5/7)$
$= 1/0.5$
$= 2$

5. Specificity $= TN / (TN + FP)$
$= 2 / (2 + 1)$
$= 2/(3)$
$= 2/3$