Exercise 1

1 a) $S_B = \begin{pmatrix} (\mu_{1x} - \mu_{2x}) \\ (\mu_{1y} - \mu_{2y}) \end{pmatrix} \begin{pmatrix} (\mu_{1x} - \mu_{2x}) & (\mu_{1y} - \mu_{2y}) \end{pmatrix}$

$$= \begin{pmatrix} -2.33 \\ -2.40 \end{pmatrix} \begin{pmatrix} -2.33 & -2.40 \end{pmatrix}$$

$$= \begin{pmatrix} (-2.33)^2 & (-2.33)(-2.40) \\ (-2.40)(-2.33) & (-2.40)^2 \end{pmatrix}$$

$$= \begin{pmatrix} 5.43 & 5.59 \\ 5.59 & 5.76 \end{pmatrix}$$

b) rank of $S_B = 1$

2 a)

figure	$(\tilde{\mu}_1 - \tilde{\mu}_2)^2$	$\tilde{\sigma}_1^2$	$\tilde{\sigma}_2^2$	w	$J(w)$
3 (w_1)	11.23	1.40	2.09	$\begin{pmatrix} 0.70 \\ 0.72 \end{pmatrix}$	3.21
4 (w_2)	0	0.28	0.74	$\begin{pmatrix} 0.72 \\ -0.70 \end{pmatrix}$	0
5 (w_3)	3.04	0.16	0.06	$\begin{pmatrix} 0.25 \\ -0.97 \end{pmatrix}$	13.81

b) figure 4 has $(\tilde{\mu}_1 - \tilde{\mu}_2)^2 = 0$, as both the projection lies on the same point i.e. $(0,0)$

figure 3 has $(\tilde{\mu}_1 - \tilde{\mu}_2)^2 = (\mu_1 - \mu_2)^2$, since the projection is exactly where it lies, thus the original distance is what we need:

$$\therefore \left\| \begin{pmatrix} 2.34 \\ 2.4 \end{pmatrix} \right\|_2^2 = 5.475 + 5.76 = 11.23$$

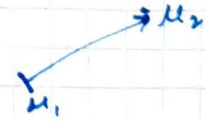


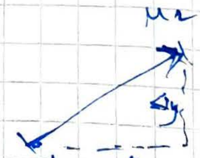
fig 5: $\tilde{\mu}_1 = w^T \cdot \mu_1 = (0.25\ \ -0.97) \begin{pmatrix} 1.17 \\ 1.20 \end{pmatrix} = -0.872$

$\tilde{\mu}_2 = w^T \cdot \mu_2 = 0.872$

$\therefore (\mu_1 - \mu_2)^2 = (-1.744)^2 = 3.04$

c) $w_1 = \frac{\Delta x}{\|\mu_2 - \mu_1\|}, \frac{\Delta y}{\|\mu_2 - \mu_1\|} = \frac{2.34}{\sqrt{11.23}}, \frac{2.40}{\sqrt{11.23}}$

$= 0.698,\ 0.716$



w_2: Since $w_2^T \cdot \mu_1 = w_2^T \cdot \mu_2 = 0$. Thus the $\vec{w_2} \perp \vec{w_1}$, from perceptron chapter.

d) $$J(w) = \frac{(\tilde{\mu}_1 - \tilde{\mu}_2)^2}{\tilde{\sigma}_1^2 + \tilde{\sigma}_2^2}$$

Values are shown ~~at~~ in the table

3). As can be seen in figure 3. Even after ~~ada~~ achieving a maximum squared mean distance, class seperation is not perfect. We do have some misclassification.
Therefore as shown in figure 5, we also need to minimize the variances within the class.

4) a) $\lambda_2 = 0$, because there can be at most 1 ~~eigenvalue~~ non-zero eigen value as rank of $S_B = 1$.

This is not specific to this subset but rather a general rule.

b) $\tilde{x}_1 = (0.25 \quad -0.97) \begin{pmatrix} 0.38 \\ -0.42 \end{pmatrix} = 0.095 + 1.39 = 1.485$

$\tilde{x}_2 \quad (0.25 \quad -0.97) \begin{pmatrix} 1.88 \\ 1.08 \end{pmatrix} = 0.47 - 1.047 = -0.58$

c) y is more important for class discrimination because, ~~x~~ ~~produces~~ ~~high resultant value~~ ~~than~~ for same x but different y, the points can belong to different class in ~~pro~~ projected plane, but for same y & different x we cannot say so. Also because in the resultant $\tilde{x}_1$, y's contribution is more than x's contribution

d) $w_0 = (-0.05)$