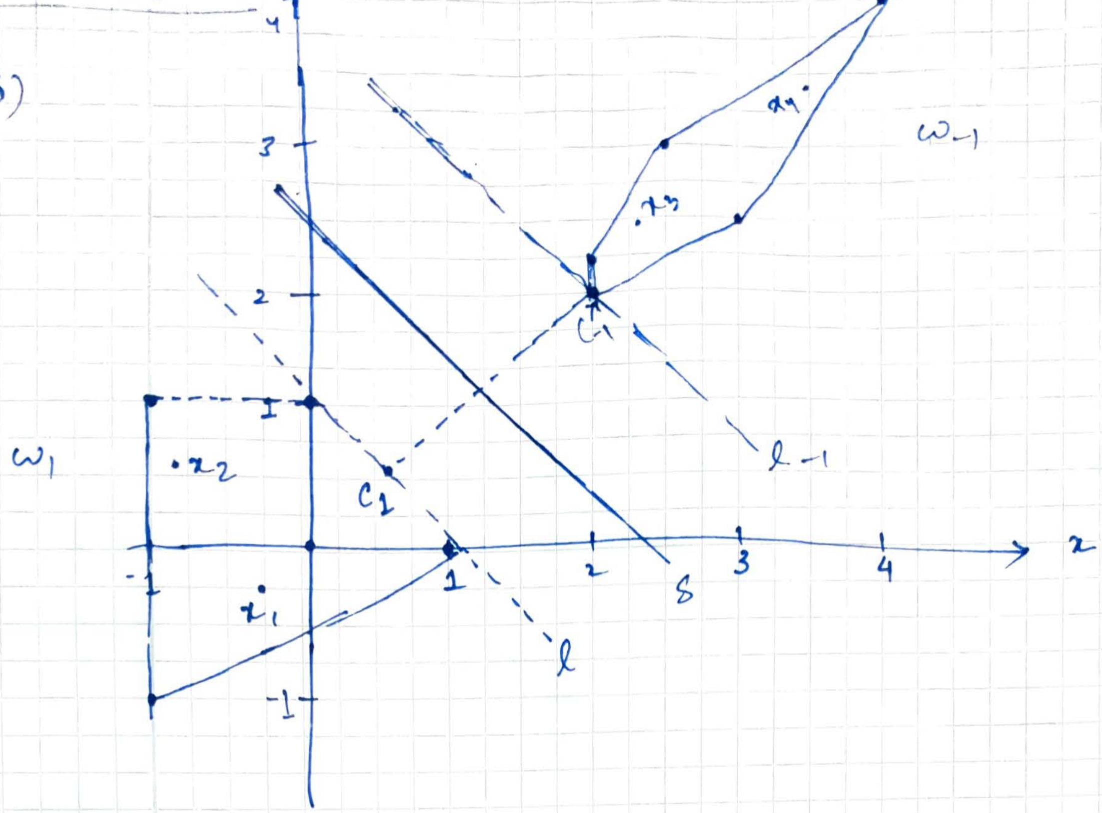
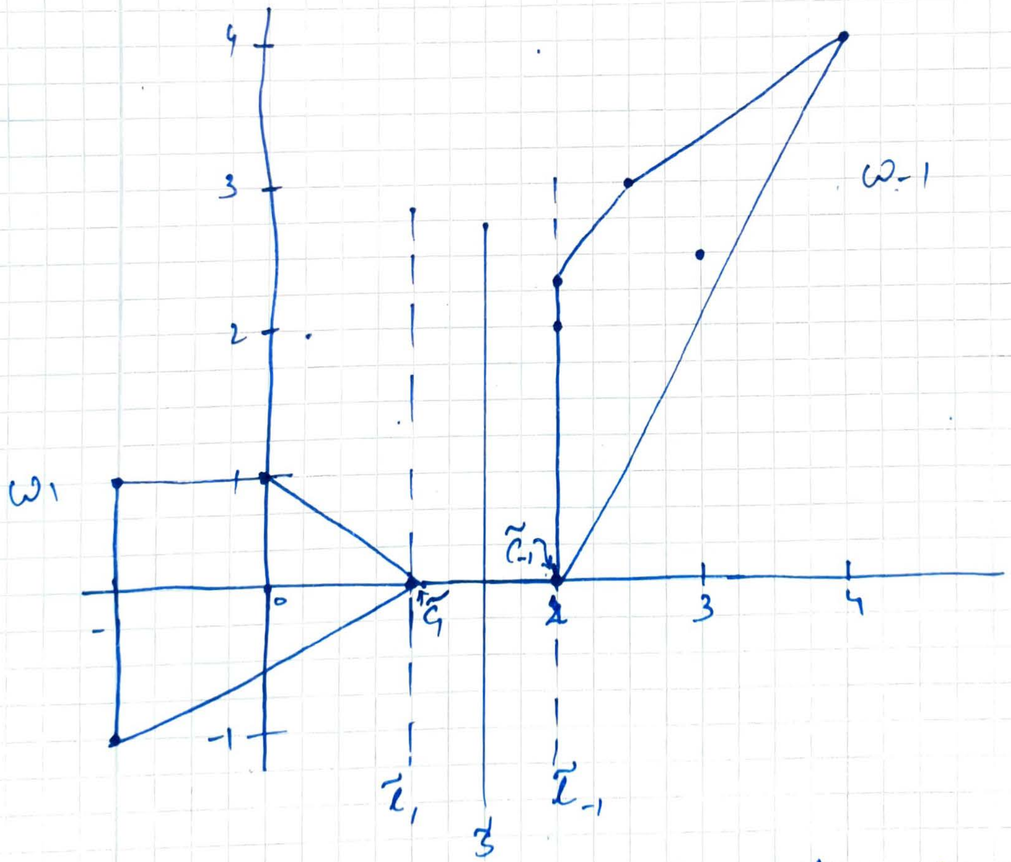
## Exercise 1:

(1-5)




6)




7) In both approach we try to find hyperplanes which divides the linearly seperable classes data points. The only difference is that in SVM, we find a separation plane with maximal distance to the data. But in perception, we just find hyperplane without any restriction.

Exercise 2 :

$$x_1 = \begin{pmatrix}1\\1\end{pmatrix} \in \omega_1 \qquad x_2 = \begin{pmatrix}2\\3\end{pmatrix} \in \omega_2$$

$$\omega_1 = \{x_i \mid T_i = 1\} \qquad \omega_2 = \{x_i \mid T_i = -1\}$$

$$T_1 = 1 \qquad T_2 = -1$$

Given: $L(\omega, \omega_0, \alpha_1, \alpha_2) = L(\alpha_1, \alpha_2) = \sum_{i=1}^{2} \alpha_i - \frac{1}{2}\sum_{i=1}^{2}\sum_{j=1}^{2} \alpha_i \alpha_j T_i T_j x_i^T x_j$

constraint : $\sum_{i=1}^{2} \alpha_i T_i = 0$

1). Opening up the summation :

$$L(\alpha_1, \alpha_2) = \sum_{i=1}^{2} \alpha_i - \frac{1}{2}\left(\alpha_1 \alpha_1 T_1 T_1 x_1^T \cdot x_1 + \alpha_1 \alpha_2 T_1 T_2 x_1^T x_2 + \alpha_2 \cdot \alpha_1 T_2 T_1 x_2^T \cdot x_1 + \alpha_2 \cdot \alpha_2 T_2 T_2 x_2^T \cdot x_2\right)$$

$$= \alpha_1 + \alpha_2 - \frac{1}{2}\left(\alpha_1^2 T_1^2 \cdot x_1^T \cdot x_1 + \alpha_1 \alpha_2 T_1 T_2 x_1^T \cdot x_2 + \alpha_2 \alpha_1 T_2 T_1 x_2^T \cdot x_1 + \alpha_2^2 \cdot T_2^2 \cdot x_2^T \cdot x_2\right)$$

(Simplifying & using the constraint given ( $\sum_{i=1}^{2} \alpha_i T_i = 0$ )

$$\downarrow$$

( $\alpha_1 T_1 + \alpha_2 T_2 = 0 \Rightarrow \alpha_1 - \alpha_2 = 0 \Rightarrow \alpha_1 = \alpha_2$

But before that, ~~simplifying~~ substituting the value of $T_i$ & $x_i$.

$$= \alpha_1 + \alpha_2 - \frac{1}{2}\left(\alpha_1^2 (1)^2 \cdot (1\ 1)\begin{pmatrix}1\\1\end{pmatrix} + \alpha_1 \alpha_2 (1)(-1)(1\ 1)\begin{pmatrix}2\\3\end{pmatrix} + \alpha_2 \alpha_1 (-1)(1)(2\ 3)\begin{pmatrix}1\\1\end{pmatrix} + \alpha_2^2 \cdot (-1)^2 \cdot (2\ 3)\begin{pmatrix}2\\3\end{pmatrix}\right)$$

$$= \alpha_1 + \alpha_2 - \frac{1}{2}\left(2\alpha_1^2 - 5\alpha_1\alpha_2 - 5\alpha_2\alpha_1 + 13\alpha_2^2\right)$$

$$= \alpha_1 + \alpha_2 - \alpha_1^2 + 5\alpha_1\alpha_2 - \frac{13}{2}\alpha_2^2 \quad \text{with constraint } \alpha_1 = \alpha_2$$

$$= -\alpha_1^2 + 5\alpha_1^2 - \frac{13}{2}\alpha_1^2 + 2\alpha_1$$

$$= \left(\frac{-2 + 10 - 13}{2}\right)\alpha_1^2 + 2\alpha_1$$

$$= -\frac{5}{2}\alpha_1^2 + 2\alpha_1$$

2).

$$\Lambda(\alpha_1,\alpha_2,\lambda) = L(\alpha_1,\alpha_2) + \lambda\left(\sum_{i=1}^{2}\alpha_i T_i\right)$$

$$= \alpha_1+\alpha_2-\alpha_1^2+5\alpha_1\alpha_2-\frac{13}{2}\alpha_2^2 + \lambda\left(\sum_{i=1}^{2}\alpha_i T_i\right)$$

$$\frac{\partial\Lambda}{\partial\alpha_1} = 1-2\alpha_1+5\alpha_2+\lambda = 0 \quad -(i)$$

$$\frac{\partial\Lambda}{\partial\alpha_2} = 1+5\alpha_1-13\alpha_2-\lambda = 0 \quad -(ii)$$

$$\frac{\partial\Lambda}{\partial\lambda} = \sum_{i=1}^{2}\alpha_i T_i = 0 \quad -(iii) \quad \text{(already given as constraint)}$$

$\Rightarrow \alpha_1 = \alpha_2$ from equation (iii)

Substituting in (i) & (ii) and adding (i) & (ii)

$$1-2\alpha_1+5\alpha_1+\lambda \qquad 1+3\alpha_1+\lambda$$

$$1+5\alpha_1-13\alpha_1-\lambda \qquad 1-8\alpha_1-\lambda$$

$$2-5\alpha_1 = 0$$

$$\alpha_1^* = \frac{2}{5} = \alpha_2^* \quad \text{(due to constraint)}$$

3) $$w = \sum_{i=1}^{N}\alpha_i T_i x_i \quad \text{(cf. 100)}$$

$$= \alpha_1 T_1 x_1 + \alpha_2 T_2 x_2 = \frac{2}{5}\cdot 1\cdot\begin{pmatrix}1\\1\end{pmatrix} + \frac{2}{5}\cdot(-1)\cdot\begin{pmatrix}2\\3\end{pmatrix}$$

$$= \begin{pmatrix}2/5\\2/5\end{pmatrix} - \begin{pmatrix}4/5\\6/5\end{pmatrix} = \begin{pmatrix}-2/5\\-4/5\end{pmatrix}$$

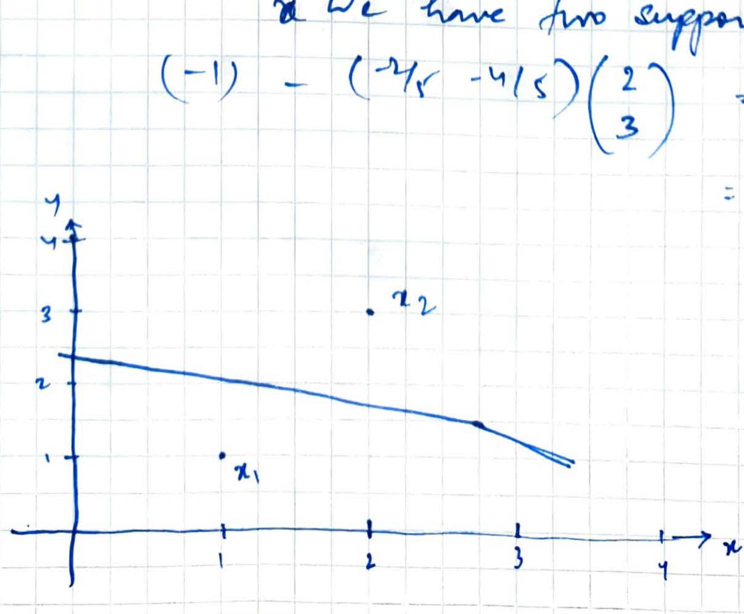
$$w_0 := \frac{1}{T_{i_0}} - w^T x_{i_0} = T_{i_0} - w^T\cdot x_{i_0} \quad \text{(cf. 102)}$$

We have two support vectors, taking one

$$(-1) - (-2/5\ \ -4/5)\begin{pmatrix}2\\3\end{pmatrix} = -1 + \left(\frac{4}{5}+\frac{12}{5}\right)$$

$$= -1 + \frac{16}{5} = 11/5$$

4)



5) $x_3 = (2,2)$

a) Graphically $\rightarrow$ $\omega_{-1}$

b) $f(x) = sig\left(\sum_{i=1}^{N} \alpha_i^* T_i (x_i^T x) + w_0^*\right)$

$= sig\left(\left(\alpha_1^* \cdot (1)\left((2\ 1)\begin{pmatrix}2\\2\end{pmatrix}\right) + \frac{11}{5}\right) + \left(\alpha_2^*(-1)\left((2\ 3)\begin{pmatrix}2\\2\end{pmatrix}\right) + \frac{11}{5}\right)\right)$

$= sig\left(\frac{2}{5} \times 4 - \frac{2 \cdot 10}{5}\right) + \frac{11}{5}$

$sig\left(\frac{8}{5} - \frac{20}{5} + \frac{11}{5}\right)$

$sig\left(\frac{-1}{5}\right) = -1$, So, label is $\omega_{-1}$

6) a) Equation 1 would expand ~~expon~~ polynomially upon simplifying or expanding the summation. In ~~other~~ words, for 2 points, we have 4 ~~parts~~ terms, for 3 inputs we will have 9 terms, and as the points increases we get $N^2$ terms. (Range of summation is from $1 \dots N$)

$$L(w, w_0, \alpha_1, \alpha_2) = \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j T_i T_j x_i^T x_j$$

b) If $\alpha_i \neq 0$, then $x_i$ lies ~~on the~~ at the margin and thus contributes to the separation line. That is why these vectors are also called support vectors (c.f. 100 in the script)