Exercise 1

We have $L(\theta, \theta_0, \lambda) = \frac{1}{2}\theta^T\theta - \sum_{i=1}^{N} \lambda_i \left[y_i (\theta^T x_i + \theta_0) - 1 \right]$ (1)

Taking the partial derivatives and set them to 0

$$\frac{\partial L(\theta, \theta_0, \lambda)}{\partial \theta} = \theta - \sum_{i=1}^{N} \lambda_i y_i x_i = 0 \Rightarrow \boxed{\theta = \sum_{i=1}^{N} \lambda_i y_i x_i} \quad (2)$$

$$\frac{\partial L(\theta, \theta_0, \lambda)}{\partial \theta_0} = -\sum_{i=1}^{N} \lambda_i y_i = 0 \Rightarrow \boxed{\sum \lambda_i y_i = 0} \quad (3)$$

Substituting (2), (3) into (1)

$$L(\theta, \theta_0, \lambda) = \frac{1}{2} \sum_{i=1}^{N} \lambda_i y_i x_i \sum_{j=1}^{N} \lambda_j y_j x_j - \sum_{i=1}^{N} \lambda_i y_i x_i \sum_{j=1}^{N} \lambda_j y_j x_j - \cancel{\sum_{i=1}^{N} \lambda_i y_i \theta_0}^{\,(3)\ 0} + \sum_{i=1}^{N} \lambda_i =$$

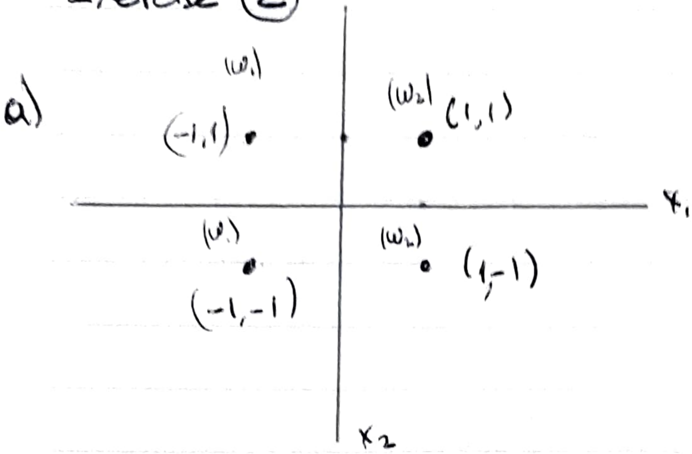
$$= \sum \lambda_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \lambda_i \lambda_j y_i y_j x_i \cdot x_j$$

subject to $\sum \lambda_i y_i = 0$

$\lambda \geq 0$

Exercise 2

a)



The linear svm classifier shoud return the line $x_1 = 0$

b) $J(\lambda) = \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 - \frac{1}{2}\big[\lambda_1\lambda_1 \cdot 2 + \cancel{\lambda_1\lambda_2 \cdot 0} + \lambda_1\lambda_3 \cdot 2 - \cancel{\lambda_1\lambda_4 0}$
$+ \cancel{\lambda_2\lambda_1 0} + \lambda_2\lambda_2 \cdot 2 + \cancel{\lambda_2\lambda_3 0} + \lambda_2\lambda_4 2$
$+ \lambda_3\lambda_1(-1)(-2) + \cancel{\lambda_3\lambda_2(-1)0} + \lambda_3\lambda_3(1)(2) + \cancel{\lambda_3\lambda_4(1)0}$
$+ \cancel{\lambda_4\lambda_1(-1)(0)} + \lambda_4\lambda_2(-1)(-2) + \cancel{\lambda_4\lambda_3(1)(0)} + \lambda_4\lambda_4(1)(2)\big]$

$\Rightarrow J(\lambda) = \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 - \lambda_1^2 - \lambda_2^2 - \lambda_3^2 - \lambda_4^2 - 2\lambda_1\lambda_3 - 2\lambda_2\lambda_4$

Taking the partial derivatives w.r.t. $\lambda_1, \lambda_2, \lambda_3, \lambda_4$

$$\frac{\partial J(\lambda)}{\partial \lambda_1} = 1 - 2\lambda_1 - 2\lambda_3 = 0 \Rightarrow \lambda_1 = \frac{1 - 2\lambda_3}{2}$$

$$\frac{\partial J(\lambda)}{\partial \lambda_2} = 1 - 2\lambda_2 - 2\lambda_4 = 0 \Rightarrow \lambda_2 = \frac{1 - 2\lambda_4}{2}$$

$$\frac{\partial J(\lambda)}{\partial \lambda_3} = 1 - 2\lambda_3 - 2\lambda_1 = 0 \Rightarrow \lambda_3 = \frac{1 - 2\lambda_1}{2}$$

$$\frac{\partial J(\lambda)}{\partial \lambda_4} = 1 - 2\lambda_4 - 2\lambda_2 = 0 \Rightarrow \lambda_4 = \frac{1 - 2\lambda_2}{2}$$

We have that

$$\theta = \sum_{i} \alpha_i y_i x_i$$

$$\Rightarrow \theta = \frac{1-2\lambda_3}{2}\begin{bmatrix}-1\\1\end{bmatrix} + \lambda_2\begin{bmatrix}-1\\-1\end{bmatrix} - \lambda_3\begin{bmatrix}1\\-1\end{bmatrix} - \frac{1-2\lambda_2}{2}\begin{bmatrix}1\\1\end{bmatrix} =$$

$$= \begin{bmatrix}\frac{2\lambda_3-1}{2}\\ \frac{1-2\lambda_3}{2}\end{bmatrix} + \begin{bmatrix}-\lambda_2\\-\lambda_2\end{bmatrix} - \begin{bmatrix}\lambda_3\\-\lambda_3\end{bmatrix} - \begin{bmatrix}\frac{1-2\lambda_2}{2}\\ \frac{1-2\lambda_2}{2}\end{bmatrix}$$

So $\theta_1 = \frac{2\lambda_3-1}{2} - \lambda_2 - \lambda_3 - \frac{1-2\lambda}{2} = \frac{\cancel{2\lambda_3} - 1 - \cancel{2\lambda_2} - \cancel{2\lambda_3} - 1 + \cancel{2\lambda_2}}{2} = -1$

$\theta_2 = \frac{1-2\lambda_3}{2} - \lambda_2 + \lambda_3 \; \frac{-1+2\lambda_2}{2} = \frac{\cancel{1} - \cancel{2\lambda_3} - \cancel{2\lambda_2} + \cancel{2\lambda_3} \cancel{-1} + \cancel{2\lambda_2}}{2} = 0$

so $\theta = [-1 \;\; 0]^T$

In order to find θ_0 we can say

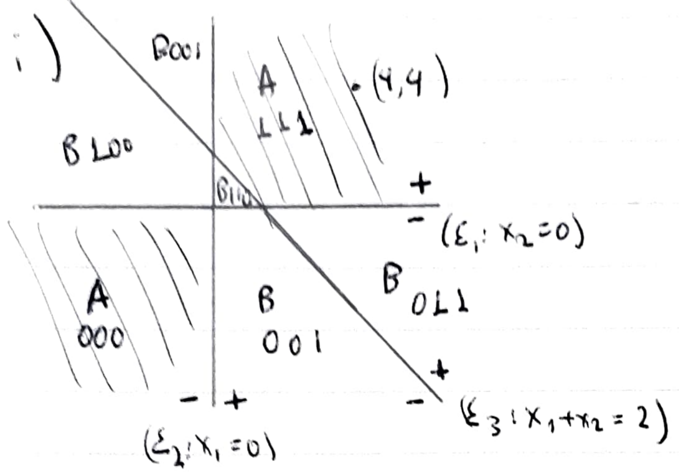
$\theta^T x + \theta_0 = 0$ for $x = [1 \;\; 1]^T$

$[-1 \;\; 0]\begin{bmatrix}1\\1\end{bmatrix} + \theta_0 = 0 \Rightarrow \theta_0 = 0$

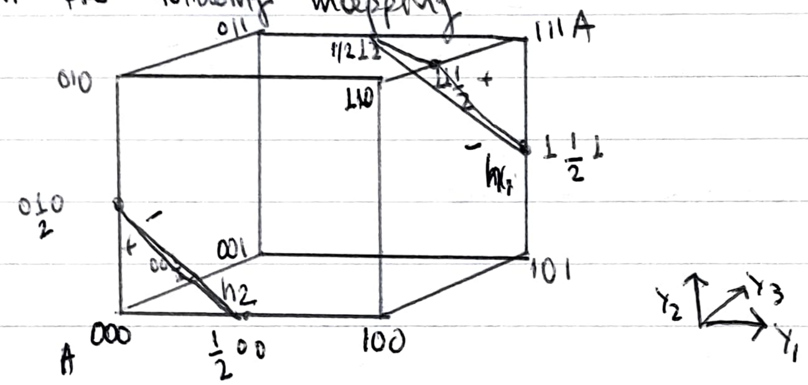
so our hyperplane is indeed the line $x_1 = 0$

Exercise ③

i)



ii) Based on the lines above, we can transform the space to R^3 with the following mapping



Even in our transformed space, we cannot separate class A from class B using a linear classifier. We will create 2 hyperplanes and leave the datapoints that belong to class A on their positive side (drawn above)

We know that the equation of a plane that passes through the points (x_{11}, x_{12}, x_{13}), (x_{21}, x_{22}, x_{23}), (x_{31}, x_{32}, x_{33}) is:

$$\begin{vmatrix} x_1 & x_2 & x_3 & 1 \\ x_{11} & x_{12} & x_{13} & 1 \\ x_{21} & x_{22} & x_{23} & 1 \\ x_{31} & x_{32} & x_{33} & 1 \end{vmatrix} = 0$$

- For the first hyperplane we have

$$\begin{vmatrix} x_1 & x_2 & x_3 & 1 \\ 1/2 & 0 & 0 & 1 \\ 0 & 1/2 & 0 & 1 \\ 0 & 0 & 1/2 & 1 \end{vmatrix} = 0$$

$$\Rightarrow x_1 \begin{bmatrix} 0 & 0 & 1 \\ 1/2 & 0 & 1 \\ 0 & 1/2 & 1 \end{bmatrix} - x_2 \begin{vmatrix} 1/2 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 1/2 & 1 \end{vmatrix} + x_3 \begin{bmatrix} 1/2 & 0 & 1 \\ 0 & 1/2 & 1 \\ 0 & 0 & 1 \end{bmatrix} - 1 \begin{vmatrix} 1/2 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1/2 \end{vmatrix} = 0$$

$$\frac{1}{4}x_1 + \frac{1}{4}x_2 + \frac{1}{4}x_3 - \frac{1}{8} = 0 \Rightarrow \boxed{x_1 + x_2 + x_3 - \frac{1}{2} = 0}$$

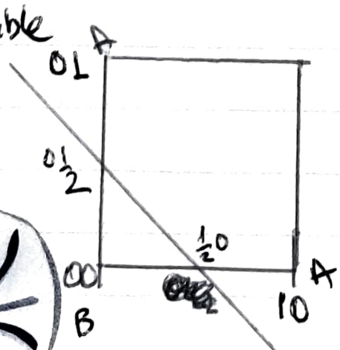
- For the second hyperplane we have

$$\begin{vmatrix} x_1 & x_2 & x_3 & 1 \\ 1 & 1 & 1/2 & 1 \\ 1 & 1/2 & 1 & 1 \\ 1/2 & 1 & 1 & 1 \end{vmatrix} = 0$$

$$\Rightarrow x_1 \begin{vmatrix} 1 & 1/2 & 1 \\ 1/2 & 1 & 1 \\ 1 & 1 & 1 \end{vmatrix} - x_2 \begin{vmatrix} 1 & 1/2 & 1 \\ 1 & 1 & 1 \\ 1/2 & 1 & 1 \end{vmatrix} + x_3 \begin{vmatrix} 1 & 1 & 1 \\ 1 & 1/2 & 1 \\ 1/2 & 1 & 1 \end{vmatrix} - 1 \begin{vmatrix} 1 & 1 & 1/2 \\ 1 & 1/2 & 1 \\ 1/2 & 1 & 1 \end{vmatrix}$$

$$-\frac{1}{4}x_1 - \frac{1}{4}x_2 - \frac{1}{4}x_3 + \frac{5}{8} = 0 \Rightarrow x_1 + x_2 + x_3 - \frac{5}{2} = 0$$

Finally in our tranformed space the classes are lineary separable



The equation of the line is

$$x_1 + x_2 - \frac{1}{2} = 0$$

Our neural network can be depicted as: