Homework 1

1. Consider the following four data points:
   Class 1: $\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}$
   Class 2: $\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}$

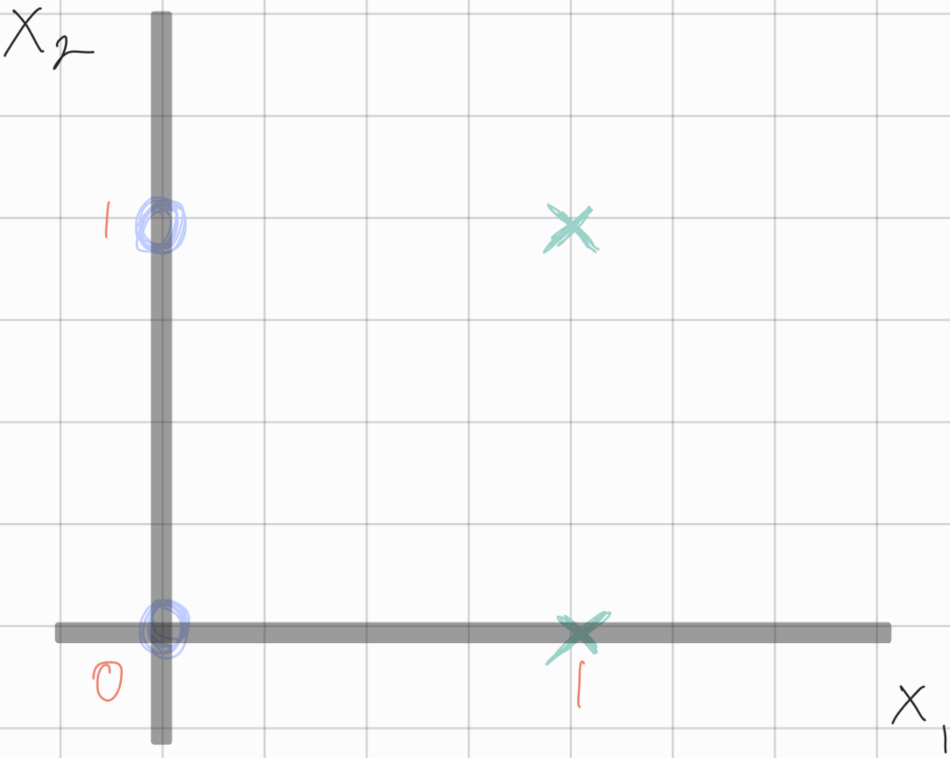
   (a) Are they linearly separable?
   (b) Find the linear decision boundary by applying fixed-increment single-sample perception procedure with initial weight vector **0** and plot it in the feature space.
   (c) Repeat (b) by using MSE procedure.
   (d) Implement the Ho-Kashyap algorithm in MATLAB and repeat (b).
   (e) Repeat (a) through (d) with the following data points:
   Class 1: $\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix}$
   Class 2: $\begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}$
   (f) Comment on the three procedures used.

2. Suppose that the following are a set of points in two classes:
   Class 1: $\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}1\\2\end{pmatrix}, \begin{pmatrix}2\\1\end{pmatrix}$
   Class 2: $\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}$

   Plot them and find the optimal separating line. What are the support vectors and what is the margin?

do e & f

1)



a) They are linearly seperable

b)

$$Y = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ -1 & -1 & -1 \\ -1 & -1 & 0 \end{bmatrix} \begin{matrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{matrix} \begin{matrix} \left.\begin{matrix} \\ \\ \end{matrix}\right\} \text{class 1} \\ \left.\begin{matrix} \\ \\ \end{matrix}\right\} \text{class 2} \end{matrix}$$

$$a = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}^T$$

$$Y_1 \cdot a = 0 \overset{?}{>} 0 \qquad \text{No}$$

$$a_{new} = a_{old} + (Y_1)^T$$

$$a_{new} = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^T$$

$$Y_2 \cdot a = \overset{1\cdot 1}{1} + \overset{0\cdot 0}{0} + \overset{1\cdot 0}{0} = 1 > 0 \qquad \text{yes}$$

$$Y_3 \cdot a = \overset{-1\cdot 1}{-1} + \overset{-1\cdot 0}{0} + \overset{-1\cdot 0}{0} = -1 \overset{?}{>} 0 \quad \text{No}$$

$$a_{new} = a_{old} + (Y_3)^T$$

$$a_{new} = \begin{bmatrix} 0 & -1 & -1 \end{bmatrix}$$

$$-1\cdot 0 + -1\cdot -1 + 0\cdot -1$$

$$Y_4 \cdot a = 0 \quad 1 \quad 0 = 1 \overset{?}{>} 0 \quad \text{yes}$$

$$Y_1 \cdot a = \overset{1\cdot 0}{0} + \overset{0\cdot -1}{0} + \overset{0\cdot -1}{0} = 0 \overset{?}{>} 0 \quad NO$$

$$a_{new} = a_{old} + (Y_1)^T$$

$$a_{new} = [\,1 \quad -1 \quad -1\,]^T$$

$$Y_2 \cdot a = 1\cdot 1 + 0\cdot -1 + 1\cdot -1$$

$$1 + 0 + -1$$

$$0 \overset{?}{>} 0 \qquad NO$$

$$a_{new} = a_{old} + (Y_2)^T$$

$$a_{new} = [\,2 \quad -1 \quad 0\,]$$

$$Y_3 \cdot a = -1\cdot 2 + -1\cdot -1 + -1\cdot 0$$

$$-2 + 1 + 0$$

$$-1 \overset{?}{>} 0 \qquad No$$

$$a_{new} = a_{old} + (Y_3)^T$$

$$a_{new} = [\,1 \quad -2 \quad -1\,]$$

$$Y_4 \cdot a = \underset{-1}{-1\cdot 1} + \underset{2}{-1\cdot -2} + \underset{0}{0\cdot -1} = 1 \overset{?}{>} 0 \qquad yes$$

$$Y_1 \cdot a = 1\cdot 1 + 0\cdot -2 + 0\cdot -1 = 1 \overset{?}{>} 0 \qquad yes$$

$Y_2 \cdot a = \overset{1}{1\cdot 1} + \overset{0}{0\cdot -2} + \overset{-1}{1\cdot -1} = 0 \overset{?}{>} 0$ No

$a_{new} = a_{old} + (Y_2)^T$

$a_{new} = [2 \quad -2 \quad 0]$

$Y_3 \cdot a = \overset{-2}{-1\cdot 2} + \overset{2}{-1\cdot -2} + \overset{0}{-1\cdot 0} = 0 \overset{?}{>} 0$ No

$a_{new} = a_{old} + (Y_3)^T$

$a_{new} = [1 \quad -3 \quad -1]$

$Y_4 \cdot a = \overset{-1}{-1\cdot 1} + \overset{3}{-1\cdot -3} + \overset{0}{0\cdot -1} = 2 > 0$ yes

---

$Y_1 \cdot a = 1\cdot 1 + 0\cdot -3 + 0\cdot -1 = 1 > 0$ yes

$Y_2 \cdot a = \overset{1}{1\cdot 1} + \overset{0}{0\cdot -3} + \overset{-1}{1\cdot -1} = 0 > 0$ No

$a_{new} = a_{old} + (Y_2)^T$

$a_{new} = [2 \quad -3 \quad 0]$

$Y_3 \cdot a = \overset{-2}{-1\cdot 2} + \overset{3}{-1\cdot -3} + \overset{0}{-1\cdot 0} = 1 > 0$ yes

---

$Y_4 \cdot a = \overset{-2}{-1\cdot 2} + \overset{3}{-1\cdot -3} + \overset{0}{0\cdot -1} = 1 \overset{?}{>} 0$ yes

---

$Y_1 \cdot a = 1\cdot 2 + 0\cdot -3 + 0\cdot -1 = 2 \overset{?}{>} 0$ yes

---

$Y_2 \cdot a = 1\cdot 2 + 0\cdot -3 + 1\cdot 0 = 2 \overset{?}{>} 0$ yes

Final: $a = [2 \quad -3 \quad 0]^T$

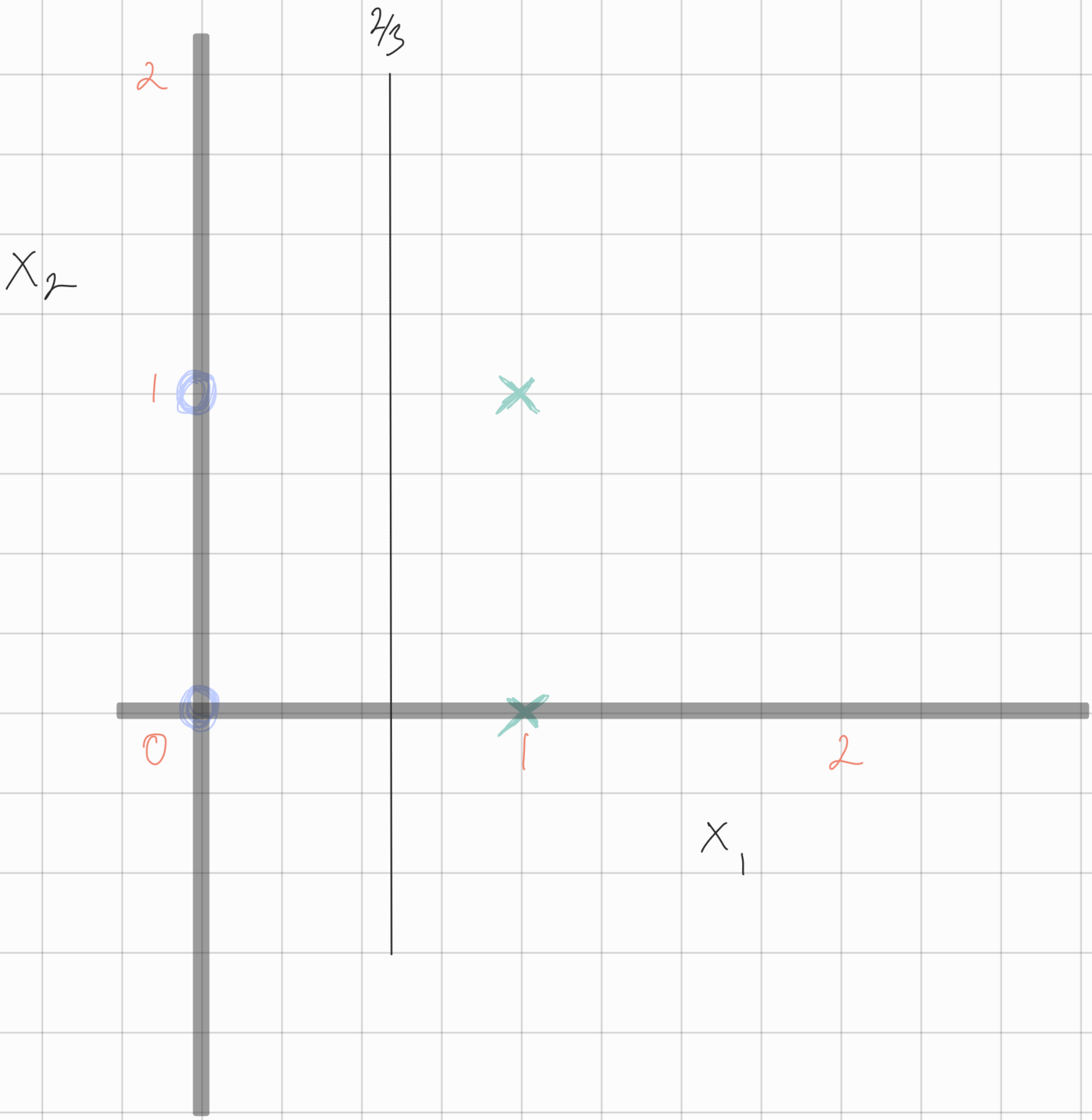
when $g(x) = 0$

$$0 = w_0 + w_1 x_1 + w_2 x_2$$

$$2 + -3x_1 + 0x_2$$

$$-2 = -3x_1$$

$$2/3 = x_1$$



c) <u>MSE</u>

$$\underline{Y} \cdot \underline{a} = \underline{b} \qquad J_s(a) = \| Ya - b \|^2$$
$$= \sum_{i=1}^{n} (a^T y_i - b_i)^2$$

$$b = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}^T \qquad \nabla J_s(a) = 2Y^T(Ya-b) = 0$$
$$\rightarrow \quad a = (Y^T Y)^{-1} Y^T b$$

$$Y = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ -1 & -1 & -1 \\ -1 & -1 & 0 \end{bmatrix} \qquad Y^T = \begin{bmatrix} 1 & 1 & -1 & -1 \\ 0 & 0 & -1 & -1 \\ 0 & 1 & -1 & 0 \end{bmatrix}$$

$$Y^T Y = \begin{bmatrix} 4 & 2 & 2 \\ 2 & 2 & 1 \\ 2 & 1 & 2 \end{bmatrix} \rightarrow \quad (Y^T Y)^{-1} = \begin{bmatrix} 0.75 & -0.5 & -0.5 \\ -0.5 & 1 & 0 \\ -0.5 & 0 & 1 \end{bmatrix}$$

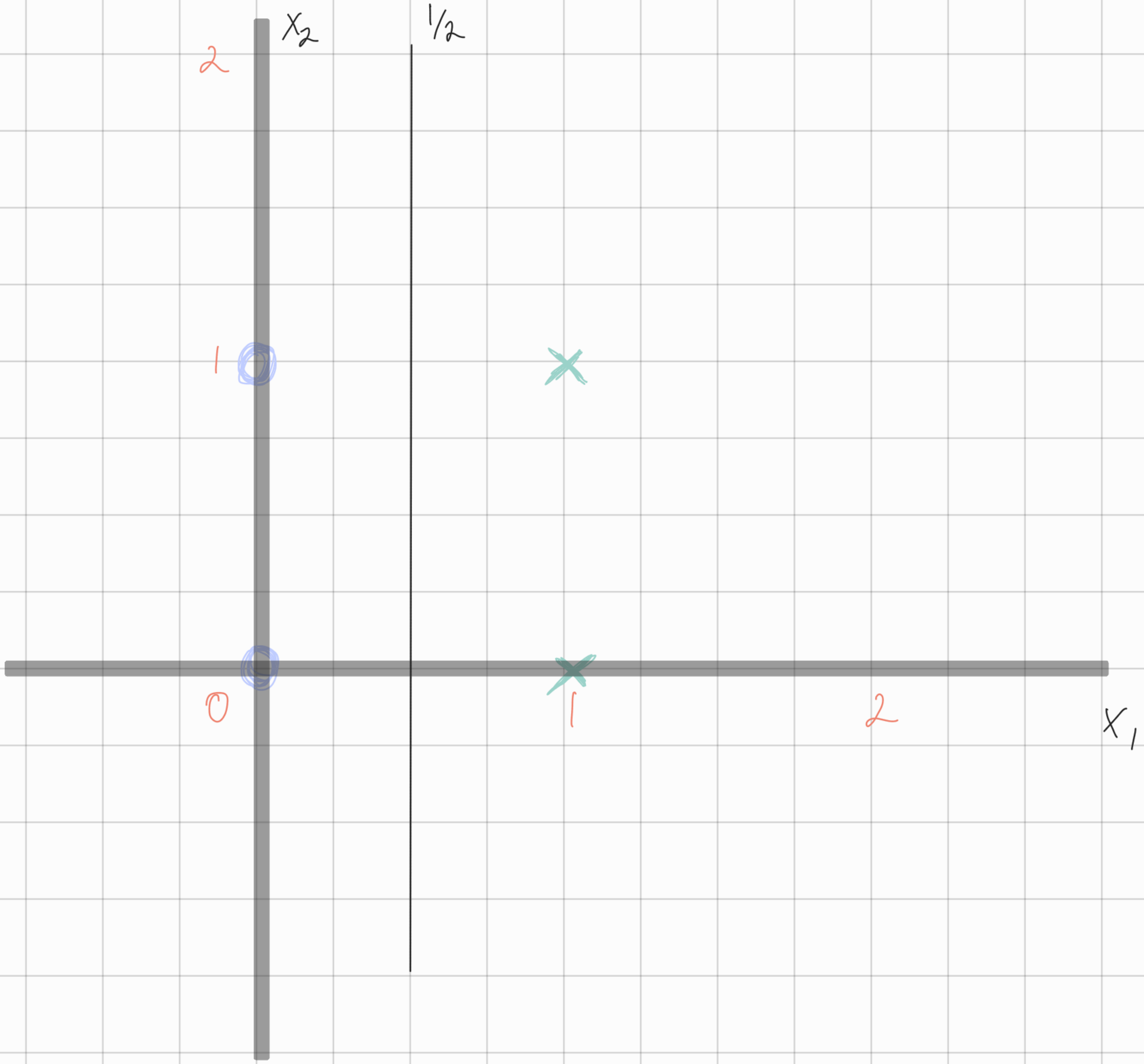
$$(Y^T Y)^{-1} Y^T = \begin{bmatrix} 0.75 & 0.25 & 0.25 & -0.25 \\ -0.5 & -0.5 & -0.5 & -0.5 \\ -0.5 & 0.5 & -0.5 & 0.5 \end{bmatrix}$$

$$a = (Y^T Y)^{-1} Y^T b = \begin{bmatrix} 1 & -2 & 0 \end{bmatrix}^T$$

$$0 = 1 - 2x_1 + 0x_2$$

$$-1 = 2x_1$$

$$1/2 = x_1$$



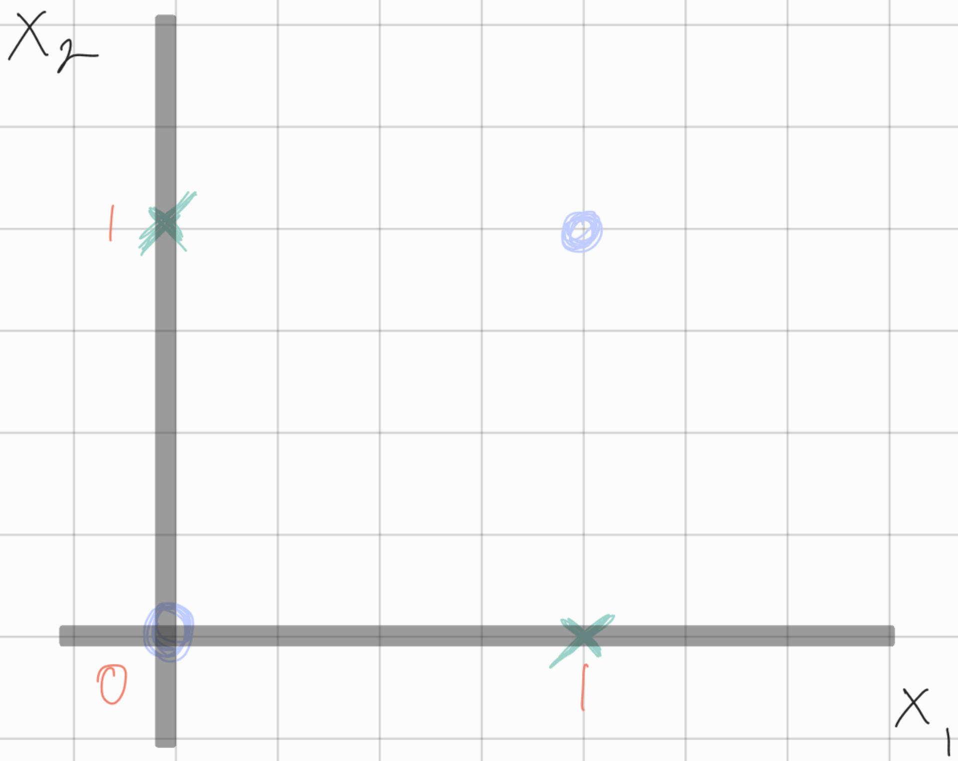
d) done in matlab

final a vector was the

same as in MSE

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

e)



a) they are not linearly seperable

b)

$$Y = \begin{array}{c} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ -1 & 0 & -1 \\ -1 & -1 & 0 \end{bmatrix} \end{array} \begin{array}{l} \left.\begin{array}{l} y_1 \\ y_2 \end{array}\right\} \text{class 1} \\ \left.\begin{array}{l} y_3 \\ y_4 \end{array}\right\} \text{class 2} \end{array}$$

$$a = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}^T$$

$$y_1 \cdot a = 1\cdot 0 + 0\cdot 0 + 0\cdot 0 = 0 \overset{?}{>} 0 \quad \text{NO}$$

$$a_{new} = a_{old} + (y_1)^T = [1 \quad 0 \quad 0]$$

$$y_2 \cdot a = 1\cdot 1 + 1\cdot 0 + 1\cdot 0 = 1 \overset{?}{>} 0 \quad \text{yes}$$

$$y_3 \cdot a = -1\cdot 1 + 0\cdot 0 + -1\cdot 0 = -1 \overset{?}{>} 0 \quad \text{NO}$$

$$a_{new} = a_{old} + (y_3)^T = [0 \quad 0 \quad -1]$$

$$y_4 \cdot a = -1\cdot 0 \quad -1\cdot 0 \quad 0\cdot -1 = 0 \overset{?}{>} 0 \quad \text{NO}$$

$$a_{new} = a_{old} + (y_4)^T = [-1 \quad -1 \quad -1]$$

$$y_1 \cdot a = 1\cdot -1 \quad 0\cdot -1 \quad 0\cdot -1 = 0 \overset{?}{>} 0 \quad \text{No}$$

$$a_{new} = a_{old} + (y_1)^T = [0 \quad -1 \quad -1]$$

$$y_2 \cdot a = 1\cdot 0 + 1\cdot -1 + 1\cdot -1 = -2 \overset{?}{>} 0 \quad \text{NO}$$

$$a_{new} = a_{old} + (y_2)^T = [1 \quad 0 \quad 0]$$

Loop found, will not terminate without iteration cap or decreasing learning rate.

MSE

$\underline{Y} \cdot \underline{a} = \underline{b}$ $\qquad J_s(a) = \| Ya - b \|^2$

$$= \sum_{i=1}^{n} (a^T y_i - b_i)^2$$

$b = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}^T$ $\qquad \nabla J_s(a) = 2Y^T(Ya - b) = 0$

$$\rightarrow a = (Y^T Y)^{-1} Y^T b$$

$$Y = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ -1 & 0 & -1 \\ -1 & -1 & 0 \end{bmatrix} \qquad Y^T = \begin{bmatrix} 1 & 1 & -1 & -1 \\ 0 & 1 & 0 & -1 \\ 0 & 1 & -1 & 0 \end{bmatrix}$$

$$Y^T Y = \begin{bmatrix} 4 & 2 & 2 \\ 2 & 2 & 1 \\ 2 & 1 & 2 \end{bmatrix} \rightarrow (Y^T Y)^{-1} = \begin{bmatrix} 0.75 & -0.5 & -0.5 \\ -0.5 & 1 & 0 \\ -0.5 & 0 & 1 \end{bmatrix}$$

$$(Y^T Y)^{-1} Y^T = \begin{bmatrix} 0.75 & -0.25 & -0.25 & -0.25 \\ -0.5 & 0.5 & 0.5 & -0.5 \\ -0.5 & 0.5 & -0.5 & 0.5 \end{bmatrix}$$

$$a = (Y^T Y)^{-1} Y^T b = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}^T$$

d) done in matlab

final a vector was the

same as in MSE

$a = [0 \quad 0 \quad 0]^T$

f) single-sample perceptron was most labor by hand & on matlab. Gave acceptable solution when possible

MSE was quicker because it was an analytical solution rather than iterative method. If all computation for pseudo inverse had to be done by hand, might be a different story.

Ho-Kashyap was least effort & simplest in understanding after getting through previous two. don't have to worry about selecting b values because it will get optimized

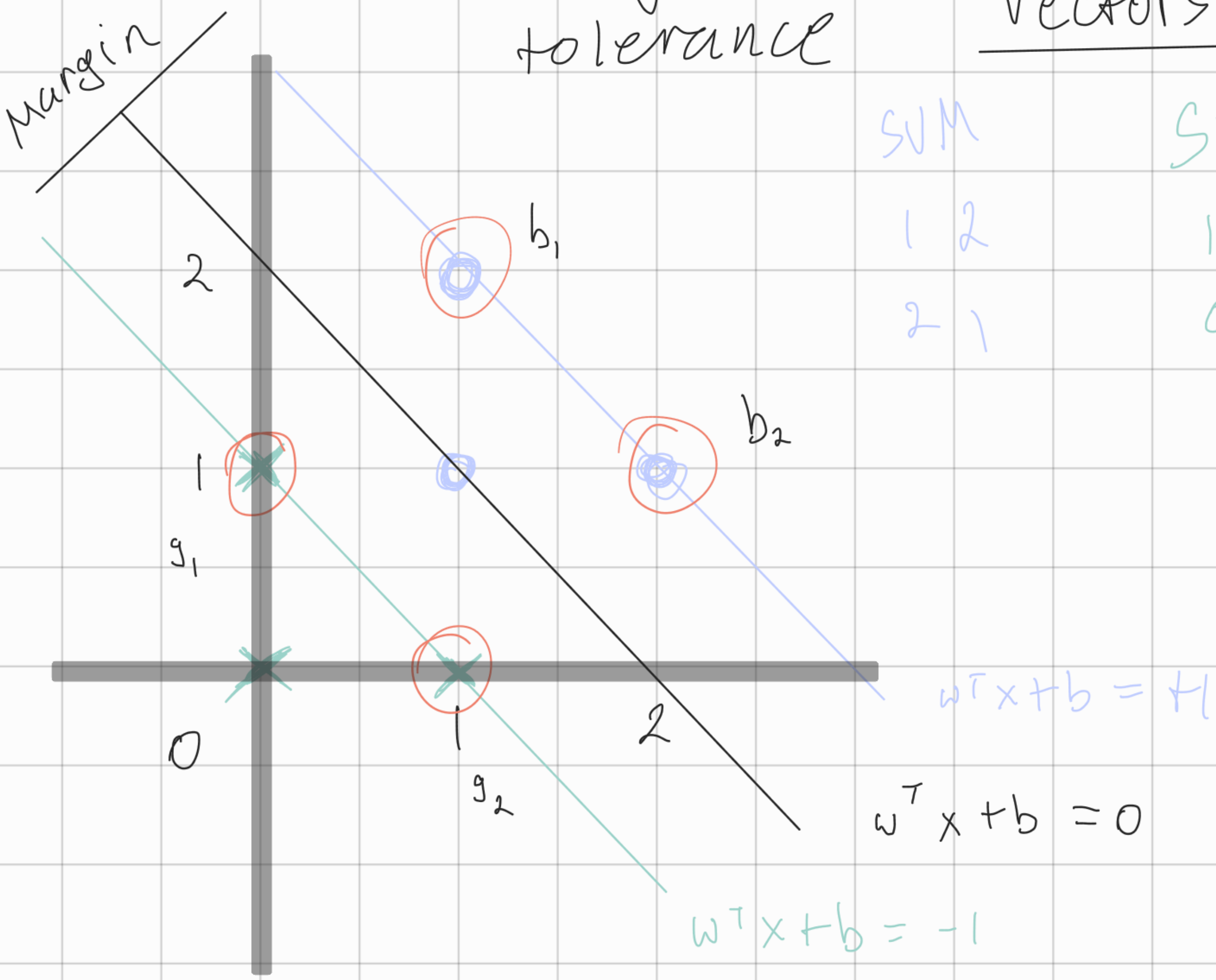
2)

Class 1: $\binom{1}{1}, \binom{1}{2}, \binom{2}{1}$

Class 2: $\binom{0}{0}, \binom{1}{0}, \binom{0}{1}$

$$Y = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \\ 1 & 2 & 1 \\ -1 & 0 & 0 \\ -1 & -1 & 0 \\ -1 & 0 & -1 \end{bmatrix}$$

Assuming some tolerance



Vectors

SVM		SVM	
1	2	1	0
2	1	0	1

distance(b1, g2)

$d = 2 - 0 = 2$

$y_2 - y_1 = m(x_2 - x_1)$

$m = \frac{1-2}{2-1} = \frac{-1}{1} = -1$

$y = -(x-1) + 2$

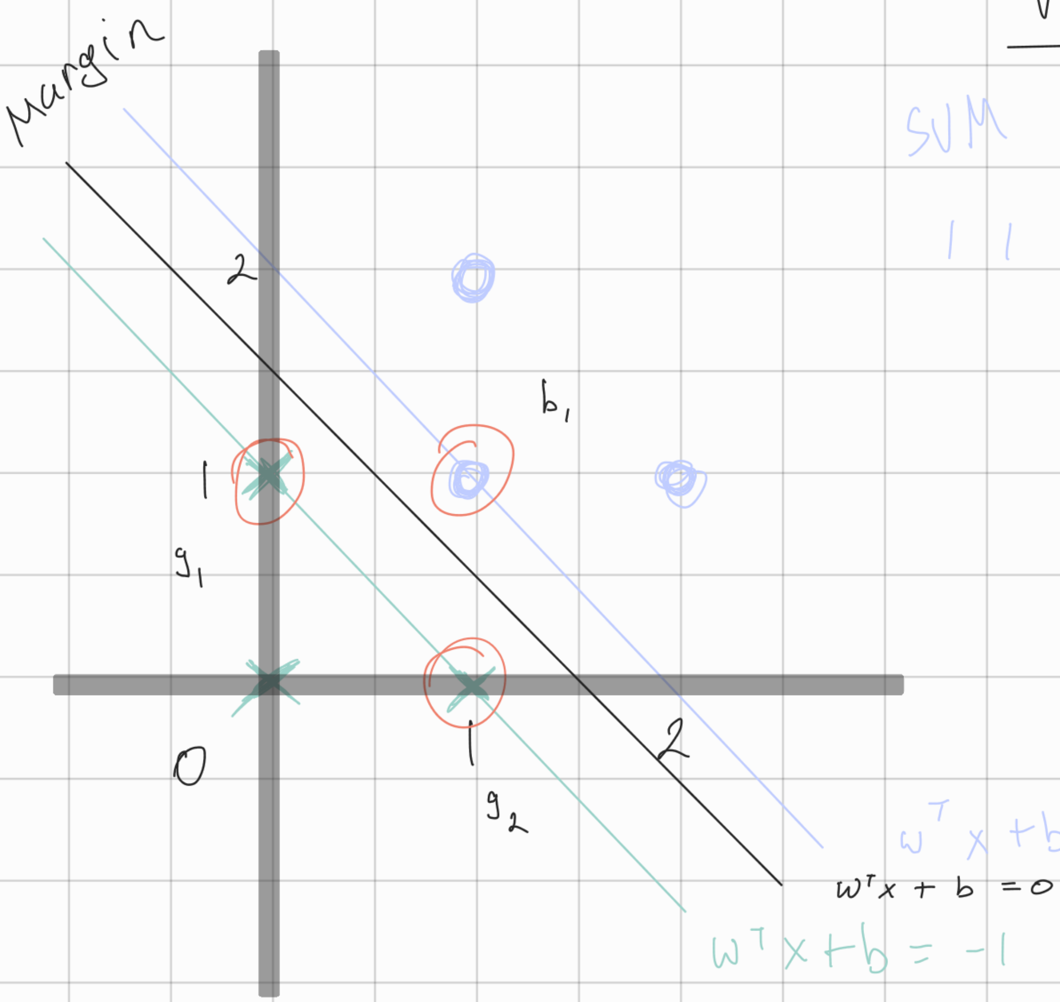
$-x + 1 + 2$

$y = -x + 3$

$0 = w^T x + 3$

$\frac{2}{\|-1\|} = 2 = \text{margin}$

No tolerance



Vectors

SVM		SVM	
1	1	1	0
		0	1