3. Here we explore the maximal margin classifier on a toy data set.

(a) We are given $n = 7$ observations in $p = 2$ dimensions. For each observation, there is an associated class label.

Obs.	X_1	X_2	Y
1	3	4	Red
2	2	2	Red
3	4	4	Red
4	1	4	Red
5	2	1	Blue
6	4	3	Blue
7	4	1	Blue

Sketch the observations.

(b) Sketch the optimal separating hyperplane, and provide the equation for this hyperplane (of the form (9.1)).

(c) Describe the classification rule for the maximal margin classifier. It should be something along the lines of "Classify to Red if $\beta_0 + \beta_1 X_1 + \beta_2 X_2 > 0$, and classify to Blue otherwise." Provide the values for β_0, β_1, and β_2.

(d) On your sketch, indicate the margin for the maximal margin hyperplane.

(e) Indicate the support vectors for the maximal margin classifier.

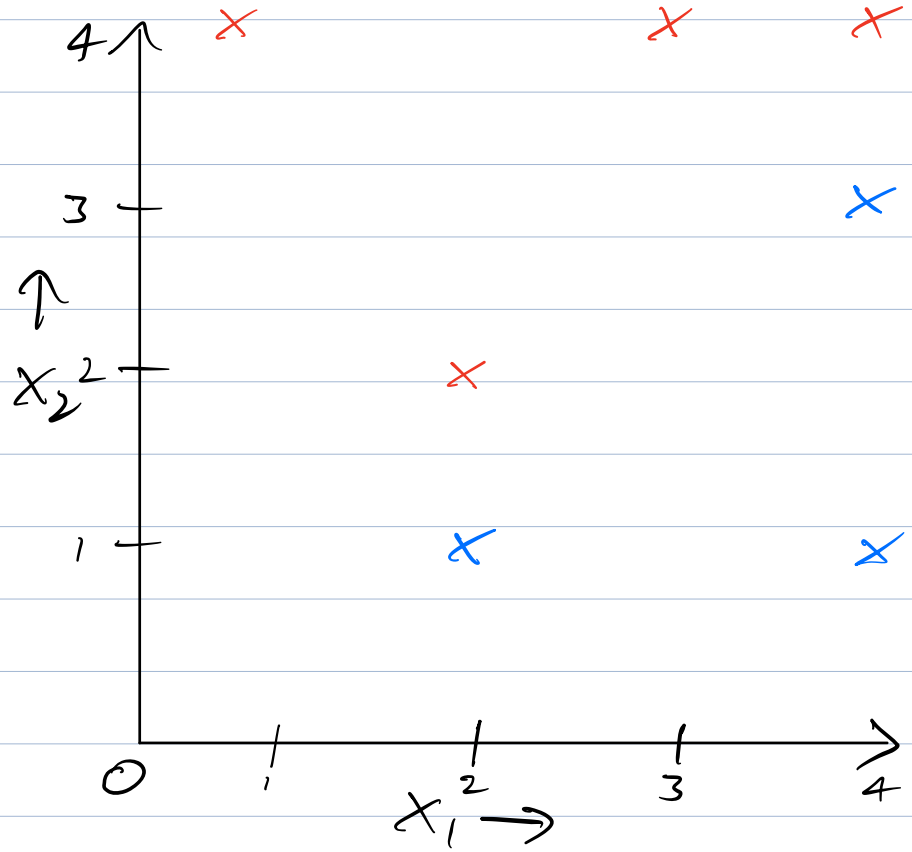
(f) Argue that a slight movement of the seventh observation would not affect the maximal margin hyperplane.

(g) Sketch a hyperplane that is *not* the optimal separating hyperplane, and provide the equation for this hyperplane.

(h) Draw an additional observation on the plot so that the two classes are no longer separable by a hyperplane.

Answer

a)



b) The optimal separating hyper plane lies between points (2, 1), (4, 3), (2, 2), (4, 4)

∴ The hyper plane passes through points

$$(2, 1.5),\ (4, 3.5)$$

$$\therefore \text{slope} = \frac{3.5 - 1.5}{4 - 2} = \frac{2}{2} = 1$$

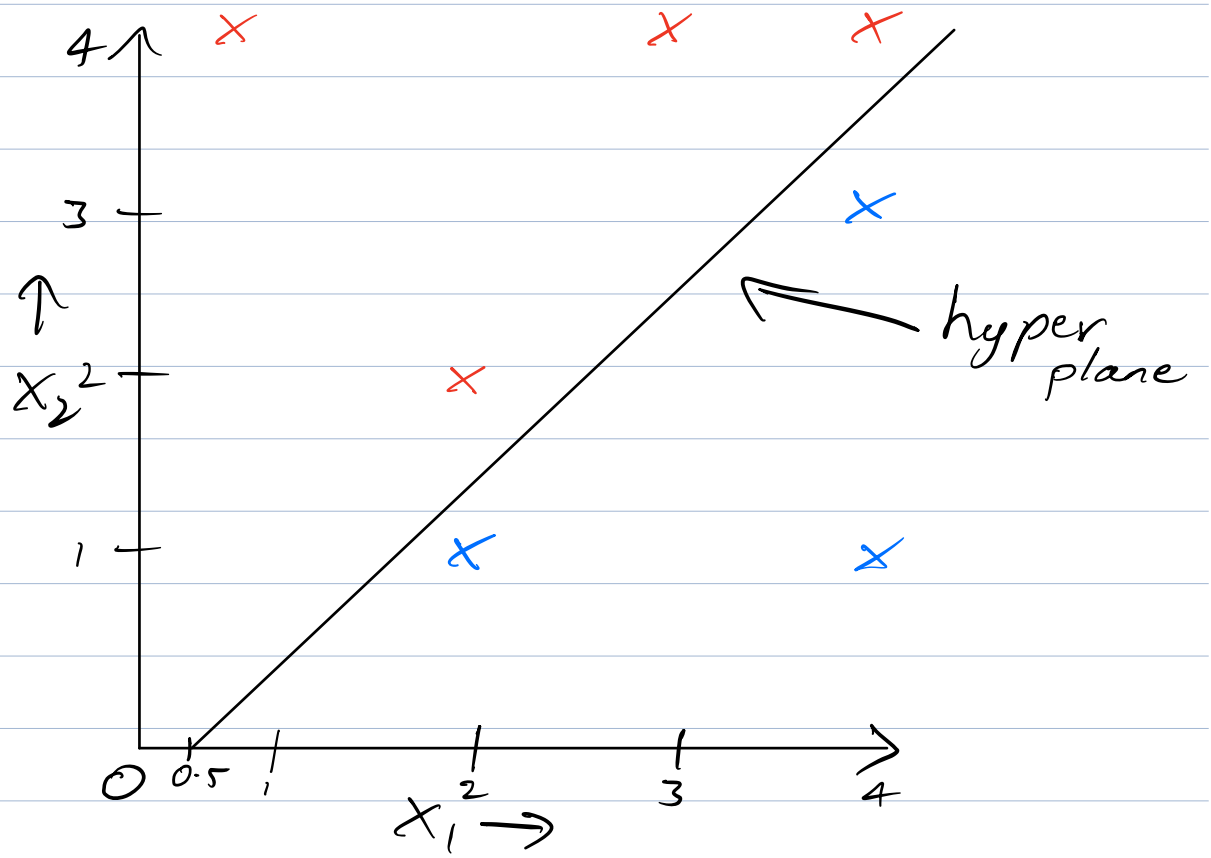
$y = mx + c$ | $(2, 1.5)$ (x, y)

$1.5 = 1(2) + c$

$\therefore \; c = 1.5 - 2 = -0.5$

$\therefore$ Intercept $= -0.5$

Equation of hyperplane: $x_2 = x_1 - 0.5$



c) $\beta_0 + \beta_1 x_1 + \beta_2 x_2 > 0$

But the hyperplane is

$$0.5 - x_1 + x_2 = 0$$

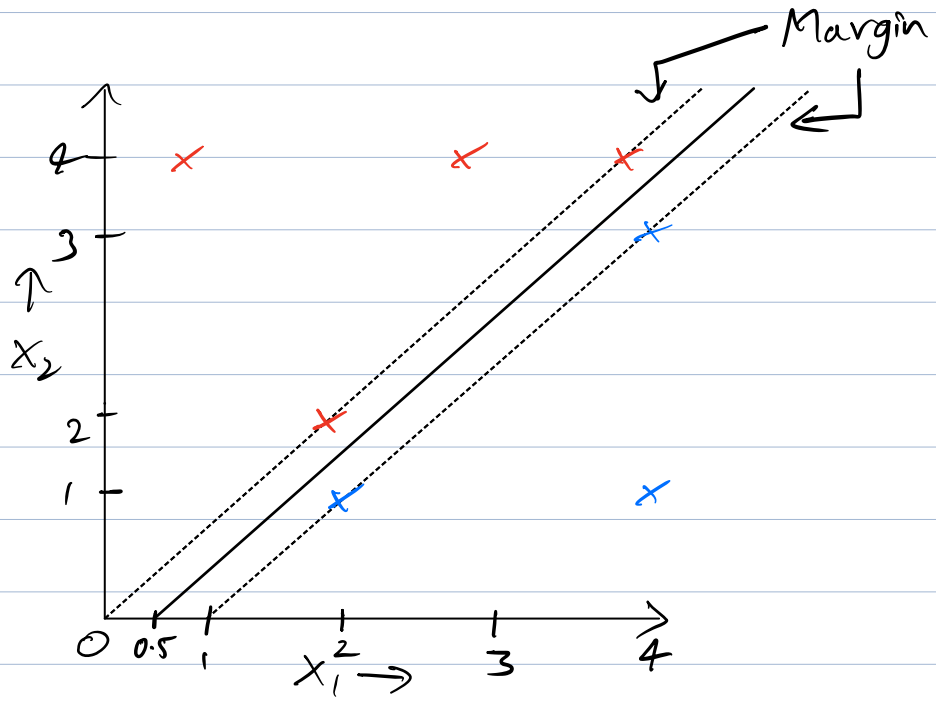
∴ Classification Rules

Classify to Red if $0.5 - x_1 + x_2 > 0$

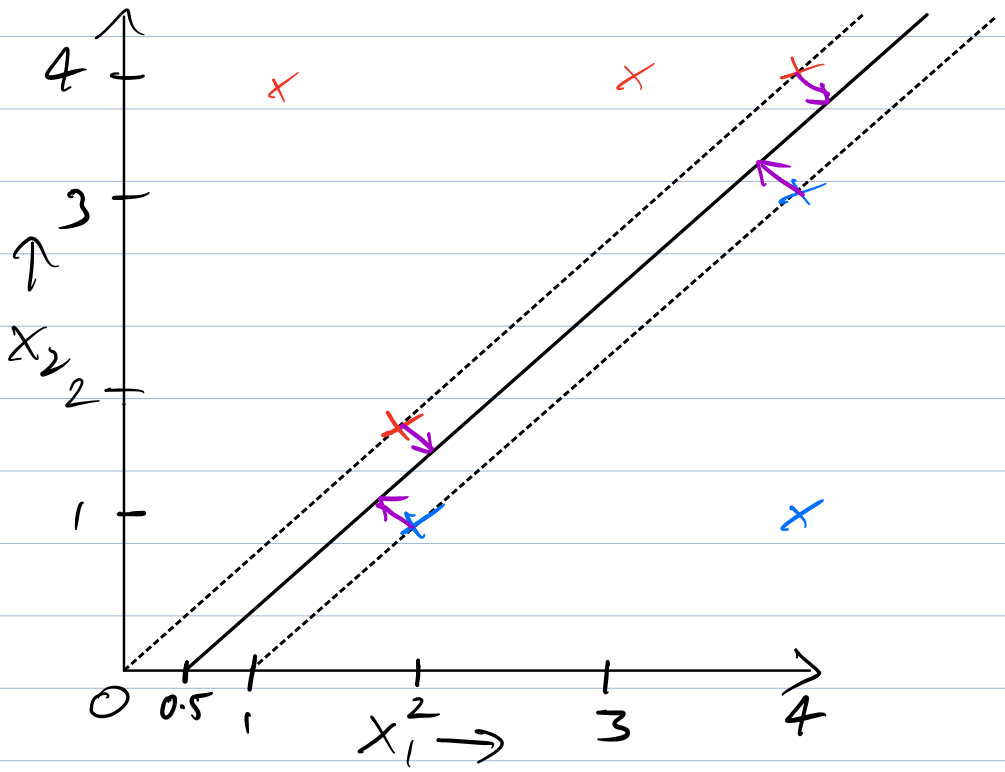
Classify to Blue if $0.5 - x_1 + x_2 < 0$

$(\beta_0 = 0.5, \beta_1 = -1, \beta_2 = 1)$

d)



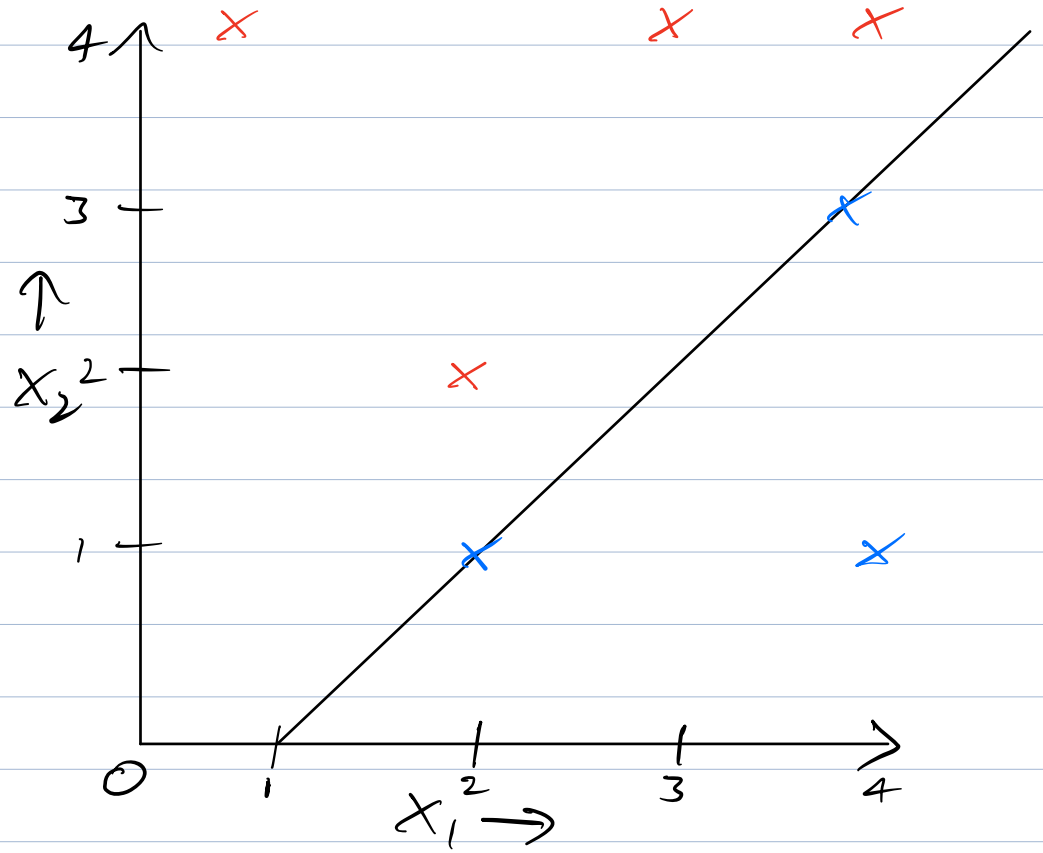
e)



The purple lines denote the support vectors.

f) A slight movement of the seventh observation (4, 1) would not have any effect on the maximal margin hyperplane since the point is far from any of the hyperplane margins, and its movement would always be outside the margin.

g)



The hyperplane that lies on the points (2,1) and (4,3) is not optimal.

$$m = \frac{3-1}{4-2} = \frac{2}{2} = 1$$

$y = mx + c$ | (2,1)

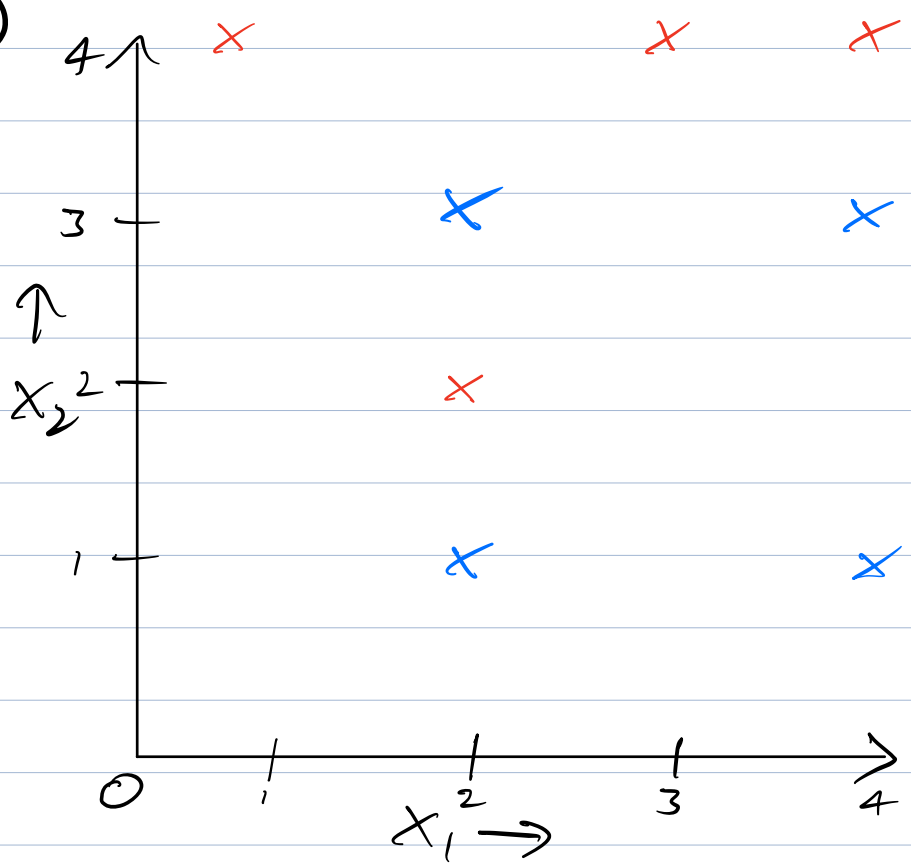
$\Rightarrow 1 = 2 + c \Rightarrow c = 1 - 2 = -1$

∴ Equation of hyperplane:

$$y = x - 1$$

This hyperplane is not optimal.

h)



The 2 classes are no longer separable.