An Introduction to Statistical Learning: Chapter 9

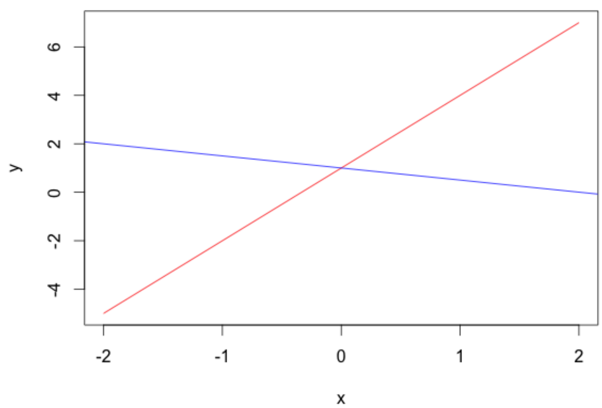
Due on January 25,2016 at 3:10pm

G. James et.al.

Tony Jiang

1. This problem involves hyperplanes in two dimensions.
(a). Sketch the hyperplane $1 + 3X_1 - X_2 = 0$. Indicate the set of points for which $1 + 3X_1 - X_2 > 0$, as well as the set of points for which $1 + 3X_1 - X_2 < 0$.
(b). On the same plot, sketch the hyperplane $-2 + X_1 + 2X_2 = 0$. Indicate the set of points for which $-2 + X_1 + 2X_2 > 0.$, as well as the set of points for which $-2 + X_1 + 2X_2 < 0$.

Answers:



Red line= $1 + 3X_1 - X_2 = 0$; Blue line=$-2 + X_1 + 2X_2 = 0$. (a)
set of points for $1 + 3X_1 - X_2 > 0$: (0,2)
set of points for $1 + 3X_1 - X_2 > 0$: (0,-1)
(b)
set of points for$-2 + X_1 + 2X_2 > 0.$: (0,2)
set of points for$-2 + X_1 + 2X_2 < 0.$: (0,-2)

2. We have seen that in $p = 2$ dimensions, a linear decision boundary takes the form $\beta_0 + \beta_1 X1 + \beta_2 X_2 = 0$. We now investigate a non-linear decision boundary.
(a) Sketch the curve

$$(1 + X_1)^2 + (2 - X_2)^2 = 4.$$

(b) On your sketch, indicate the set of points for which

$$(1 + X_1)^2 + (2 - X_2)^2 > 4.$$

as well as the set of points for which

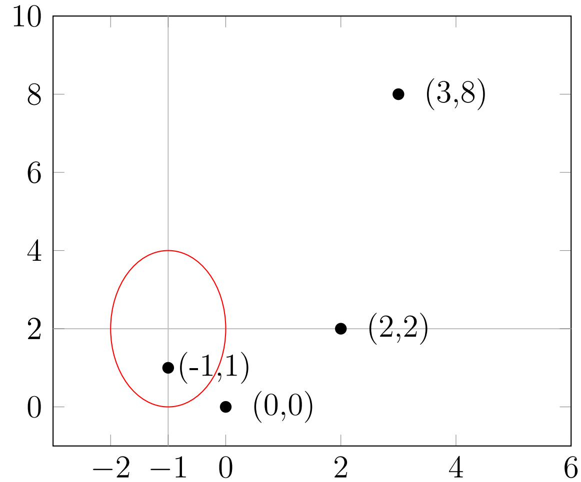
$$(1 + X_1)^2 + (2 - X_2)^2 \leq 4.$$

(c) Suppose that a classifier assigns an observation to the blue class if

$$(1 + X_1)^2 + (2 - X_2)^2 > 4.$$

and to the red class otherwise. To what class is the observation $(0,0)$classified? $(-1,1)$?$(2,2)$?$(3,8)$?
(d) Argue that while the decision boundary in (c) is not linear in terms of X_1 and X_2, it is linear in terms of X_1, X_1^2, X_2, and X_2^2.

Answers:
(a).
It would be an eclipse entering on (-1,2) long axis=2, short axis=1



(b). point inside the eclipse satisfy $(1+X_1)^2 + (2-X_2)^2 \le 4$.
point outside the eclipse satisfy $(1+X_1)^2 + (2-X_2)^2 > 4$.
(c)
$(0,0)$ is outside the eclipse $->$ blue class
$(-1,1)$ is inside the eclipse $->$red class
$(2,2)$ is outside the eclipse $->$ blue class
$(3,8)$ is outside the eclipse $->$ blue class

(d)
This can be easily shown by expanding the terms

$$\begin{aligned} f(x) &= (1+X_1)^2 + (2-X_2)^2 - 4 \\ &= X_1^2 + 2X_1 + X_2^2 - 4X_2 + 1 \end{aligned}$$

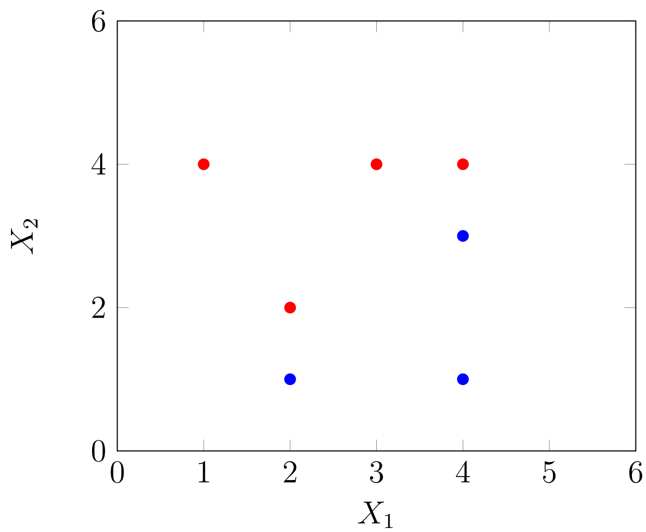
it is a linear combination of the terms X_1, X_1^2, X_2, and X_2^2
3. Here we explore the maximal margin classifier on a toy data set.
(a) We are given $n = 7$ observation in $p = 2$ dimensions. For each observation, there is an associated class label.

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

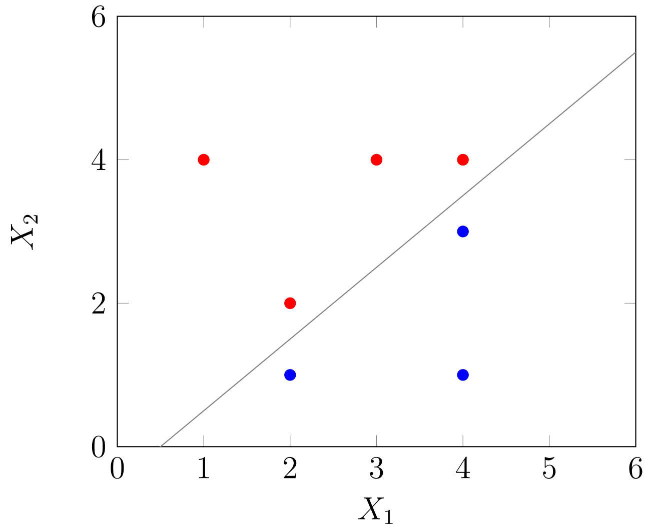
Obs.	X_1	X_2	Y
1	3	4	Red
2	2	2	Red
3	4	4	Red
4	1	4	Red
5	2	1	Blue
6	4	3	Blue
7	4	1	Blue

(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.

Answers:
(a).



(b)

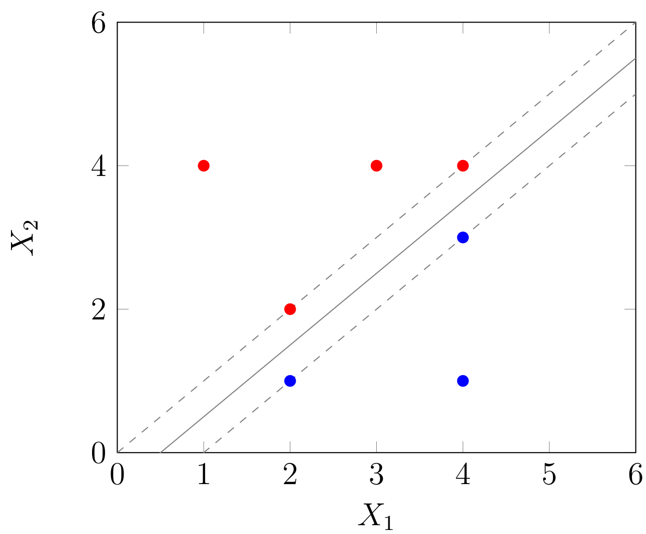


(c)

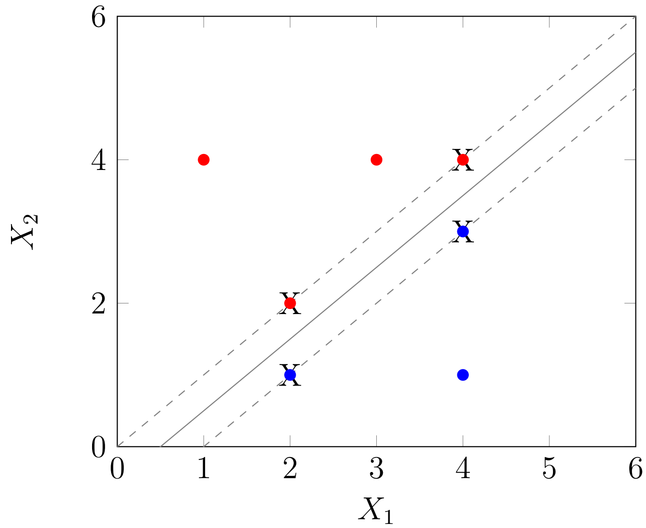
Classify to Red if $-X_1 + X_2 + 0.5 > 0$, and classify to Blue otherwise. $\beta_0 = 05, \beta_1 = -1$ and $\beta_2 = 1$

(d)

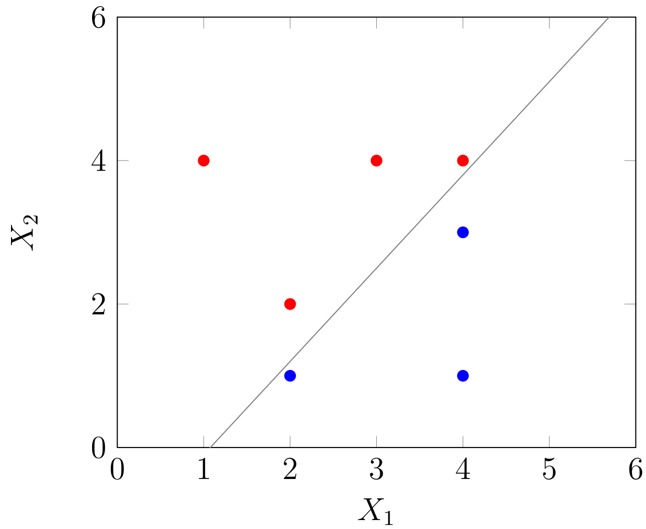
see dashed lines



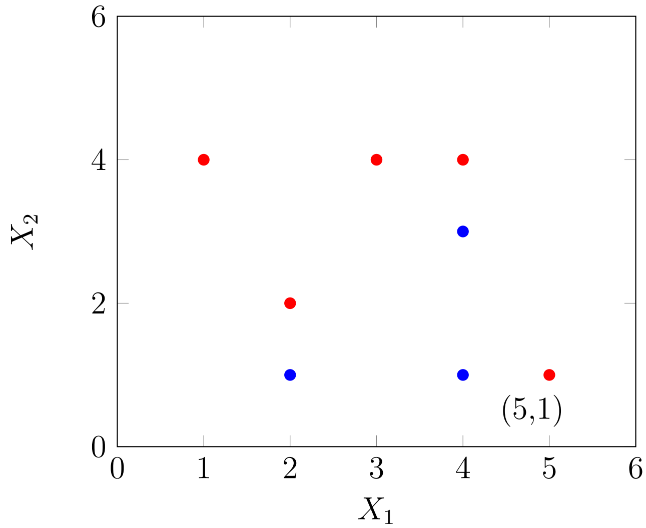
(e)



(f)
7th observation is not a support vector,so a slight movement (so slight that it still is on the down side of the boundary to be classified as blue) will not affect the maximal margin hyperplane.
(g)



$1.4 + 1.3X_1 - X_2 = 0$. (h)


6
4
2
0
0
1
2
3
4
5
6
X_1
X_2
(5,1)