Assignment 2 Machine Learning COS4852

Adriaan Louw (53031377)

July 2, 2018

Contents

1 Question 1 **1**
1.1 Question 1(a) 1
1.2 Question 1(b) 3
1.3 Question 1(c) 5
1.4 Question 1(d) 5

2 Question 2 **5**
2.1 Question 2(a) 5
2.2 Question 2(b) 6
2.3 Question 2(c) 6

3 Question 3 **7**
3.1 Question 3(a) 7
3.2 Question 3(b) 7
3.3 Question 3(c) 7
3.4 Question 3(d) 7
3.5 Question 3(e) 7

1 Question 1

1.1 Question 1(a)

Firstly we calculate the line

$$x_2 = mx_1 + c \tag{1}$$

for the intersect points (2,0) and (0,6).
Calculating slope m,

$$\begin{aligned} m &= \frac{6-0}{0-2} \\ &= -3 \end{aligned} \tag{2}$$

x_2 intercept c is 6.
This makes equation 1

$$x_2 = -3x_1 + 6 \tag{3}$$

Nils J Nilsson (1998) gives the equation for the hyperplane as

$$\sum_{i=1}^{n} x_i\omega_i \geq \theta \tag{4}$$

which in this case gives the equation for the hyperplane to be

$$\omega_1 x_1 + \omega_2 x_2 + \omega_3 = 0 \tag{5}$$

We need to get equation 5 in the form of equation 1

$$\begin{aligned} \omega_1 x_1 + \omega_2 x_2 + +\omega_3 = 0 \\ \omega_2 x_2 &= -\omega_1 x_1 - \omega_3 \\ x_2 &= -\frac{\omega_1 x_1}{\omega_2} - \frac{\omega_3}{\omega_2} \end{aligned} \tag{6}$$

Comparing coefficients m and c from equation 3 to 6 we get

$$\begin{aligned} -\frac{\omega_1}{\omega_2} &= -3 \\ \omega_1 &= 3\omega_2 \end{aligned} \tag{7}$$

and

$$\begin{aligned} -\frac{\omega_3}{\omega_2} &= 6 \\ \omega_3 &= -6\omega_2 \end{aligned} \tag{8}$$

If we choose $\omega_3 = -2$ then $\omega_1 = 1$ and $\omega_2 = \frac{1}{3}$. This makes the hyperplane equation from equation 5

$$x_1 + \frac{x_2}{3} - 2 = 0 \tag{9}$$

Now we need to test this hyperplane. For positive instance (2,6)

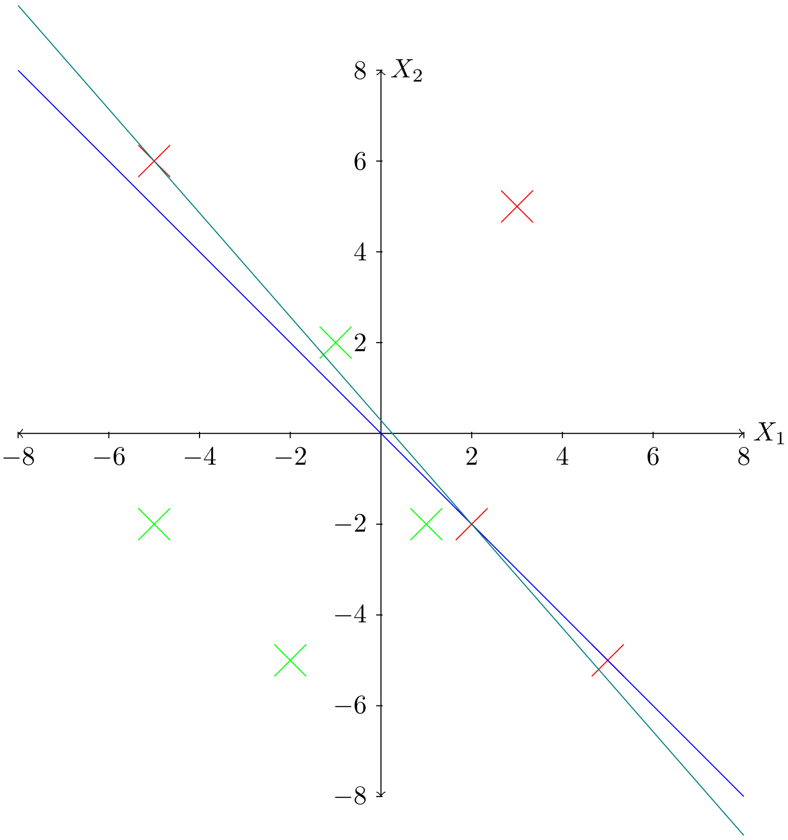
$$\begin{aligned} x_1 + \frac{x_2}{3} - 2 &= \\ 2 + \frac{6}{2} - 2 &= \\ 2 \end{aligned} \tag{10}$$

Which is as expected.
And the negative instance (-1,2)

$$\begin{aligned} x_1 + \frac{x_2}{3} - 2 &= \\ -1 + \frac{2}{3} - 2 &= \\ -\frac{7}{3} \end{aligned} \tag{11}$$

This is also as expected. The perceptron now classifies the the data correctly

1.2 Question 1(b)



From the above image we can see that any that it is not possible to create a hyperplane that correctly classifies all negative instances and positive instances. The blue line is the line $x_2 = -x_1$ and the teal line is the line $x_2 = -\frac{7}{8}x_1 + \frac{2}{7}$. The any minimum plane that correctly classifies all the negative instances will classify the positive instance (-1,2) incorrectly as negative.

We can create a hyperplane from regression from all the points close to where the hyperplane should be. Using negative points (-5,6),(2,-2),(5,-5) and positive points (-1,2),(1,-2)

For the equation of the regressed line $x_2 = mx_1 + c$

$$\begin{aligned} m &= r\frac{S(x_2)}{S(x_1)} \\ m &= \frac{\sum((x_1 - \bar{x_1})(x_2 - \bar{x_2}))}{\sqrt{\sum(x_1 - \bar{x_1})^2 \sum(x_2 - \bar{x_2})^2}} \frac{\sqrt{\frac{\sum(x_2 - \bar{x_2})^2}{n-1}}}{\sqrt{\frac{\sum(x_1 - \bar{x_1})^2}{n-1}}} \end{aligned} \tag{12}$$

where r is Pearsons Correlation Coefficient and S is standard deviation of axis x_2 or x_1.

Here follows the calculation

x_1	x_2	$x_1 - \bar{x_1}$	$x_2 - \bar{x_2}$	$(x_1 - \bar{x_1})(x_2 - \bar{x_2})$	$(x_1 - \bar{x_1})^2$	$(x_2 - \bar{x_2})^2$
-5	6	-5.4	6.2	-33.48	29.16	38.44
2	-2	1.6	-1.8	-2.88	2.56	3.24
5	-5	4.6	-4.8	-22.08	21.16	23.04
-1	2	-1.4	2.2	-3.08	1.96	4.84
1	-2	0.6	-1.8	-1.08	0.36	3.24

From the above table we have $\bar{x_1} = 0.4, \bar{x_2} = -0.2, \sum((x_1 - \bar{x_1})(x_2 - \bar{x_2})) = -62.6, (x_1 - \bar{x_1})^2 = 55.2, and (x_2 - \bar{x_2})^2 = 72.8$

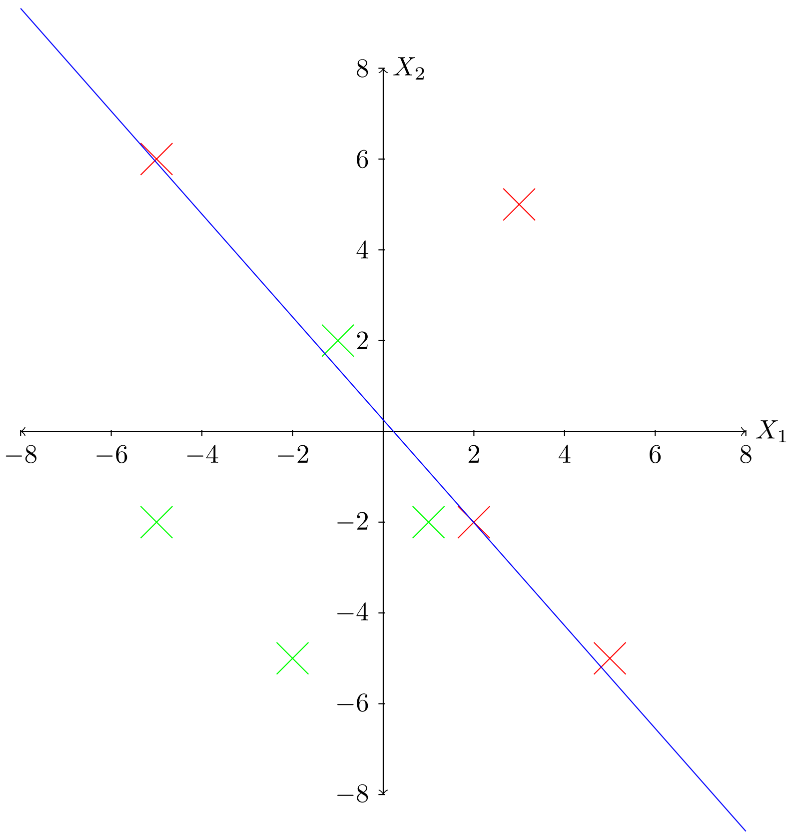
Passing these into equation 12 we get m = -1.13

$$\begin{aligned} c &= \bar{x_2} - m\bar{x_1} \\ &= -0.2 - (-1.13)(0.4) \\ &= 0.25 \end{aligned} \tag{13}$$

Which gives us equation

$$x_2 = -1.13x_1 + 0.25 \tag{14}$$

Visually it would be



This will minimise the error even though it would incorrectly classify (-1,2)

1.3 Question 1(c)

1.4 Question 1(d)

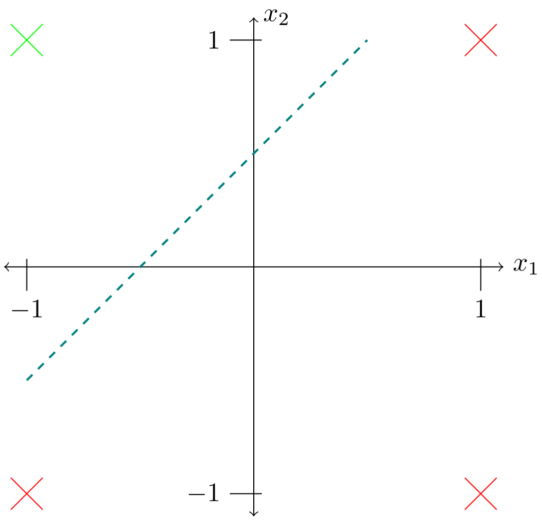
2 Question 2

2.1 Question 2(a)

The truth table for this function is

x_1	x_2	$\neg x_1$	$\neg x_1 \vee x_2$
-1	-1	1	-1
-1	1	1	1
1	-1	-1	-1
1	1	-1	-1

Visually this is:



The dashed line represents the function

$$x_2 = x_1 + 0.5 \tag{15}$$

This hyperplane will classify the boolean function correctly because it separates all the positive instances from the negative ones.

The equation for the hyperplane with weights are given by equation 5 and we know from equation 6 what the relationship is from the weight equation to the line equation. Thus for case

$$\begin{aligned} m &= -\frac{\omega_1}{\omega_2} \\ 1 &= -\frac{\omega_1}{\omega_2} \\ \omega_2 &= -\omega_1 \end{aligned} \tag{16}$$

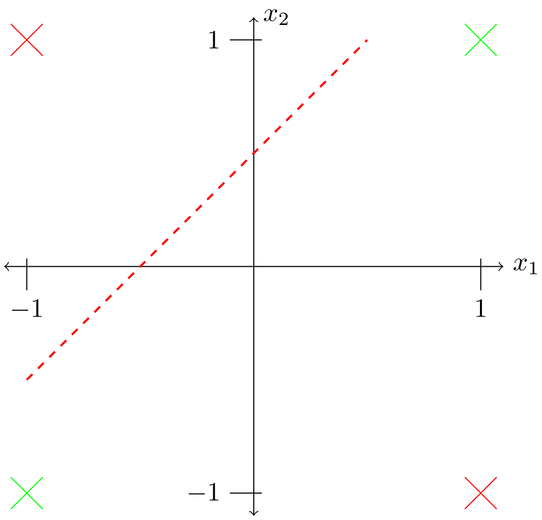
and

$$
\begin{aligned}
c &= -\frac{\omega_0}{\omega_2} \\
0.5 &= -\frac{\omega_0}{\omega_2} \\
\omega_0 &= -\frac{\omega_2}{2}
\end{aligned}
\tag{17}
$$

2.2 Question 2(b)

x_1	x_2	$x_1 \bigoplus x_2$	$\neg x_1 \bigoplus x_2$
-1	-1	-1	1
-1	1	1	-1
1	-1	1	-1
1	1	-1	1

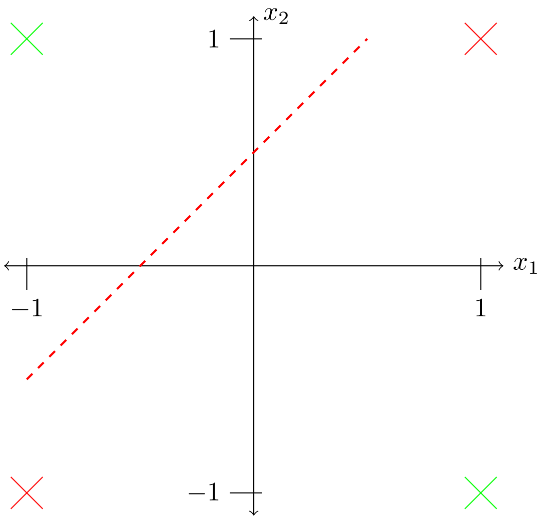
Visually this is:



2.3 Question 2(c)

x_1	x_2	$\neg x_1$	$\neg x_2$	$x_1 \vee x_2$	$\neg x_1 \vee \neg x_2$	$(x_1 \vee x_2) \wedge (\neg x_1 \vee \neg x_2)$
-1	-1	1	1	-1	1	-1
-1	1	1	-1	1	1	1
1	-1	-1	1	1	1	1
1	1	-1	-1	1	-1	-1

Visually this is:



3 Question 3

3.1 Question 3(a)

3.2 Question 3(b)

3.3 Question 3(c)

3.4 Question 3(d)

3.5 Question 3(e)

References

Nils J Nilsson. (1998). *Introduction to Machine Learning.* Retrieved from http://robotics.stanford.edu/people/nilsson/MLBOOK.pdf