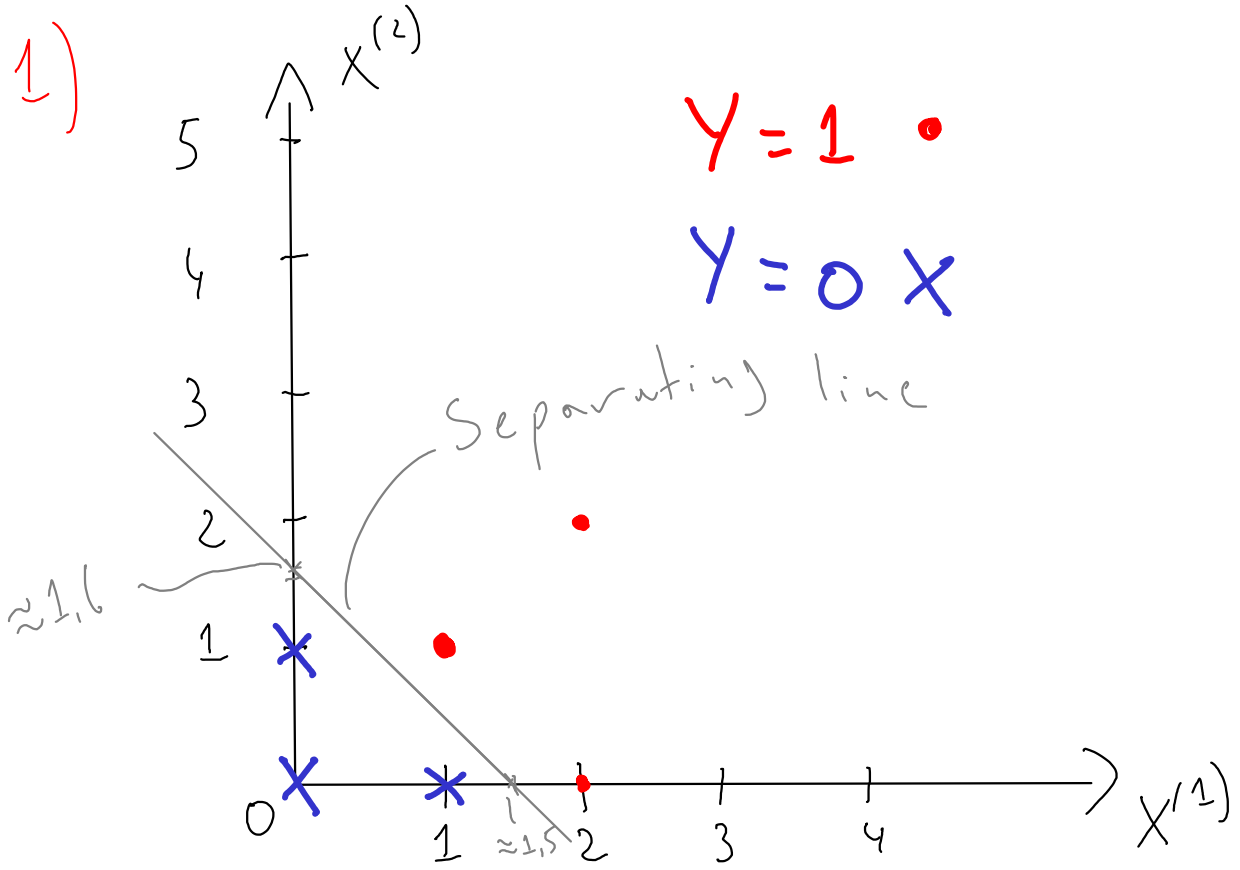
# Introduction to AI



## Assignment 3: Logistic regression

## Solutions to part A

A) 1)




⇒ Visually the classes are Separable.

2) Equation of the separating line

$$\beta_2 x^{(2)} + \beta_1 x^{(1)} + \beta_0 = 0$$

$$\Downarrow$$

$$x^{(2)} = \underbrace{-\frac{\beta_1}{\beta_2}}_{a} x^{(1)} \underbrace{- \frac{\beta_0}{\beta_2}}_{b}$$

$$a = -\frac{19,64}{18,06} \approx -1,08 \qquad b = \frac{28,55}{18,06} \approx 1,58$$

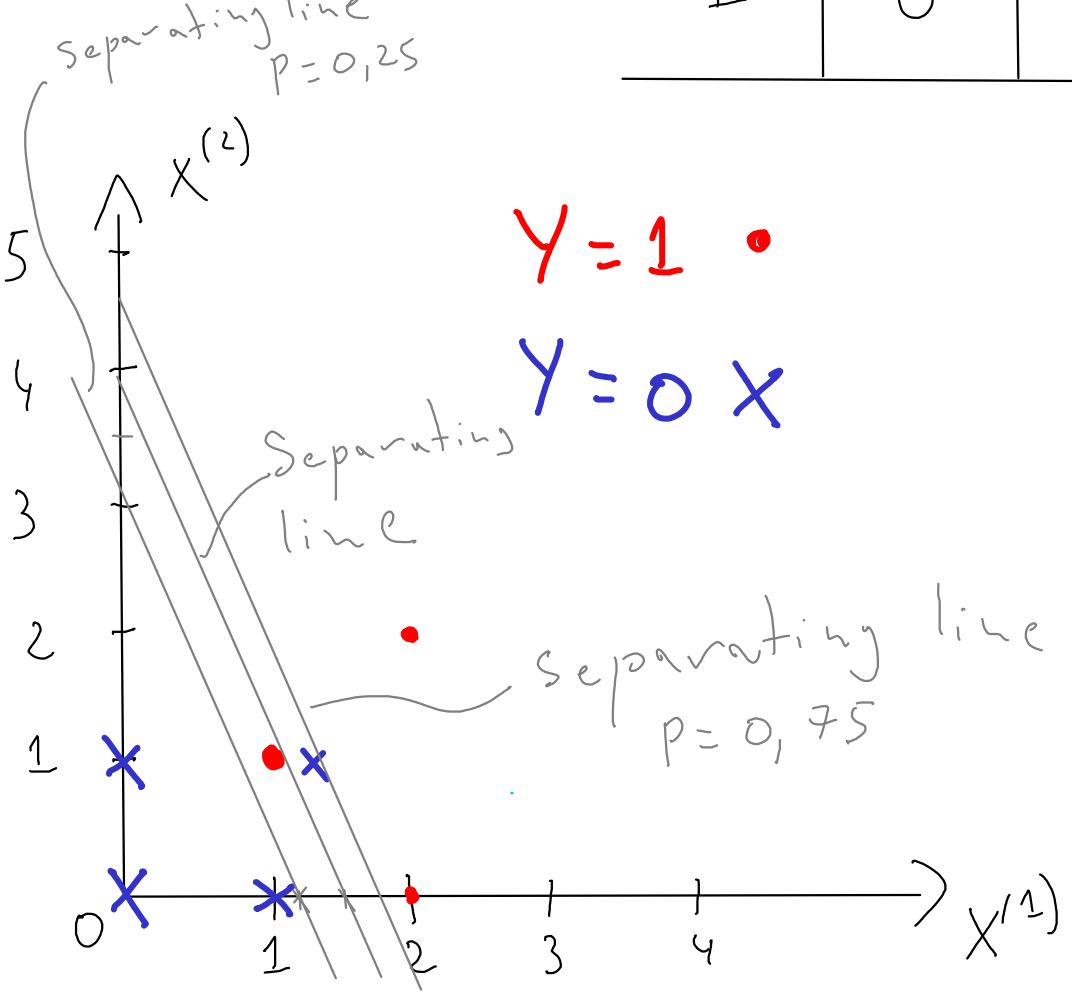
3) 0 classification errors

$$Acc_{\hat{y}(x)}(X, y) = 1$$

Confusion matrix:

| True \ Predicted | 0 | 1 |
|---|---|---|
| 0 | 3 | 0 |
| 1 | 0 | 3 |

4)




Y = 1 •

Y = 0 X

⇒ The classes are not linearly separable anymore.

5) $a = -\frac{3,87}{1,5} \approx -2,58$ $\quad b = \frac{5,85}{1,50} \approx 3,9$ (3)

See figure above for the line.

6) 2 classification errors over 7 observations.

$$Acc_{\hat{Y}(x)}(X,Y) = 1 - \frac{2}{7} \approx 0,71$$

Confusion matrix:

| True \ Predicted | 0 | 1 |
|---|---|---|
| 0 | 3 | 1 |
| 1 | 1 | 2 |

B) 1) If we require $p > 0,5$, we would be more conservative on deciding $\hat{y} = 1$ than for $p = 0,5$ and less conservative on deciding $\hat{Y} = 0$.

For $p < 0,5$, we have the opposite. We would more conservative on deciding $\hat{Y} = 0$ than for $p = 0,5$ and less conservative on deciding $\hat{Y} = 1$.

B) 2) For $(x^{(1)}, x^{(2)})$ on the separation line we have (4)

$$\frac{1}{1+e^{-(\beta_0+\beta_1 x^{(1)}+\beta_2 x^{(2)})}} = p$$

Therefore,

$$e^{-(\beta_0+\beta_1 x^{(1)}+\beta_2 x^{(2)})} = \frac{1}{p} - 1 = \frac{1-p}{p}$$

Taking $-\log(\cdot)$:

$$\beta_0+\beta_1 x^{(1)}+\beta_2 x^{(2)} = -\log\left(\frac{1-p}{p}\right)$$

$$\beta_0+\beta_1 x^{(1)}+\beta_2 x^{(2)} = \log\left(\frac{p}{1-p}\right)$$

Isolating $x^{(2)}$:

$$x^{(2)} = \underbrace{-\left(\frac{\beta_1}{\beta_2}\right)}_{a} x^{(1)} + \underbrace{\left(\frac{\log\left(\frac{p}{1-p}\right)-\beta_0}{\beta_2}\right)}_{b}$$

For $p = 0,25$

$a \approx -2,58$

$b = -\left(\frac{1,1-5,85}{1,5}\right) \simeq 3,17$

For $p = 0,75$

$a \approx -2,58$

$b = -\left(\frac{-1,1-5,85}{1,5}\right) \simeq 9,63$

3) See figure above for the lines. (5)

For $p = 0,25$

1 classification error over 7 observations.

$$Acc_{\hat{y}(x)}(X,y) = 1 - \frac{1}{7} \approx 0,86$$

Confusion matrix:

| True \ Predicted | 0 | 1 |
|---|---|---|
| 0 | 3 | 1 |
| 1 | 0 | 3 |

For $p = 0,75$

1 classification error over 7 observations.

$$Acc_{\hat{y}(x)}(X,y) = 1 - \frac{1}{7} \approx 0,86$$

Confusion matrix:

| True \ Predicted | 0 | 1 |
|---|---|---|
| 0 | 4 | 0 |
| 1 | 1 | 2 |

Results are in accordance with intuition.

when $p < 0,5$, we are more conservative for deciding $\hat{Y} = 0$, thus reducing False Negatives.

while when $p > 0,5$, we are more conservative for deciding $\hat{Y} = 1$, thus reducing False Positives.