Homework Assignment N°1

Thibault Douzon
Rajavarman Mathivanan

September 4th, 2018

Contents

1 Exercise 1: Geometry

1.1 Part a

We have:

$$w_0 + 3w_1 + w_2 = 0$$
$$w_0 - w_1 + 2w_2 = 0$$

Normalize by w_2 it gives

$$W_0 + 3W_1 = -1$$
$$W_0 - W_1 = -2$$

Where $W_0 = \frac{w_0}{w_2}$ and $W_1 = \frac{w_1}{w_2}$
Cramer's rule gives us the results:

$$W_0 = \frac{(-1 \times -1) - (3 \times -2)}{(1 \times -1) - (1 \times 3)} = -\frac{7}{4}$$

$$W_1 = \frac{(1 \times -2) - (-1 \times 1)}{(1 \times -1) - (1 \times 3)} = \frac{1}{4}$$

Thus, we can chose $w_2 = 1$ and then

$$w_0 = -\frac{7}{4}$$

$$w_1 = \frac{1}{4}$$

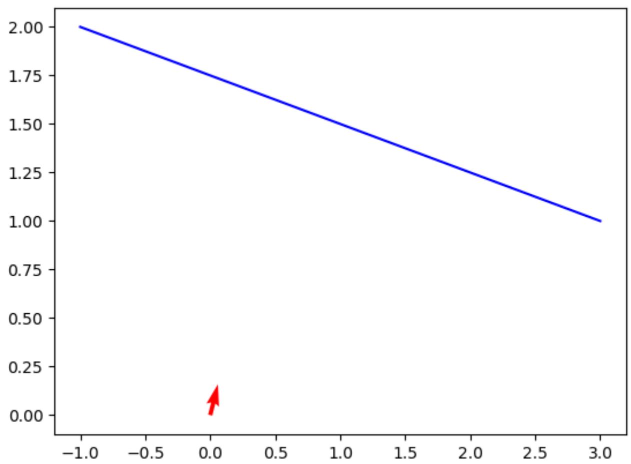
$$w_2 = 1$$

We can check our first system is solved with this solution. Hence the equation of the line is the following:

$$-\frac{7}{4} + \frac{x_1}{4} + x_2 = 0$$

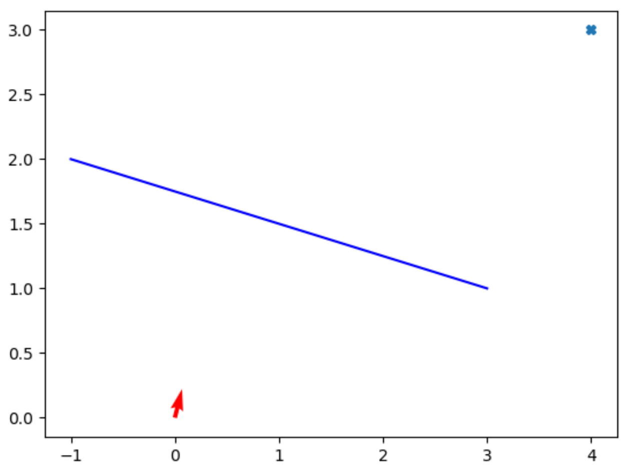
1.2 Part b

Our figure is the following:



1.3 Part c

The point $(4, 3)$ is in the positive side of the line, the most simple «proof» is this plot:



It shows our point «above» the blue line, hence it is in the positive side of the line. For this point, because we saw it was on the positive side, $w_0 + w_1x_1 + w_2x_2$ should give a positive result.

$$-\frac{7}{4} + \frac{4}{4} + 3 = \frac{9}{4} > 0$$

1.4 Part d

First we must notice that (w_1, w_2) is orthogonal to our line. To figure out the direction vector of the line, we can rewrite its equation like this

$$-w_1 x - w_0 = w_2 y$$

Hence its direction vector is $(w_2, -w_1)$ and now we can proove that (w_1, w_2) is orthogonal to the line by showing that

$$(w_1, w_2) \cdot (w_2, -w_1) = 0$$

Now we know the perpendicular direction to the line, we can compute the distance between the point and the line.

$$d_{line}(P) = |[(x_1, x_2) - (3, 1)] \cdot (w_1, w_2)| \,/\, \|(w_1, w_2)\|$$

$$d_{line}(P) = |x_1 w_1 + x_2 w_2 - 3w_1 - 1w_2| \,/\, \|(w_1, w_2)\|$$

Or because $w_0 + 3w_1 + 1w_2 = 0$ ($(3, 1)$ is on the line) we can add $w_0 + 3w_1 + 1w_2$ without changing the value. Hence:

$$d_{line}(P) = \frac{|w_0 + x_1 w_1 + x_2 w_2|}{\|(w_1, w_2)\|}$$

1.5 Part e

We use what we learned from part d:

$$d_{line}((0, 0)) = \frac{|w_0|}{\|(w_1, w_2)\|}$$

Hence we have

$$d_{line}((0, 0)) = \frac{|w_0|}{\sqrt{{w_1}^2 + {w_2}^2}} = \frac{\frac{7}{4}}{\sqrt{\frac{1}{4}^2 + 1}} \approx 1.7$$

2 Exercice 2: Calculus

$$\sigma(y) = \frac{1}{1 + e^{-y}}$$

2.1 Part a

Calculus says $\left(\frac{u}{v}\right)' = \frac{u'v - uv'}{v^2}$, then:

$$\sigma'(y) = \frac{e^{-y}}{(1 + e^{-y})^2}$$

$$\sigma'(y) = \frac{(1 + e^{-y}) - 1}{(1 + e^{-y})^2}$$

$$\sigma'(y) = \sigma(y) \times \frac{(1 + e^{-y}) - 1}{1 + e^{-y}}$$

$$\sigma'(y) = \sigma(y) \times (1 - \sigma(y))$$

2.2 Part b

We know that $\lim_{y\to\infty} e^{-y} = 0$ thus because the denominator does not tend to zero:

$$\lim_{y\to\infty} \sigma(y) = \frac{1}{1 + \lim_{y\to\infty} e^{-y}} = 1$$

and now we know the limit exists and its value:

$$\lim_{y\to\infty} \sigma'(y) = \lim_{y\to\infty} \sigma(y) \times (1 - \sigma(y)) = 1 \times (1 - 1) = 0$$

2.3 Part c

$$\frac{\partial \sigma}{\partial x_1} = \frac{\partial \sigma}{\partial y} \times \frac{\partial y}{\partial x_1}$$

$$\frac{\partial \sigma}{\partial x_1} = (\sigma \times (1 - \sigma)) \times w_1$$

Hence:

$$\frac{\partial \sigma}{\partial x_1}(y) = \sigma(y)(1 - \sigma(y) \times w_1$$

2.4 Part d

Same as in Part c:
For w_1

$$\frac{\partial \sigma}{\partial w_1} = \frac{\partial \sigma}{\partial y} \times \frac{\partial y}{\partial w_1}$$

$$\frac{\partial \sigma}{\partial w_1} = (\sigma \times (1 - \sigma)) \times x_1$$

Hence:

$$\frac{\partial \sigma}{\partial w_1}(y) = \sigma(y)(1 - \sigma(y) \times x_1$$

Same procedure for w_2 and w_0:
For w_2

$$\frac{\partial \sigma}{\partial w_2}(y) = \sigma(y)(1 - \sigma(y) \times x_2$$

For w_0

$$\frac{\partial \sigma}{\partial w_0}(y) = \sigma(y)(1 - \sigma(y) \times 1$$

Finally we can compute the gradient:

$$\nabla_w \sigma(y) = \begin{bmatrix} \frac{\partial \sigma}{\partial w_0} \\ \\ \frac{\partial \sigma}{\partial w_1} \\ \\ \frac{\partial \sigma}{\partial w_2} \end{bmatrix} (y) = \begin{bmatrix} \sigma(y)(1 - \sigma(y) \\ \\ \sigma(y)(1 - \sigma(y))x_1 \\ \\ \sigma(y)(1 - \sigma(y))x_2 \end{bmatrix} = \sigma(y)(1 - \sigma(y))x$$

Where $x = \begin{bmatrix} 1 & x_1 & x_2 \end{bmatrix}^\top$

2.5 Part e

When $x_1 > 0$, if $w_1 \gg 1$ (and assuming w_1 larger than w_0 and w_2) then $y = w_0 + x_1 w_1 + x_2 w_2 \gg 1$
Then what we showed in Part b applies:

$$\lim_{y \to \infty} \frac{\partial \sigma}{\partial w_1}(y) = \lim_{y \to \infty} \sigma(y)(1 - \sigma(y))x_1 = 0 \times x_1 = 0$$

When $x_1 < 0$, under the same assumptions, we have $\lim_{w_1 \to \infty} y = -\infty$. That's why we first need to find the value of $\lim_{y \to -\infty} \sigma'(y)$

$$\lim_{y \to -\infty} \sigma(y) = \lim_{y \to -\infty} \frac{1}{1 + e^{-y}} = 0$$

$$\lim_{y \to -\infty} \sigma'(y) = \lim_{y \to -\infty} \sigma(y)(1 - \sigma(y)) = 0$$

Then

$$\lim_{y \to -\infty} \frac{\partial \sigma}{\partial w_1}(y) = \lim_{y \to -\infty} \sigma'(y) \times x_1 = 0 \times x_1 = 0$$

3 Exercise 3

3.1 Part a