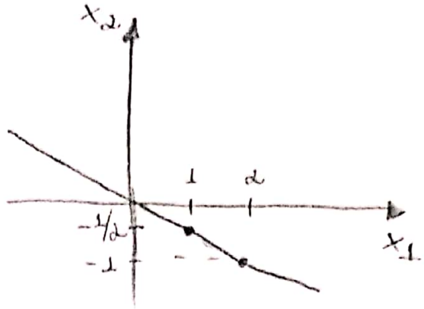
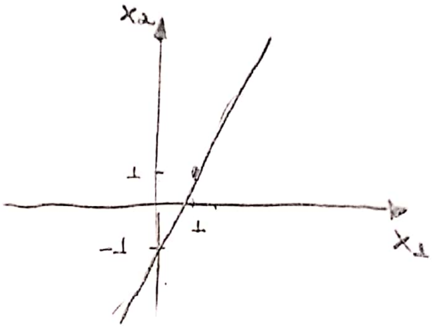
LISTA TEÓRICA I - DEEP LEARNING

THIAGO MALTA COUTINHO - 2014123335

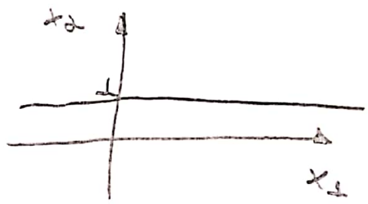
1 a) $2x_2 + x_1 = 0 \Rightarrow x_2 = \frac{-x_1}{2}$



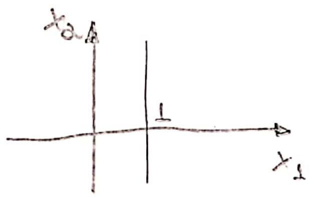
b) $x_2 - 2x_1 + 1 = 0 \Rightarrow x_2 = 2x_1 - 1$



c) $x_2 - 1 = 0 \Rightarrow x_2 = 1$



d) $x_1 - 1 = 0 \Rightarrow x_1 = 1$



)

$x_2 = 3 + 2x_1$

Retas paralelas: $x_2 = b + 2x_1 \quad \forall\, b \in \mathbb{R}$

Retas perpendiculares: $a \cdot 2 = -1 \Rightarrow a = -1/2$

$\therefore x_2 = b - \frac{1}{2} x_1 \quad \forall\, b \in \mathbb{R}$

(5) a) $x = \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix}$; $A \in \mathbb{R}^{m \times m}$

$$x' A x = [x_1 \cdots x_m] \begin{bmatrix} A_1 \cdots A_m \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} = [x_1 \cdots x_m] \underbrace{\left(x_1 \begin{bmatrix} A_1 \end{bmatrix} + \cdots + x_m \begin{bmatrix} A_m \end{bmatrix} \right)}_{\begin{bmatrix} \sum_i x_i A_{1i} \\ \vdots \\ \sum_i x_i A_{mi} \end{bmatrix}} =$$

$$= x_1 \sum_i x_i A_{1i} + \cdots + x_m \sum_i x_i A_{mi} = \boxed{\sum_{ij} x_i x_j A_{ij}}$$

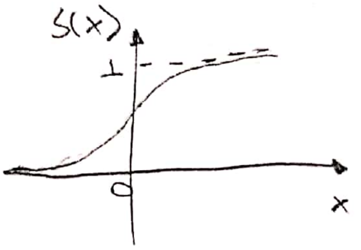
b) $$[x_1 \cdots x_m] \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} = x_1 \cdot x_1 + \cdots + x_m x_m = \boxed{\sum_i (x_i)^2}$$

c) $$\begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} [x_1 \cdots x_m] = \begin{bmatrix} x_1 \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} & x_2 \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} & \cdots & x_m \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} \end{bmatrix} =$$

$$= \begin{bmatrix} x_1^2 & x_2 x_1 & \cdots \\ x_1 x_2 & x_2^2 & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}$$

Scanned by CamScanner

⑧ a)

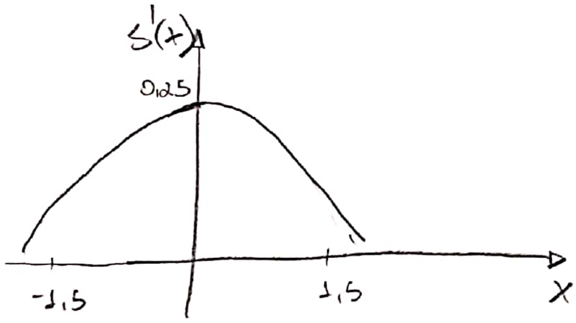


b) $s(-x) = \frac{1}{1+e^{x}}$; $1 - s(x) = 1 - \frac{1}{1+e^{-x}} = \frac{1+e^{-x}-1}{1+e^{-x}} = \frac{e^{-x}}{1+e^{-x}} \cdot \frac{e^{x}}{e^{x}} =$

$$= \frac{e^{-x}e^{x}}{e^{x}+e^{-x}e^{x}} = \frac{1}{1+e^{x}} \quad \therefore \boxed{s(-x) = 1 - s(x)}$$

c) $$s'(x) = \left[\left(1+e^{-x}\right)^{-1}\right]' = \frac{1 \cdot e^{-x}}{(1+e^{-x})^2} = \underbrace{\frac{1}{1+e^{-x}}}_{s(x)} \cdot \underbrace{\frac{e^{-x}}{1+e^{-x}}}_{1-s(x)}$$

$$\therefore \boxed{s'(x) = s(x)[1-s(x)]}$$



d) $\arg\max_{x} s'(x) = 0 \longrightarrow s'(0) = 1/4$

e) $s(z) = (1+e^{-z})^{-1}$; $z = b + w_1 x$

$$\frac{\partial s}{\partial b} = -1(1+e^{-z})^{-2} \cdot e^{-z} \cdot \frac{\partial -z}{\partial b} = \frac{e^{-z}}{(1+e^{-z})^2}$$

$$\frac{\partial s}{\partial w_1} = -1(1+e^{-z})^{-2} \cdot e^{-z} \frac{\partial -z}{\partial w_1} = \frac{e^{-z} \cdot x}{(1+e^{-z})^2}$$

f) $s(h(x)) = \left[1+e^{-h(x)}\right]^{-1}$

$$\frac{\partial s}{\partial x} = \left(1+e^{-h(x)}\right)^{-2} \cdot e^{-h(x)} \cdot \frac{\partial h(x)}{\partial x}$$

(9) $X\beta = Y \longrightarrow (X^TX)\beta = X^TY$

$$(X^TX)^{-1}(X^TX)\beta = (X^TX)^{-1}X^TY \Rightarrow \boxed{\beta = (X^TX)^{-1}X^TY}$$

---

(11) $w_0 = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$; $m = 0{,}1$

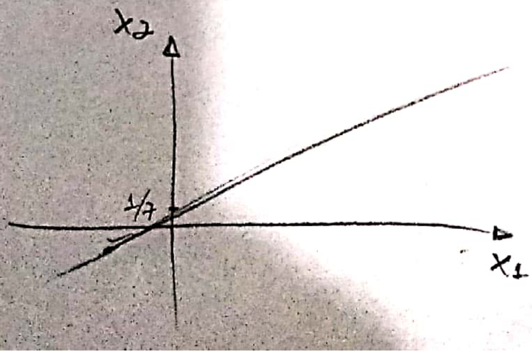
$$X = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2{,}3 & 2 \\ 1 & 2{,}4 & 1{,}5 \\ 1 & 2 & 1 \\ 1 & 1{,}5 & 0{,}5 \\ 1 & 0{,}7 & 2{,}7 \\ 1 & 0{,}3 & 2 \\ 1 & 1 & 2 \\ 1 & 0{,}5 & 1 \end{bmatrix}$$

$$w_{m+1} = w_m + m(y - \hat{y}_0)x$$

$$\hat{y}_0 = Xw_0 > 0 = \begin{bmatrix} -2+3>0 \\ -2{,}3+2>0 \\ -2{,}4+1{,}5>0 \\ -2+1>0 \\ -1{,}5+0{,}5>0 \\ -0{,}7+2{,}7>0 \\ -0{,}3+2>0 \\ -1+2>0 \\ -0{,}5+1>0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

$$w_1 = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} + 0{,}1 \times \begin{bmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \Rightarrow w_1 = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} + \begin{bmatrix} -0{,}1 \\ -0{,}2 \\ -0{,}3 \end{bmatrix} = \begin{bmatrix} -0{,}1 \\ -1{,}2 \\ 0{,}7 \end{bmatrix}$$

$$-0{,}1 - 1{,}2x_1 + 0{,}7x_2 = 0 \Rightarrow x_2 = \frac{1{,}2x_1}{0{,}7} + \frac{0{,}1}{0{,}7}$$



Scanned by CamScanner

(1b) $P(Y=1|b,w_1) = \left[1+e^{-(b+w_1x)}\right]^{-1}$ → p/1 amostra, igual à $p(x_i)$

$\theta = \langle b, w_1 \rangle$

Likelihood: $L(\theta) = \prod_{i=1}^{m} p(x_i)^{y_i} (1-p(x_i))^{1-y_i}$

$$\log l(\theta) = \sum_m y_i \log(p(x_i)) + (1-y_i)\log(1-p(x_i))$$

$$\frac{\partial \log l(\theta)}{\partial b} = \sum_m \underbrace{y_i \frac{\partial \log(p(x_i))}{\partial b}} + \underbrace{(1-y_i)\frac{\partial \log(1-p(x_i))}{\partial b}}$$

$$\frac{\partial \log l(\theta)}{\partial b} = \sum_m y_i \frac{1}{p(x_i)} \cdot \frac{\partial p(x_i)}{\partial b} + \frac{(1-y_i)}{1-p(x_i)} \underbrace{\frac{\partial}{\partial b} 1-p(x_i)}_{\frac{\partial -p(x_i)}{\partial b}}$$

$$= \sum_m \underbrace{\frac{\partial p(x_i)}{\partial b}} \left(\frac{y_i}{p(x_i)} - \frac{(1-y_i)}{1-p(x_i)}\right)$$

$$\frac{\partial}{\partial b}\left[1+e^{-(b+w_1x)}\right]^{-1} = \frac{(-1)\cdot e^{-(b+w_1x)}(-1)}{\left[1+e^{-(b+w_1x)}\right]^2} = \frac{1}{1+e^{-(b+w_1x)}} \cdot \frac{e^{-(b+w_1x)}}{1+e^{-(b+w_1x)}}$$

$$= p(x_i) \cdot 1-p(x_i)$$

$$= \sum_m \frac{p(x_i)}{p(x_i)}(1-p(x_i)) \cdot y_i - \left(1-y_i \cdot p(x_i) \cdot \frac{1-p(x_i)}{1-p(x_i)}\right) =$$

$$\boxed{= \sum_m y_i - p(x_i)}$$

Scanned by CamScanner

$$\frac{\partial \log l(\theta)}{\partial w_1} = \sum_m \left(\frac{y_i}{p(x_i)} - \frac{1-y_i}{1-p(x_i)} \right) \cdot \underbrace{\frac{\partial p(x_i)}{\partial w_1}}$$

$$\frac{\partial}{\partial w_1}\left[1+e^{-(b+w_1x)}\right]^{-1} = \frac{(-1) \cdot -x_i \cdot e^{-(b+w_1x)}}{\left[1+e^{-(b+w_1x)}\right]^2} = p(x_i)(1-p(x_i)) \cdot x_i$$

$$\boxed{\frac{\partial \log l(\theta)}{\partial w_1} = \sum_m \left[y_i - p(x_i)\right] x_i}$$

---

(20) $f(x) = -(x-3)^4$ com $x_0 = 1$

$f'(x) = g(x)$; TAYLOR: $g(x) \approx g(x_0) + g'(x_0) \cdot (x - x_0) = 0$

$$x_1 = x_0 - \frac{g(x_0)}{g'(x_0)} \Rightarrow x_1 = x_0 - \frac{f'(x_0)}{f''(x_0)}$$

$f'(x) = -4(x-3)^3$; $f''(x) = -12(x-3)^2$; $\frac{f'(x)}{f''(x)} = \frac{-4}{-12}\frac{(x-3)^3}{(x-3)^2} = \frac{+1}{3}(x-3)$

<u>PASSO 1:</u> $x_1 = 1 - \left(\frac{1}{3} \cdot -2\right) = 5/3$

<u>PASSO 2:</u> $x_2 = \frac{5}{3} - \left(\frac{1}{3} \cdot \frac{-4}{3}\right) = \frac{19}{9} \approx 2{,}2$