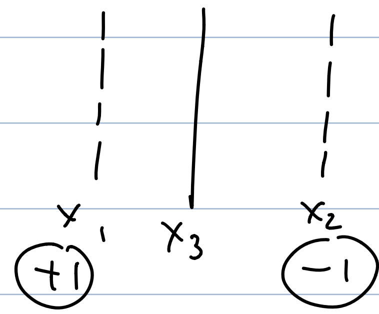
Phang Teng Fone 1003296 DL HW2

1. $x_1 = (-5, 1, 3)$ $y_1 = +1$

$x_2 = (2, 2, -3)$ $y_2 = -1$



$w x_3 + b = 0 \Rightarrow$ hyperplane

$x_3 = \frac{1}{2}(x_1 + x_2)$

$= (-1.5, 1.5, 0)$

$w = x_1 - x_3$

$= (-3.5 \quad -0.5 \quad -3)$

$$(-3.5 \quad -0.5 \quad -3)\begin{array}{c} x_3 \\ \begin{bmatrix} -1.5 \\ 1.5 \\ 0 \end{bmatrix} \end{array} + b = 0$$

$$b = -4.5$$

$\therefore x_1 w + b = [-3.5 \quad -0.5 \quad -3]\begin{bmatrix} -5 \\ 1 \\ 3 \end{bmatrix} - 4.5$

$= 3.5$

$\therefore x_2 w + b = [-3.5 \quad -0.5 \quad -3]\begin{bmatrix} 2 \\ 2 \\ -3 \end{bmatrix} - 4.5$

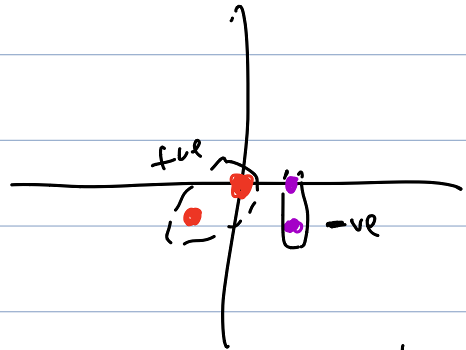
$= -3.5$

$\therefore$ After normalize.

$$f(x) = \frac{1}{3.5}\begin{bmatrix} -3.5 \\ -0.5 \\ -3 \end{bmatrix} x_{new} - \frac{4.5}{3.5}$$

2.

$$g(z) = \begin{cases} z & \text{if } |z| \le 1 \\ \text{sign}(z) & \text{otherwise} \end{cases}$$



$z = w_1 x_1 + w_2 x_2 - W_0$

$z = w_1 x_1 + w_2 x_2 + 1$

+ve = $(0,0)$ $(0,-1)$

By guess & check:

-ve = $(1,0)$ $(1,-1)$

$z = -3x_1 + 0.5x_2 + 1$

$\therefore$ $z = -3x_1 + \frac{1}{2}x_2 + 1$ #

3.1 $\dfrac{\partial \log(s(a))}{\partial a} = 1 - s(a)$

LHS: $s(a) = \dfrac{\exp(a)}{1+\exp(a)}$

$= \dfrac{1}{e^{-a}+1}$

$\log(s(a)) = \log\left(\dfrac{1}{e^{-a}+1}\right)$

$= \cancel{\log(1)} - \log(e^{-a}+1)$

$= -\log(e^{-a}+1)$

$\dfrac{\partial \log(s(a))}{\partial a} = -\dfrac{1}{e^{-a}+1}(-e^{-a})$

$= \dfrac{e^{-a}}{e^{-a}+1}$

$= 1 - \dfrac{1}{e^{-a}+1}$

$\therefore 1 - s(a)$ # (PROVED)

$$\log(1-s(a)) = \log\left(\frac{e^{-a}}{e^{-a}+1}\right)$$
$$= \log(e^{-a}) - \log(e^{-a}+1)$$
$$= -a - \log(e^{-a}+1)$$

$$\frac{\partial \log(1-s(a))}{\partial a} = -1 - \left[\frac{1}{e^{-a}+1} \cdot -e^{-a}\right]$$

$$= -1 + \frac{e^{-a}}{e^{-a}+1}$$

$$= -\left(\frac{e^{-a}+1}{e^{-a}+1}\right) + \frac{e^{-a}}{e^{-a}+1}$$

$$= \frac{-e^{-a} - 1 + e^{-a}}{e^{-a}+1}$$

$$= -\frac{1}{e^{-a}+1}$$

$$= -s(a) \quad \text{(PROVE)}$$

3.2 $\nabla_w L = \sum_{i=1}^{n} x_i (s(w.x_i) - y_i) = \sum_{i=1}^{n} x_i (h(x_i) - y_i) \quad \because h(x) = s(w.x)$

LHS: $L = (-1) . \sum_{i=1}^{n} y_i \log[h(x_i)) + (1-y_i) \log (1-h(x_i)]$

Sub $h(x_i)$ into $s(a)$ from 3.1

$L = (-1) \sum_{i=1}^{n} y_i \log[s(w.x_i) + (1-y_i) \log (1 - s(w.x_i))]$

$$\nabla_w L = (-1)\sum_{i=1}^{n} y_i \left[\left(1 - S(w \cdot x_i)\right)(x_i) + (1-y_i)\left(-S(w_i \cdot x_i)\right)x_i\right]$$

$$= \sum_{i=1}^{n} -x_i \left(y_i - \cancel{y_i S(w \cdot x_i)} - S(w \cdot x_i) + \cancel{y_i S(w \cdot x_i)} \right)$$

$$= \sum_{i=1}^{n} -x_i \left(y_i - S(w \cdot x_i) \right)$$

$$= \sum_{i=1}^{n} x_i \left(S(w \cdot x_i) - y_i \right)$$

$$= \sum_{i=1}^{n} x_i \left(h(x_i) - y_i \right) \quad \text{(PROVE)}$$

4. $$\frac{\partial L}{\partial n_4} = \frac{\partial L}{\partial n_2} \cdot \frac{\partial n_2}{\partial n_4} + \frac{\partial L}{\partial n_3} \cdot \frac{\partial n_3}{\partial n_4}$$

$$\frac{\partial L}{\partial n_5} = \frac{\partial L}{\partial n_4} \cdot \frac{\partial n_4}{\partial n_5} = \frac{\partial L}{\partial n_2} \cdot \frac{\partial n_2}{\partial n_4} \frac{\partial n_4}{\partial n_5} + \frac{\partial L}{\partial n_3} \cdot \frac{\partial n_3}{\partial n_4} \cdot \frac{\partial n_4}{\partial n_5}$$

$$\frac{\partial L}{\partial n_6} = \frac{\partial L}{\partial n_4} \cdot \frac{\partial n_4}{\partial n_6} + \frac{\partial L}{\partial n_3} \cdot \frac{\partial n_3}{\partial n_6} = \frac{\partial L}{\partial n_3} \cdot \frac{\partial n_3}{\partial n_6} + \frac{\partial L}{\partial n_2} \cdot \frac{\partial n_2}{\partial n_4} \cdot \frac{\partial n_4}{\partial n_6} + \frac{\partial L}{\partial n_3} \frac{\partial n_3}{\partial n_4} \frac{\partial n_4}{\partial n_6}$$

$$\frac{\partial L}{\partial n_7} = \frac{\partial L}{\partial n_5} \cdot \frac{\partial n_5}{\partial n_7} + \frac{\partial L}{\partial n_6} \cdot \frac{\partial n_6}{\partial n_7}$$

$$= \frac{\partial L}{\partial n_2} \cdot \frac{\partial n_2}{\partial n_4} \frac{\partial n_4}{\partial n_5} \frac{\partial n_5}{\partial n_7} + \frac{\partial L}{\partial n_3} \cdot \frac{\partial n_3}{\partial n_4} \cdot \frac{\partial n_4}{\partial n_5} \frac{\partial n_5}{\partial n_7} +$$

$$\frac{\partial L}{\partial n_3} \cdot \frac{\partial n_3}{\partial n_6} \frac{\partial 6}{\partial n_7} + \frac{\partial L}{\partial n_2} \cdot \frac{\partial n_2}{\partial n_4} \cdot \frac{\partial n_4}{\partial n_6} \frac{\partial n_6}{\partial n_7} + \frac{\partial L}{\partial n_3} \frac{\partial n_3}{\partial n_4} \frac{\partial n_4}{\partial n_6} \frac{\partial n_6}{\partial n_7}$$

$$\frac{\partial L}{\partial x} = \frac{\partial L}{\partial n_7} \cdot \frac{\partial n_7}{\partial x}$$

$$= \frac{\partial L}{\partial n_2} \cdot \frac{\partial n_2}{\partial n_4} \frac{\partial n_4}{\partial n_5} \frac{\partial n_5}{\partial n_7} \frac{\partial n_7}{\partial n_x} + \frac{\partial L}{\partial n_3} \cdot \frac{\partial n_3}{\partial n_4} \cdot \frac{\partial n_4}{\partial n_5} \frac{\partial n_5}{\partial n_7} \frac{\partial n_7}{\partial n_x} +$$

$$\frac{\partial L}{\partial n_3} \cdot \frac{\partial n_3}{\partial n_6} \frac{\partial 6}{\partial n_7} \frac{\partial n_7}{\partial n_x} + \frac{\partial L}{\partial n_2} \cdot \frac{\partial n_2}{\partial n_4} \cdot \frac{\partial n_4}{\partial n_6} \frac{\partial n_6}{\partial n_7} \frac{\partial n_7}{\partial n_x}$$

$$+ \frac{\partial L}{\partial n_3} \frac{\partial n_3}{\partial n_4} \frac{\partial n_4}{\partial n_6} \frac{\partial n_6}{\partial n_7} \frac{\partial n_7}{\partial n_x}$$

5

```
#Q5.1
q1a = torch.randn(3,2,5)
q1b = torch.randn(3)
torch.einsum('ijk,i->jk',[q1a,q1b])
```

```
tensor([[-1.4106,  0.0211, -2.0845, -3.1763, -1.1233],
        [-2.4421, -1.0829,  0.6744,  2.8113,  0.4890]])
```

```
[3] #Q5.2
q2 = torch.randn(3,2,5,3)
torch.einsum('ijkl->ik',[q2])
```

```
tensor([[-1.5118, -0.0108, -0.5652, -0.2242,  2.2020],
        [ 5.3677, -2.1148,  0.3285, -2.1441, -2.5794],
        [ 2.6283,  3.4045, -0.8401,  6.2966,  3.0774]])
```

```
[4] #Q5.3
q3 = torch.randn(3,2,5,3)
torch.einsum('ijkl->ki',[q3])
```

```
tensor([[-5.3693,  1.5106, -1.5315],
        [-3.6612,  1.1715, -0.9572],
        [ 0.7339, -2.0529, -0.6262],
        [ 1.0893, -0.0686, -0.4280],
        [ 4.2637,  0.4691, -2.0942]])
```

```
[5] #Q5.4
q4 = torch.randn(3,2,5)
torch.einsum('ijk,ijk->i',[q4,q4])
```

```
tensor([10.8261,  6.0697,  7.3739])
```

```
[6] #Q5.5
q5a = torch.arange(6).reshape(2, 3)
q5gt = torch.arange(6).reshape(3, 2)
q5b = torch.arange(6).reshape(2, 3)
torch.einsum('de,ef,fl->dl',[q5a,q5gt,q5b])
```

```
tensor([[ 39,  62,  85],
        [120, 188, 256]])
```

2-tensor order