Name: Asadollah Hill Galib

CSE 891: Deep learning

Homework 1

(Q-1)

$$W^{(1)} = \begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix}, \quad b^{(1)} = [0 \ 0 \ 0]$$

$$W^{(2)} = [1 \ 1 \ 1], \quad b^{(2)} = [-2 \cdot 0]$$

Intuition: $W^{(1)}$ is dealing with comparing ~~distinct~~ successive pair in each row:

$$W^{(1)} \begin{cases} \text{1st row: } & -x_1 + x_2 \\ \text{2nd row: } & -x_2 + x_3 \\ \text{3rd row: } & -x_3 + x_4 \end{cases} \Rightarrow$$

We want all of them to be greater than 0: which will evaluate to $[1 \ 1 \ 1]$ after activation.

$W^{(2)}$ is taking sum of the output from the previous layer. We want this sum to be 3, which indicates satifying all pairs $(x_1 < x_2, x_2 < x_3, x_3 < x_4)$.

$b^{(2)}$ is for checking whether the sum is 3 or lesser. If, the sum is 3, it will evaluate to $3 - 2 = 1 \Rightarrow \phi(1) = 1$

Otherwise, it will evaluate to 0 or lesser $\Rightarrow \phi(0) = 0$.

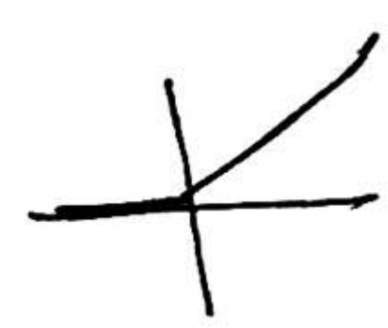
Q-2

$\frac{dL}{dw_1}$: Yes

$$\frac{dL}{dw_1} = y' \frac{dy}{dw_1} = y' \phi'(z)\, h_1 = 0 \qquad [\text{as } h_1 = 0]$$

$\frac{dL}{dw_2}$: Yes

$$\frac{dL}{dw_2} = h_1' \frac{dh_1}{dw_2} = h_1' \phi'(z)\, h_3$$



$$= h_1' \phi'(-1)\, h_3$$

$$= 0 \qquad [\text{as } \mathrm{Relu}'(-1) = 0]$$

$\frac{dL}{dw_3}$: No

$$\frac{dL}{dw_3} = h_3' \frac{dh^3}{dw_3}$$

$$= (w_2' + w_4') \frac{dh^3}{dw_3}$$

~~So~~ It's not ~~gurranteed~~ guaranteed to be 0, as it has two parts (w_2' & w_4') and we do not know anything about w_4'.

So, $\frac{dL}{dw_3}$ can be 0 or other.

Q-3

1. $$\hat{f}_\tau(x) = w_1 a(w_0 x + b_0) + b_1$$

$$a(y) = \mathbb{I}(y \geq 0)$$

$$g(h, a, b, x) = h.\ \mathbb{I}(a \leq x \leq b)$$

Let, $n = 2$

$$W_0 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad b_0 = \begin{bmatrix} -a \\ b \end{bmatrix} \Bigg\} \text{ to check } ~~compare~~ \Big\langle \begin{matrix} x \geq a \\ \text{and} \\ x \leq b \end{matrix}$$

$$w_1 = \begin{bmatrix} h \\ h \end{bmatrix}, \quad b_1 = -h \Bigg\} \begin{matrix} \text{to get } h \text{ when } a \leq x \leq b \\ \text{and } 0 \text{ when } x < a \\ \text{or } x > b \end{matrix}$$

Then, $$w_0 x + b_0 = \begin{bmatrix} x - a \\ b - x \end{bmatrix}$$

$$\Rightarrow a(w_0 x + b_0) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \text{ if } \begin{cases} x \geq a \text{ and} \\ x \leq b \end{cases}$$

$$\Rightarrow w_1 a(w_0 x + b_0) + b_1 = h + h - h$$
$$= h = g(h, a, b, x) \; [\text{when, } a \leq x \leq b]$$

In any other cases, we will get $a(w_0 x + b_0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ or $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ which will evaluate to $h - h = 0 = g(h, a, b, x)$ [when $x > b$ or $x < a$]

(*) Here, both $x < a$ and $x > b$ case is not possible, as $a \leq b$.

Q-3

(2) $f(x) = -x^2 + 1$

$\hat{f}_0(x) = 0$

New function, $\hat{f}_1(x) = \hat{f}_0(x) + g(h_1, a_1, b_1, x)$

where, $a_1 = -1$, $b_1 = 1$ and

$h_1 = \frac{1}{2} \max(f(x))$ [$\max(f(x))$ can be found by setting $f'(x) = 0$:

$= \frac{1}{2} \cdot 1 = \frac{1}{2}$ $\Rightarrow -2x = 0 \Rightarrow x = 0$

$\Rightarrow f(0) = 1$]

We need to show,

$$\|f - \hat{f}_1\| \leq \|f - \hat{f}_0\|$$

$\|f - \hat{f}_0\|$

$= \int_{-1}^{1} |f(x) - 0| \, dx$

$= \int_{-1}^{1} |-x^2 + 1| \, dx$

$= \left[-\frac{x^3}{3} + x\right]_{-1}^{1}$

$= -\frac{2}{3} + 2$

$= \frac{4}{3}$

$\|f - \hat{f}_1\|$

$= \int_{-1}^{1} |f(x) - \hat{f}_1(x)| \, dx$

$= \int_{-1}^{1} |-x^2 + 1 - h_1| \, dx$ $\left[\hat{f}_1(x) = g(h_1, -1, 1, x) = h_1 \cdot \mathbb{I}(-1 \leq x \leq 1) \right]$

$= \int_{-1}^{1} |-x^2 + 1 - \frac{1}{2}| \, dx$

$= \left[-\frac{x^3}{3} + \frac{1}{2}x\right]_{-1}^{1}$

$= -\frac{2}{3} + 1$

$= \frac{1}{3} < \frac{4}{3} = \|f - \hat{f}_0\|$

So, $\|f - \hat{f}_1\| \leq \|f - \hat{f}_0\|$