Page 1

Homework-4

Ketan Mehta
Kmm2304

Q 2.1.1) Based on given design we can write following equations —

$a=b=c=1$ & $w_i=w_h=1$ & $h_0=0$, $x_0=1$, $x_1=0$

~~$h_1 = \sigma(w_i x + w_h + b)$~~

$$h_1 = \sigma(w_i x_0 + w_h h_0 + b) \quad \text{--- (1)}$$

based on ~~this~~ values given we can write h_1 as —

$$h_1 = \sigma(2) \quad \text{--- (2)}$$

$$h_2 = \sigma(w_i x_1 + w_h h_1 + b) \quad \text{--- (3)}$$

After substituting values we get,

$$h_2 = \sigma(\sigma(2) + 1) \quad \text{--- (4)}$$

Also, we can write y_1 & y_2 as follows—

$$y_1 = a h_1 + c \quad \text{--- (5)}$$

$$y_2 = a h_2 + c \quad \text{--- (6)}$$

Now let's compute loss J

we can write J as follows—

$J = \sum_i J_i(\theta)$ &

$$J_1 = \frac{1}{2}(\hat{y}_1 - y_1)^2 \quad \& \quad J_2 = \frac{1}{2}(\hat{y}_2 - y_2)^2 \quad \text{--- (7)}$$

$$\therefore J = J_1 + J_2 \quad \text{--- (8)}$$

~~Now~~ ~~let~~

Now let's compute gradients of J with respect to w_i, w_h, a, b & c.

a) $\therefore \dfrac{\partial J}{\partial a} = \sum_{i} \dfrac{\partial J_i}{\partial a} = \dfrac{\partial J_1}{\partial a} + \dfrac{\partial J_2}{\partial a}$

$$\frac{\partial J_1}{\partial a} = \frac{\partial J_1}{\partial y_1} \frac{\partial y_1}{\partial a} \quad \text{--- (9)}$$

$$\frac{\partial J_1}{\partial y_1} = -(g_1 - y_1) \quad \cancel{ete}$$

$$\frac{\partial y_1}{\partial a} = h_1 \qquad \therefore \frac{\partial J_1}{\partial a} = -h_1(g_1 - y_1) \quad \text{--- (10)}$$

Similarly we can write above for $\frac{\partial J_2}{\partial y_2}$ & $\frac{\partial y_2}{\partial a}$

$$\therefore \frac{\partial J_2}{\partial y_2} = -(g_2 - y_2) \quad \& \quad \frac{\partial y_2}{\partial a} = h_2$$

$$\therefore \frac{\partial J_2}{\partial a} = -h_2(g_2 - y_2) \quad \text{--- (11)}$$

$$\therefore \boxed{\frac{\partial J}{\partial a} = -h_1(g_1 - y_1) + -h_2(g_2 - y_2)} \quad \text{--- (12)}$$

~~b)~~ $\dfrac{\partial J}{\partial c} = \sum_{i=1}^{2} \dfrac{\partial J_i}{\partial c} = \dfrac{\partial J_1}{\partial c} + \dfrac{\partial J_2}{\partial c}$

$$\frac{\partial J_1}{\partial c} = \frac{\partial J_1}{\partial y_1} \frac{\partial y_1}{\partial c} \quad \& \quad \frac{\partial J_2}{\partial c} = \frac{\partial J_2}{\partial y_2} \frac{\partial y_2}{\partial c}$$

$$\frac{\partial y_1}{\partial c} = 1 \quad \& \quad \frac{\partial y_2}{\partial c} = 1$$

$$\therefore \frac{\partial J}{\partial c} =$$

Page 3

b) $\frac{\partial J}{\partial b} = \sum_{i=1}^{2} \frac{\partial J_i}{\partial b}$

$\frac{\partial J_1}{\partial b} = \frac{\partial J_1}{\partial y_1} \frac{\partial y_1}{\partial h_1} \frac{\partial h_1}{\partial b}$

$\frac{\partial y_1}{\partial h_1} = a$ & $\frac{\partial h_1}{\partial b} = 1(h_1)(1-h_1)$

Let's put these values in $\frac{\partial J_1}{\partial b}$ to get —

$\frac{\partial J_1}{\partial b} = -(g_1-y_1) \times a \times 1 = -a(g_1-y_1)(h_1)(1-h_1)$

Similarly, $\frac{\partial J_2}{\partial b} = \frac{\partial J_2}{\partial y_2} \frac{\partial y_2}{\partial h_2} \frac{\partial h_2}{\partial b}$

where, $\frac{\partial y_2}{\partial h_2} = a$ &

$\frac{\partial h_2}{\partial b} = \frac{\partial h_2}{\partial b} + \frac{\partial h_2}{\partial h_1} \frac{\partial h_1}{\partial b} =$

$\frac{\partial h_2}{\partial b} = h_2(1-h_2)$ & $\frac{\partial h_2}{\partial h_1} = h_2(1-h_2) \times w_h$

we can write $\frac{\partial h_2}{\partial b}$ as —

$\frac{\partial h_2}{\partial b} = h_2(1-h_2) + h_2(1-h_2) \times w_h \times h_1(1-h_1)$

$\frac{\partial J_2}{\partial b} = -a(g_2-y_2)\, h_2(1-h_2)\left[1 + w_h \times h_1(1-h_1)\right]$

$\therefore \frac{\partial J}{\partial b} = -a\left[(g_1-y_1)\, h_1(1-h_1) + (g_2-y_2)\, h_2(1-h_2)\left(1+w_h h_1(1-h_1)\right)\right]$ (13)

c) $\frac{\partial J}{\partial c} = \frac{\partial J_1}{\partial c} + \frac{\partial J_2}{\partial c}$

$$\frac{\partial J_1}{\partial c} = \frac{\partial J_1}{\partial y_1}\frac{\partial y_1}{\partial c} = -(g_1-y_1)\times 1 = -(g_1-y_1)$$

$$\frac{\partial J_2}{\partial c} = \frac{\partial J_2}{\partial y_2}\frac{\partial y_2}{\partial c} = -(g_2-y_2)\times 1 = -(g_2-y_2)$$

$$\boxed{\frac{\partial J}{\partial c} = -\left[(g_1-y_1)+(g_2-y_2)\right]} \quad \text{———} \quad (14)$$

d) $\frac{\partial J}{\partial w_i} = \frac{\partial J_1}{\partial w_i} + \frac{\partial J_2}{\partial w_i}$

$$\frac{\partial J_1}{\partial w_i} = \frac{\partial J_1}{\partial y_1}\frac{\partial y_1}{\partial h_1}\frac{\partial h_1}{\partial w_i}$$

$$= -(g_1-y_1)\times a\times x_0\,(h_1)(1-h_1)$$

$$= -a x_0 (g_1-y_1)(h_1)(1-h_1)$$

$$\frac{\partial J_2}{\partial w_i} = \frac{\partial J_2}{\partial y_2}\frac{\partial y_2}{\partial h_2}\frac{\partial h_2}{\partial w_i}$$

$$\frac{\partial h_2}{\partial w_i} = \cancel{-(g_2-y_2)\times a}\ \frac{\partial h_2}{\partial w_i} + \frac{\partial h_2}{\partial h_1}\frac{\partial h_1}{\partial w_i}$$

$$= h_2(1-h_2)\times x_1 + h_2(1-h_2)\times w_h\times h_1(1-h_1)x_0$$

$$\cancel{\frac{\partial J}{\partial w_i} = -a x_0(g_1-y_1)(h_1)(1-h_1) + x_1 h_2(1-h_2) + h_2(1-h_2)}$$

$$\frac{\partial J}{\partial w_i} = -a x_0 (g_1-y_1)(h_1)(1-h_1) + h_2(1-h_2)\left(x_1 + w_h h_1(1-h_1)x_0\right)$$

——— (15)

e) $\frac{\partial J}{\partial w_h} = \frac{\partial J_1}{\partial w_h} + \frac{\partial J_2}{\partial w_h}$

$$\frac{\partial J_1}{\partial w_h} = \frac{\partial J_1}{\partial y_1}\frac{\partial y_1}{\partial h_1}\frac{\partial h_1}{\partial w_h} = -(g_1-y_1)\times a\times h_1(1-h_1)\times h_0$$

$$\frac{\partial J_2}{\partial w_h} = \frac{\partial J_2}{\partial y_2}\,\frac{\partial y_2}{\partial h_2}\,\frac{\partial h_2}{\partial w_h}$$

$$\frac{\partial h_2}{\partial w_h} = \frac{\partial h_2}{\partial w_h} + \frac{\partial h_2}{\partial h_1}\,\frac{\partial h_1}{\partial w_h}$$

$$= h_2(1-h_2)h_1 + h_2(1-h_2)\,w_h \times h_1(1-h_1)h_0$$

$$\therefore \boxed{\frac{\partial J}{\partial w_h} = -a(\hat{y}_1 - y_1)(h_1)(1-h_1)h_0 + h_2 h_1 (1-h_2)\left(1 + w_h(1-h_1)h_0\right)}$$

— (16)

Now let's compute all the values—

$h_1 =$ sigmoid (2) —— from equation (2)

$$\therefore h_1 = \frac{1}{1+e^{-2}}$$

$$\boxed{h_1 = 0.88}$$

similarly, $h_2 =$ sigmoid$(1 + h_1)$

$$h_2 = \frac{1}{1+e^{-1.88}}$$

$$\therefore \boxed{h_2 = 0.87}$$

$y_1 = 0.88 + 1 = 1.88$

$\boxed{y_1 = 1.88}$ & $\boxed{y_2 = 1.87}$

a

$\therefore$ Now let's calculate $\frac{\partial J}{\partial a}$ using (12)

$$\therefore \frac{\partial J}{\partial a} = -h_1(\hat{y}_1 - y_1) + -h_2(\hat{y}_2 - y_2)$$

$$= (y_1 - \hat{y}_1)h_1 + h_2(y_2 - \hat{y}_2)$$

$$= (1.88 - 1)0.88 + 0.87(1.87 - 1)$$

$$\boxed{\frac{\partial J}{\partial a} = 1.53}$$

$$\therefore \frac{\partial J}{\partial b} = -a\left[(\hat{y}_1 - y_1)h_1(1-h_1) + (\hat{y}_2 - y_2)h_2(1-h_2)(1 + w_h h_1(1-h_1))\right]$$

— from (13)

$\therefore$ Let's substitute values to get —

$$\frac{\partial J}{\partial b} = -1\left[(-0.88)\,0.88(0.12) + (-0.87)\,0.87(0.13)(1 + 0.88 \times 0.12)\right]$$

$$\boxed{\frac{\partial J}{\partial b} = +0.2}$$

$$\frac{\partial J}{\partial c} = -\left[(\hat{y}_1 - y_1) + (\hat{y}_2 - y_2)\right] \quad \text{— from (14)}$$

$$= -\left[(1 - 1.88) + (1 - 1.87)\right]$$

$$= +\left[0.88 + 0.87\right]$$

$$\boxed{\frac{\partial J}{\partial c} = +1.75}$$

classmate

Date

Page 7

$$\frac{\partial J}{\partial w_i} = -a x_0 (\hat{y}_1 - y_1)\, h_1 (1-h_1) + h_2(1-h_2)\left(x_1 + w_h h_1 (1-h_1) x_0\right)$$

— from (15)

$$\therefore \frac{\partial J}{\partial w_i} = (0.88)\,(0.88)\,(0.12) + 0.87(0.13)\left(0.88 \times 0.12\right)$$

$$\boxed{\frac{\partial J}{\partial w_i} = 0.1}$$

$$\frac{\partial J}{\partial w_h} = -a(\hat{y}_1 - y_1)\, h_1(1-h_1)\, h_0 + h_2 h_1 (1-h_2)\left[1 + w_h (1-h_1) h_0\right]$$

— from (16)

$= \cancel{0.88 \times 0.88 \times 0.12}\ 0.88 \times 0.87 \times 0.12$

$\cancel{= 0.09}$

$$\boxed{\frac{\partial J}{\partial w_h} = 0.09}$$

We can clearly see that gradients are vanishing.