Patrick Nitschke DL-Exam-2021 18/06/21

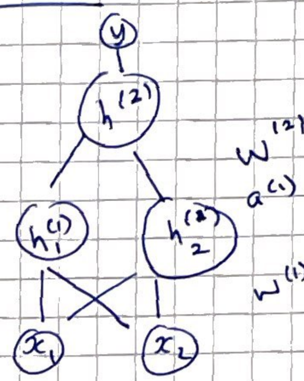
1) $h^{(1)} = W^{(1)} x + b^{(1)}$ Find $\frac{\partial J}{\partial W^{(1)}_{1,2}}$.

$a^{(1)} = \mathrm{ReLU}(h^{(1)})$

$h^{(2)} = (w^{(2)})^T a^{(1)} + b^{(2)}$

$y = \sigma(h^{(2)})$

$J = \frac{1}{2}(t-y)^2$



σ

ReLU

$$\frac{\partial J}{\partial w^{(1)}} = \underbrace{\frac{\partial J}{\partial y}}_{①} \underbrace{\frac{\partial y}{\partial h^{(2)}}}_{②} \underbrace{\frac{\partial h^{(2)}}{\partial a^{(1)}}}_{③} \underbrace{\frac{\partial a^{(1)}}{\partial h^{(1)}}}_{④} \underbrace{\frac{\partial h^{(1)}}{\partial w^{(1)}}}_{⑤}$$

Dimensions

① $J, y, \frac{\partial J}{\partial y}$: 1×1

② ~~h_2~~ $h^{(2)}$ is $\mathbb{R}$, $\frac{\partial y}{\partial h}$: 1×1

③ $a^{(1)}_r$: 2×1, $\frac{\partial h^{(2)}}{a^{(1)}} = 1 \times 2$

④ $\overset{a_{1,2}}{2} \times (\overset{h_{1,2}}{2 \times 1}) \Rightarrow 2 \times 2$

⑤ $\underset{h^{(1)}}{2} \times \underset{W}{(2 \times 2)}$.

① $\frac{\partial J}{\partial y} = -(t-y)$

② $\frac{\partial y}{\partial h^{(2)}} = \sigma(h^{(2)})(1 - \sigma(h^{(2)}))$

③ $\frac{\partial h^{(2)}}{\partial a^{(1)}} = w^{(2)T}$

④ $h^{(1)} = \begin{bmatrix} 0.75 & 0.5 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0.5 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

$\mathrm{ReLU}(x) = \max\{0, x\}$

$= \begin{bmatrix} 0.75 + 0.25 \\ -1 + 0.5 \end{bmatrix} = \begin{bmatrix} 1 \\ -0.5 \end{bmatrix}$

~~$\frac{\partial a^{(1)}}{\partial h^{(1)}} = \begin{bmatrix} \begin{cases} 1 & \text{if } h_1^{(1)} \geq 0 \\ 0 & \text{else} \end{cases} \\ \begin{cases} 1 & \text{if } h_2^{(1)} \geq 0 \\ 0 & \text{else} \end{cases} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$~~

Not 2×2

④ $$\frac{\partial a^{(1)}}{\partial h^{(1)}} = \begin{bmatrix} \frac{\partial a_1^{(1)}}{\partial h_1^{(1)}} & \frac{\partial a_1^{(1)}}{\partial h_2^{(1)}} \\ \frac{\partial a_2^{(1)}}{\partial h_2^{(1)}} & \frac{\partial a_2^{(1)}}{\partial h_2^{(1)}} \end{bmatrix}$$

Since $h_1 = \begin{bmatrix} 1 \\ -0.5 \end{bmatrix}$

$$= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

⑤ $$\frac{\partial h^{(1)}}{\partial w^{(1)}} = \begin{bmatrix} \frac{\partial h_1^{(1)}}{\partial w^{(1)}} \\ \frac{\partial h_2^{(1)}}{\partial w^{(1)}} \end{bmatrix}$$

2×2

2×2

$\therefore$ We see that $\frac{\partial a^{(1)}}{\partial h^{(1)}}$ has 0's everywhere except first element, only need to find $\frac{\partial h_1^{(1)}}{\partial w^{(1)}}$

$\cancel{\frac{\partial h_1^{(1)}}{\partial w^{(1)}}}$

$$h^{(1)} = \begin{bmatrix} w_{1,1}^{(1)} & w_{1,2}^{(1)} \\ w_{2,1}^{(1)} & w_{2,2}^{(1)} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$= \begin{bmatrix} w_{1,1}^{(1)} x_1 + w_{1,2}^{(1)} x_2 \\ w_{2,1}^{(1)} x_1 + w_{2,2}^{(1)} x_2 \end{bmatrix} = \begin{bmatrix} h_1^{(1)} \\ h_2^{(1)} \end{bmatrix}$$

$$\frac{\partial h_1^{(1)}}{\partial w^{(1)}} = \begin{bmatrix} x_1 & x_2 \\ 0 & 0 \end{bmatrix}$$

$\frac{\partial h_1^{(1)}}{\partial w_{2,1}^{(1)}} = 0$

Patrick Nitschke   DL-Exam 2021   18/6/21

Now that gradients are computed, just sum:

abuse of notation

$$\frac{\partial J}{\partial W^{(1)}} = -(t-y)\,\sigma(h^2)(1-\sigma(h^2))\,w^{(2)T} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \begin{bmatrix} x_1 & x_2 \\ 0 & 0 \end{bmatrix} \\ 0 \end{bmatrix}$$

$$h^{(2)} = \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} + (-1) = 0$$

$\sigma(0) = 0.5$

$t = 1$, $y = \sigma(0) = 0.5$

$w^{(2)} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$x = \begin{bmatrix} 1 \\ 0.5 \end{bmatrix}$

$$\frac{\partial J}{\partial W^{(1)}} = -(0.5)\,\frac{1}{2}\left(1-\frac{1}{2}\right) \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \begin{bmatrix} x_1 & x_2 \\ 0 & 0 \end{bmatrix} \\ 0 \end{bmatrix}$$

$$= -\frac{1}{8} \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} \begin{bmatrix} x_1 & x_2 \\ 0 & 0 \end{bmatrix} \\ 0 \end{bmatrix}$$

$$= -\frac{1}{8} \begin{bmatrix} x_1 & x_2 \\ 0 & 0 \end{bmatrix} \Rightarrow \frac{\partial J}{\partial W^{(1)}_{1,2}} = -\frac{1}{8} x_2 = -\frac{1}{16}$$

$0.125/2 = 0.0625$

~~Question 2~~

$$E(v,h) = -b^T v - c^T h - v^T W h$$

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

$p(x \mid \theta) = ?$

$$\frac{\partial \log p(x \mid \theta)}{\partial c_z} = \sum_{i=1}^{n} \sigma(c_z + v^T W_{:,z}) - \sum_{v} p(v)\,\sigma(c_z + v^T W_{:,z})$$