Question 2 of Assignment 1

xxxx

October 2021

1 a)

$a^{(3)} = (\psi(z^{(3)})$

$$\frac{\partial J}{\partial z^{(3)}} = \left(\frac{\partial J}{\partial a^{(3)}} \frac{\partial a^{(3)}}{\partial z^{(3)}} \right)$$

$$\frac{\partial a^{(3)}}{\partial z^{(3)}} = \left(\frac{\partial \psi(z^{(3)})}{\partial z^{(3)}} \right) = \psi(z^{(3)})(1 - \psi(z^{(3)})$$

$$\frac{\partial J}{\partial a^{(3)}} = -\frac{\partial \frac{1}{N} \log a^{(3)}}{\partial a^{(3)}} = -\frac{1}{N a^{(3)}}$$

$$\frac{\partial J}{\partial z^{(3)}} = \left(-\frac{(1 - \psi(z^{(3)}))}{N a^{(3)}} \psi(z^{(3)}) \right)$$

$$\frac{\partial J}{\partial z^{(3)}} = \left(\frac{1}{N} (\psi(z^{(3)}) - 1) \right)$$

$$\frac{\partial J}{\partial z^{(3)}} = \left(\frac{1}{N} (\psi(z^{(3)}) - \Delta_i) \right)$$

2 b)

$$\frac{\partial \tilde{J}}{\partial W^{(2)}} = \frac{\partial J}{\partial W^{(2)}} + \frac{\partial \lambda(\|W^{(1)}\|_2^{2+} \|W^{(2)}\|_2^2)}{\partial W^{(2}}$$

$$\frac{\partial J}{\partial W^{(2)}} = \frac{\partial J}{\partial z^{(3)}} \frac{\partial z^{(3)}}{\partial W^{(2)}}$$

$$\frac{\partial J}{\partial z^{(3)}} = \left(\frac{1}{N} (\psi(z^{(3)}) - \Delta_i) \right)$$

$$\frac{\partial z^{(3)}}{\partial W^{(2)}} = a_i^{(2)}$$

$$\frac{\partial \tilde{J}}{\partial W^{(2)}} =$$

$\big(\frac{1}{N}(\psi(z^{(3)}) - \Delta_i))a_i^2 \ + 2\lambda W^{(2)}$

3 c)

1.

$$\tilde{J} = \sum(\psi(z^{(3)}) + \lambda(W(1)^{2+}W(2)^2)$$

$$\frac{\partial \tilde{J}}{\partial W^{(1)}} = \frac{\partial \tilde{J}}{\partial z^{(3)}}\frac{\partial z^{(3)}}{\partial a^{(2)}}\frac{\partial a^{(2)}}{\partial z^{(2)}}\frac{\partial z^{(2)}}{\partial W^{(1)}}$$

$$\frac{\partial \tilde{J}}{\partial z^{(3)}} = \left(\frac{1}{N}(\psi(z^{(3)}) - \Delta_i)\right)$$

$$\frac{\partial z^{(3)}}{\partial a^{(2)}} = W(2)^T$$

$$\frac{\partial a^{(2)}}{\partial z^{(2)}} = \frac{\partial \phi(z^{(2)})}{\partial z^{(2)}}$$

ReLU derivative is $\begin{cases} 0 & z^{(2)} < 0 \\ 1 & z^{(2)} \geq 0 \end{cases}$

$$\frac{\partial z^{(2)}}{\partial W^{(1)}} = a(1)^T$$

Therefore:

$$\tilde{J} = \big[\frac{1}{N}\sum(\psi(z^{(3)}) - \Delta_i) + 2\lambda W^{(1)}\big]a(1)^T\frac{\partial a^{(2)}}{\partial z^{(2)}}W(2)^T$$

2.

$$\frac{\partial \tilde{J}}{\partial b^{(1)}} = \frac{\partial \tilde{J}}{\partial z^{(3)}}\frac{\partial z^{(3)}}{\partial a^{(2)}}\frac{\partial a^{(2)}}{\partial z^{(2)}}\frac{\partial z^{(2)}}{\partial b^{(1)}} = \big[\frac{1}{N}\sum(\psi(z^{(3)}) - \Delta_i\big]W(2)^T\frac{\partial \phi(z^{(2)})}{\partial z^{(2)}}$$

3.

$$\frac{\partial \tilde{J}}{\partial b^{(2)}} = \frac{\partial \tilde{J}}{\partial z^{(3)}}\frac{\partial z^{(3)}}{\partial b^{(2)}}$$

since

$$\frac{\partial z^{(3)}}{\partial b^{(2)}} = 1$$

$$\frac{\partial \tilde{J}}{\partial b^{(2)}} = \big[\frac{1}{N}\sum(\psi(z^{(3)}) - \Delta_i\big]$$