Q2

April 28, 2024

1 a)

$$z_1 = w^{(1)} \cdot x + b^{(1)}$$
$$= \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 1 \end{bmatrix} + \begin{bmatrix} -2 \\ -1 \\ 1 \end{bmatrix}$$
$$= \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} + \begin{bmatrix} -2 \\ -1 \\ 1 \end{bmatrix}$$
$$= \begin{bmatrix} -1 \\ 2 \\ 3 \end{bmatrix}$$

———

$$h_1 = ReLU(z_1)$$
$$= ReLU(\begin{bmatrix} -1 \\ 2 \\ 3 \end{bmatrix})$$
$$= \begin{bmatrix} 0 \\ 2 \\ 3 \end{bmatrix}$$

———

$$z_2 = w^{(2)} \cdot h_1 + b^{(2)}$$
$$= \begin{bmatrix} 0 & 1 & 2 \\ 1 & 1 & 1 \\ 2 & -1 & -1 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 2 \\ 3 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$
$$= \begin{bmatrix} 8 \\ 5 \\ -5 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$
$$= \begin{bmatrix} 9 \\ 6 \\ -4 \end{bmatrix}$$

—

$$\begin{aligned}\hat{y} &= w^{(3)} \cdot z_2 \\ &= \begin{bmatrix}-1 & 1 & -2\end{bmatrix} \cdot \begin{bmatrix}9 \\ 6 \\ -4\end{bmatrix} = 5\end{aligned}$$

———

$$\begin{aligned}L &= (y-\hat{y})^2/2 \\ &= (1.5-5)^2/2 \\ &= 6.125\end{aligned}$$

2 b)

$$\begin{aligned}L &= (y-\hat{y})^2/2 \\ &= (y - w^{(3)} \cdot z_2)^2/2 \\ \frac{\partial L}{\partial w_2^{(3)}} &= \frac{\partial L}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial w_2^{(3)}} \\ &= -(y-\hat{y}) \cdot \frac{\partial (w^{(3)} \cdot z_2)}{\partial w_2^{(3)}} \\ &= -(y-\hat{y}) \cdot \frac{\partial (w_1^{(3)} z_{21} + w_2^{(3)} z_{22} + w_3^{(3)} z_{23})}{w_2^{(3)}} \\ &= -(y-\hat{y}) \cdot z_{22} \\ &= -(1.5-5) \cdot 6 \\ &= 21\end{aligned}$$

———

$$\begin{aligned}\frac{\partial L}{\partial w_{22}^{(2)}} &= \frac{\partial L}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial z_2} \cdot \frac{\partial z_2}{\partial w_{22}^{(2)}} \\ &= -(y-\hat{y}) \cdot w^{(3)} \cdot \frac{\partial \begin{bmatrix} w_{11}^{(2)} h_{11} + w_{12}^{(2)} h_{12} + w_{13}^{(2)} h_{13} + b_1^{(2)} \\ w_{21}^{(2)} h_{11} + w_{22}^{(2)} h_{12} + w_{23}^{(2)} h_{13} + b_2^{(2)} \\ w_{31}^{(2)} h_{11} + w_{32}^{(2)} h_{12} + w_{33}^{(2)} h_{13} + b_2^{(2)} \end{bmatrix}}{\partial w_{22}^{(2)}} \\ &= -(y-\hat{y}) \cdot w^{(3)} \cdot \begin{bmatrix}0 \\ h_{12} \\ 0\end{bmatrix} \\ &= -(1.5-5) \cdot \begin{bmatrix}-1 & 1 & -2\end{bmatrix} \cdot \begin{bmatrix}0 \\ 2 \\ 0\end{bmatrix} \\ &= 7\end{aligned}$$

3 c)

$$\frac{\partial L}{\partial w_{21}^{(1)}} = \frac{\partial L}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial z_2} \cdot \frac{\partial z_2}{\partial h_1} \cdot \frac{\partial h_1}{\partial w_{21}^{(1)}}$$

$$= -(y - \hat{y}) \cdot w^{(3)} \cdot w^{(2)} \cdot \begin{bmatrix} \frac{\partial (w_{11}^{(1)} x_1 + w_{12}^{(1)} x_2 + b_1^{(1)})}{\partial w_{21}^{(1)}} \; if \; z_{11} \; > \; 0 \; else \; 0 \\ \frac{\partial (w_{21}^{(1)} x_1 + w_{22}^{(1)} x_2 + b_2^{(1)})}{\partial w_{21}^{(1)}} \; if \; z_{12} \; > \; 0 \; else \; 0 \\ \frac{\partial (w_{31}^{(1)} x_1 + w_{32}^{(1)} x_2 + b_3^{(1)})}{\partial w_{21}^{(1)}} \; if \; z_{13} \; > \; 0 \; else \; 0 \end{bmatrix}$$

$$= -(1.5 - 5) \cdot \begin{bmatrix} -1 & 1 & -2 \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 & 2 \\ 1 & 1 & 1 \\ 2 & -1 & -1 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}$$

$$= 14$$