CSC321 Homework 3 Ruijie Sun 1003326046

1. $W^{(1)} = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix}$ $\quad b^{(1)} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$

$W^{(2)} = (-1, -1, -1)$ $\quad b^{(2)} = 0.5$

2. ①

$$\begin{tikzcd} x \arrow[r] \arrow[rrr, bend left] & z \arrow[r] & h \arrow[r] \arrow[dr] & y \arrow[r] & s \arrow[r] & \mathcal{E} \\ & & & R \arrow[urr] & & \end{tikzcd}$$

②

$\bar{\mathcal{E}} = 1$

$\bar{s} = \bar{\mathcal{E}} \cdot 1 = 1$

$\bar{y} = \bar{s} \cdot (y - s) = y - s$

$\bar{R} = \bar{\mathcal{E}} \cdot \frac{d\mathcal{E}}{dR} = \bar{\mathcal{E}} \cdot 1 = 1$

$\bar{h} = W^{(2)T} \bar{y} + \gamma \cdot \bar{R} = W^{(2)T} \bar{y} + r$

$\bar{z} = \bar{h} \circ \sigma'(z)$

$\bar{x} = W^{(1)T} \bar{z} + \bar{y}$

3. $\frac{\partial L}{\partial w_1}$ Yes, $\frac{\partial L}{\partial w_2}$ Yes, $\frac{\partial L}{\partial w_3}$ No.

Since the meaning of $\frac{\partial L}{\partial w_i}$ is that the measure in change of L when we make an infinitesimal change to w_i while keeping others fixed. So, if we make an infinitesimal change to w_1, the L will not change since $w_1 \times$ output of h_1 is always equal to 0 based on the Relu activation function. If we make infinitesimal change to w_2, the $w_2 \times$ output of h_3 is still close to -1, the output of h_1 is still 0. So there is no change to L. However, if we make an infinitesimal change to w_3, it might change y since we don't know the activation function of h_2 and h_3 and h_2 is connected to h_3 as well. So based on this case, L may change.