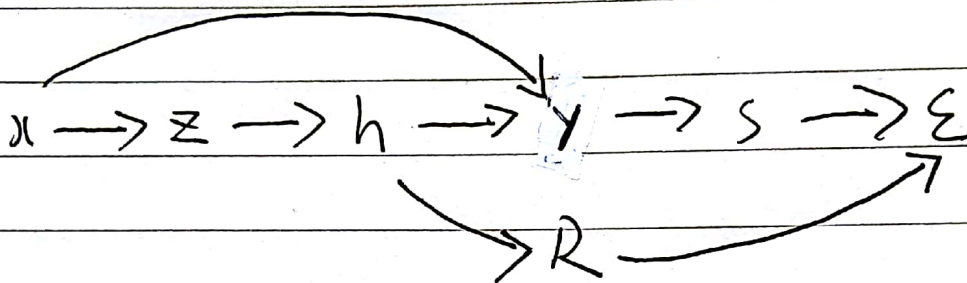


CSC321 Homework 3 Ruijie Sun 1003326046

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

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

2. ①



$$② \quad \bar{\varepsilon} = 1$$

$$\bar{s} = \bar{\varepsilon} \cdot 1 = 1$$

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

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

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

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

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

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_2 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.