CSC321: Assignment #3

Xiangyu Kong
kongxi16

January 29, 2018

Problem 1

$$\mathbf{W}^{(1)} = \begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix}$$

$$\mathbf{b}^{(1)} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

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

$$b^{(2)} = -2.5$$

Problem 2

1. As shown in Figure 1

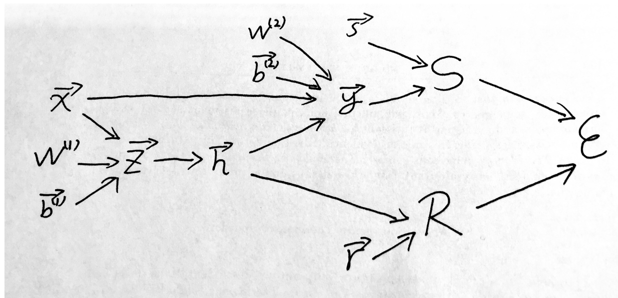

Figure 1: 2.1

2.

$$
\begin{aligned}
\overline{\mathcal{E}} &= 1 \\
\overline{\mathcal{S}} &= \overline{\mathcal{E}}\frac{d\mathcal{E}}{d\mathcal{S}} = 1 \\
\overline{\mathcal{R}} &= \overline{\mathcal{E}}\frac{d\mathcal{E}}{d\mathcal{R}} = 1 \\
\overline{\mathbf{y}} &= \overline{\mathcal{S}}\frac{d\mathcal{S}}{d\mathbf{y}} = 1||\mathbf{y}-\mathbf{s}|| = ||\mathbf{y}-\mathbf{s}|| \\
\overline{\mathbf{h}} &= \overline{\mathbf{y}}\frac{d\mathbf{y}}{d\mathbf{h}} + \overline{\mathcal{R}}\frac{d\mathcal{R}}{d\mathbf{h}} = ||\mathbf{y}-\mathbf{s}||\mathbf{W}^{(2)'} + \mathbf{r}^{\top'} \\
\overline{\mathbf{z}} &= \overline{\mathbf{h}}\frac{d\mathbf{h}}{d\mathbf{z}} = (||\mathbf{y}-\mathbf{s}||\mathbf{W}^{(2)'} + \mathbf{r}^{\top'})\sigma'(\mathbf{z}) \\
\overline{\mathbf{x}} &= \overline{\mathbf{z}}\frac{d\mathbf{z}}{d\mathbf{x}} + \overline{\mathbf{y}}\frac{d\mathbf{y}}{d\mathbf{x}} = (||\mathbf{y}-\mathbf{s}||\mathbf{W}^{(2)'} + \mathbf{r}^{\top'})\sigma'(\mathbf{z})\mathbf{W}^{(1)'} + ||\mathbf{y}-\mathbf{s}||
\end{aligned}
$$

Problem 3

1. $\dfrac{\partial \mathcal{L}}{\partial w_1}$: YES
   Using back-propagation:

$$\begin{aligned}
\overline{\mathcal{L}} &= 1 \\
\overline{y} &= \overline{\mathcal{L}}\frac{d\mathcal{L}}{dy} = f'(y) \\
\overline{w_1} &= \overline{y}\frac{dy}{dw_1} = f'(y)0 = 0
\end{aligned}$$

2. $\dfrac{\partial \mathcal{L}}{\partial w_2}$: NO
   Using back-propagation:

$$\begin{aligned}
\overline{\mathcal{L}} &= 1 \\
\overline{y} &= \overline{\mathcal{L}}\frac{d\mathcal{L}}{dy} = f'(y) \\
\overline{h_1} &= \overline{y}\frac{dy}{dh_1} = f'(y)w_1 \\
\overline{w_2} &= \overline{h_1}\frac{dh_1}{dh_1} = f'(y)w_1h_3
\end{aligned}$$

   If $f'(y)w_1h_3 \neq 0$, $\dfrac{\partial \mathcal{L}}{\partial w_2}$ does not necessarily equal to 0.

3. $\dfrac{\partial \mathcal{L}}{\partial w_3}$: NO
   Using back-propagation:

$$\begin{aligned}
\overline{\mathcal{L}} &= 1 \\
\overline{y} &= \overline{\mathcal{L}}\frac{d\mathcal{L}}{dy} = f'(y) \\
\overline{h_1} &= \overline{y}\frac{dy}{dh_1} = f'(y)w_1 \\
\overline{h_3} &= \overline{h_1}\frac{dh_1}{dh_3} = f'(y)w_1w_3 \\
\overline{w_3} &= \overline{h_3}\frac{dh_3}{dw_3} = f'(y)w_1w_3x_1
\end{aligned}$$

   If $f'(y)w_1w_3x_1 \neq 0$, $\dfrac{\partial \mathcal{L}}{\partial w_3}$ does not necessarily equal to 0.