CSC421: Written Homework 3

Due on Thrusday, Mar 7, 2019

Zhongtian Ouyang
1002341012

Problem 1

Dropout

(a)

$$\begin{aligned}
E(y) &= E[\sum_j m_j w_j x_j] \\
&= E[m_1 w_1 x_1 + m_2 w_2 x_2 + ... + m_J w_J x_J] \\
&= E[m_1 w_1 x_1] + ... + E[m_J w_J x_J] \\
&= w_1 x_1 E[m_1] + ... + w_J x_J E[m_J] \\
&= w_1 x_1 (\frac{1}{2}) + ... + m_J w_J (\frac{1}{2}) = \frac{1}{2} \sum_j w_j x_j = \frac{1}{2} \mathbf{w} \cdot \mathbf{x}
\end{aligned} \tag{1}$$

$$\begin{aligned}
Var(y) &= Var(\sum_j m_j w_j x_j) \\
&= \sum_j Var(m_j w_j x_j),\ \textit{since they are independent} \\
&= \sum_j w_j^2 x_j^2 Var(m_j) \\
&= \sum_j w_j^2 x_j^2 (E[m_j^2] - E[m_j]^2) \\
&= \sum_j w_j^2 x_j^2 (\frac{1}{2} - \frac{1}{4}),\ E[m_j^2] = E[m_j]\ \textit{because}\ m_j\ \textit{is either}\ 0\ \textit{or}\ 1,\ m_j = m_j^2 \\
&= \frac{1}{4} \sum_j w_j^2 x_j^2 = \frac{1}{4} (\mathbf{w} \cdot \mathbf{w}) \cdot (\mathbf{x} \cdot \mathbf{x})
\end{aligned} \tag{2}$$

(b)
Since $E(y) = \tilde{y} = \frac{1}{2} \sum_j w_j x_j = \sum_j (\frac{1}{2} w_j) x_j$, $\tilde{w_j} = \frac{1}{2} w_j$.

(c)

$$
\begin{aligned}
J &= \frac{1}{2N}\sum_{i=1}^{N} E[(y^{(i)} - t^{(i)})^2] \\
&= \frac{1}{2N}\sum_{i=1}^{N} E[{y^{(i)}}^2 - 2y^{(i)}t^{(i)} + {t^{(i)}}^2] \\
&= \frac{1}{2N}\sum_{i=1}^{N} (E[{y^{(i)}}^2] - E[2y^{(i)}t^{(i)}] + E[{t^{(i)}}^2]) \\
&= \frac{1}{2N}\sum_{i=1}^{N} ((Var[y^{(i)}] + E[y^{(i)}]^2) - 2t^{(i)}E[y^{(i)}] + {t^{(i)}}^2) \\
&= \frac{1}{2N}\sum_{i=1}^{N} ((\frac{1}{4}\sum_j w_j^2 x_j^2 + (\frac{1}{2}\sum_j w_j x_j^{(i)})^2) - 2t^{(i)}\frac{1}{2}\sum_j w_j x_j^{(i)} + {t^{(i)}}^2) \\
&= \frac{1}{2N}\sum_{i=1}^{N} ((\frac{1}{4}\sum_j (2\tilde{w}_j)^2 x_j^2 + (\frac{1}{2}\sum_j 2\tilde{w}_j x_j^{(i)})^2) - 2t^{(i)}\frac{1}{2}\sum_j 2\tilde{w}_j x_j^{(i)} + {t^{(i)}}^2) \\
&= \frac{1}{2N}\sum_{i=1}^{N} (\sum_j {\tilde{w}_j}^2 x_j^2 + (\sum_j \tilde{w}_j x_j^{(i)})^2 - 2t^{(i)}\sum_j \tilde{w}_j x_j^{(i)} + {t^{(i)}}^2) \\
&= \frac{1}{2N}\sum_{i=1}^{N} (\sum_j {\tilde{w}_j}^2 x_j^2 + (\tilde{y}^{(i)})^2 - 2t^{(i)}\tilde{y}^{(i)} + {t^{(i)}}^2) \\
&= \frac{1}{2N}\sum_{i=1}^{N} ((\tilde{y}^{(i)})^2 - 2t^{(i)}\tilde{y}^{(i)} + {t^{(i)}}^2) + \frac{1}{2N}\sum_{i=1}^{N}\sum_j {\tilde{w}_j}^2 x_j^2 \\
&= \frac{1}{2N}\sum_{i=1}^{N} (\tilde{y}^{(i)} - t^{(i)})^2 + R(\tilde{w}_1, ..., \tilde{w}_D)
\end{aligned}
\tag{3}
$$

Problem 2

RNN

$$U = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix}$$

$$V = \begin{bmatrix} 1 & -1 & 2 \end{bmatrix}$$

$$W = \begin{bmatrix} h_1^{(t)} \to h_1^{(t+1)} & h_2^{(t)} \to h_1^{(t+1)} & h_3^{(t)} \to h_1^{(t+1)} \\ h_1^{(t)} \to h_2^{(t+1)} & h_2^{(t)} \to h_2^{(t+1)} & h_3^{(t)} \to h_2^{(t+1)} \\ h_1^{(t)} \to h_3^{(t+1)} & h_2^{(t)} \to h_3^{(t+1)} & h_3^{(t)} \to h_3^{(t+1)} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

$$b_h = \begin{bmatrix} -0.5 \\ -1.5 \\ -2.5 \end{bmatrix}$$

$$b_y = -0.5$$