521153S assignment 2 - Derivative notes for reference only

Lam Huynh

November 11, 2020

1 Our model

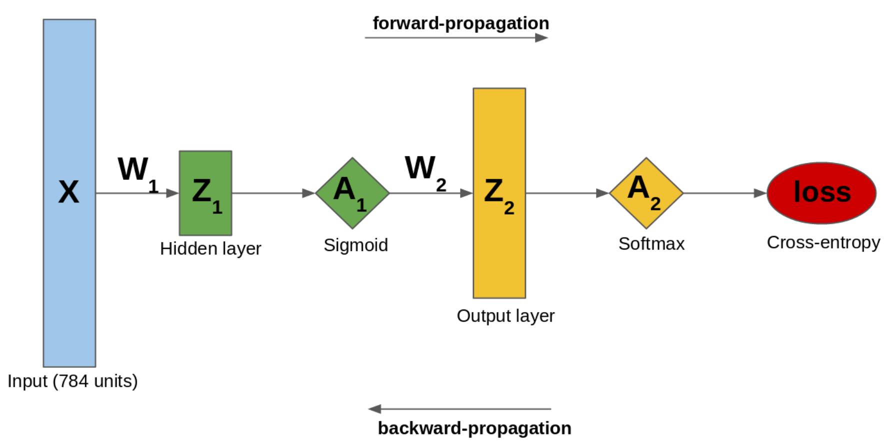

Figure 1: Structure of our network.

From assignment 2, we have:

$$\frac{\partial L}{\partial W_{2i}} = \frac{\partial L}{\partial A_{2i}} \cdot \frac{\partial A_{2i}}{\partial Z_{2i}} \cdot \frac{\partial Z_{2i}}{\partial W_{2i}} \tag{6}$$

The derivative of A_{2i} with respect to Z_{2i} is

$$\frac{\partial A_{2i}}{\partial Z_{2j}} = \frac{\partial \frac{e^{Z_{2i}}}{\sum_{l=0}^{k} e^{Z_{2l}}}}{\partial Z_{2j}}$$

If $i = j$, we have:

$$\begin{aligned}
\frac{\partial \frac{e^{Z_{2i}}}{\sum_{l=0}^{k} e^{Z_{2l}}}}{\partial Z_{2j}} &= \frac{e^{Z_{2i}} \sum_{l=0}^{k} e^{Z_{2l}} - e^{Z_{2i}} e^{Z_{2j}}}{(\sum_{l=0}^{k} e^{Z_{2l}})^2} \\
&= \frac{e^{Z_{2i}} (\sum_{l=0}^{k} e^{Z_{2l}} - e^{Z_{2j}})}{(\sum_{l=0}^{k} e^{Z_{2l}})^2} \\
&= \frac{e^{Z_{2i}}}{\sum_{l=0}^{k} e^{Z_{2l}}} \cdot \frac{(\sum_{l=0}^{k} e^{Z_{2l}} - e^{Z_{2j}})}{\sum_{l=0}^{k} e^{Z_{2l}}} \\
&= \frac{e^{Z_{2i}}}{\sum_{l=0}^{k} e^{Z_{2l}}} \cdot (1 - \frac{e^{Z_{2j}}}{\sum_{l=0}^{k} e^{Z_{2l}}})
\end{aligned} \tag{19}$$

We have:

$$A_2 = \frac{e^{Z_2}}{\sum_{l=0}^{k} e^{Z_{2l}}}$$

Therefore Eq.(19) become:

$$\frac{\partial \frac{e^{Z_{2i}}}{\sum_{l=0}^{k} e^{Z_{2l}}}}{\partial Z_{2j}} = A_{2i} \cdot (1 - A_{2j}) \tag{20}$$

If i $\neq$ j, we have:

$$\begin{aligned}
\frac{\partial \frac{e^{Z_{2i}}}{\sum_{l=0}^{k} e^{Z_{2l}}}}{\partial Z_{2j}} &= \frac{0 - e^{Z_{2i}} e^{Z_{2j}}}{(\sum_{l=0}^{k} e^{Z_{2l}})^2} \\
&= \frac{-e^{Z_{2i}}}{\sum_{l=0}^{k} e^{Z_{2l}}} \cdot \frac{e^{Z_{2j}}}{\sum_{l=0}^{k} e^{Z_{2l}}} \\
&= -A_{2i} A_{2j}
\end{aligned} \tag{21}$$

From assignment 2, we have the derivative of L with respect to Z_{2i} is

$$\begin{aligned}
\frac{\partial L}{\partial Z_{2i}} &= -\sum_j Y_j \frac{\partial log(A_{2j})}{\partial Z_{2i}} \\
&= -\sum_j Y_j \frac{\partial log(A_{2j})}{\partial A_{2j}} \cdot \frac{\partial A_{2j}}{\partial Z_{2i}} \\
&= -\sum_j Y_j \frac{1}{A_{2j}} \cdot \frac{\partial A_{2j}}{\partial Z_{2i}}
\end{aligned}$$

Substitute with Eq.(20) and (21) we have:

$$
\begin{aligned}
\frac{\partial L}{\partial Z_{2i}} &= -Y_i \frac{1}{A_{2i}} A_{2i}(1 - A_{2i}) - \sum_{j \neq i} Y_j \frac{1}{A_{2j}} (-A_{2i} A_{2j}) \\
&= -Y_i(1 - A_{2i}) + \sum_{j \neq i} Y_j A_{2i} \\
&= -Y_i + Y_i A_{2i} + \sum_{j \neq i} Y_j A_{2i} \\
&= A_{2i}(Y_i + \sum_{j \neq i} Y_j) - Y_i \\
&= A_{2i} \sum_j Y_j - Y_i
\end{aligned}
$$

we have $\sum_j Y_j = 1.0$ so

$$
\frac{\partial L}{\partial Z_{2i}} = A_{2i} - Y_i \tag{8}
$$

The derivative of Z_{2i} with respect to W_{2i} is

$$
\frac{\partial Z_{2i}}{\partial W_{2i}} = A_{1i} \tag{9}
$$

Hence Eq.(6) become:

$$
\frac{\partial L}{\partial W_{2i}} = \frac{\partial L}{\partial A_{2i}} \cdot \frac{\partial A_{2i}}{\partial Z_{2i}} \cdot \frac{\partial Z_{2i}}{\partial W_{2i}} = (A_{2i} - Y_i) A_{1i} \tag{6}
$$

Derive equation (7), we have:

$$
\frac{\partial L}{\partial W_{1i}} = \frac{\partial L}{\partial A_{2i}} \cdot \frac{\partial A_{2i}}{\partial Z_{2i}} \cdot \frac{\partial Z_{2i}}{\partial A_{1i}} \cdot \frac{\partial A_{1i}}{\partial Z_{1i}} \cdot \frac{\partial Z_{1i}}{\partial W_{1i}} \tag{7}
$$

The derivative of Z_{2i} with respect to A_{1i} is

$$
\frac{\partial Z_{2i}}{\partial A_{1i}} = W_{2i}
$$

The derivative of L with respect to A_{1i} is

$$
\frac{\partial L}{\partial A_{1i}} = \frac{\partial L}{\partial Z_{2i}} \cdot \frac{\partial Z_{2i}}{\partial A_{1i}} = (A_{2i} - Y_i) W_{2i} \tag{10}
$$

The derivative of A_{1i} with respect to Z_{1i}:

$$\begin{aligned}
\frac{\partial A_{1i}}{\partial Z_{1i}} &= \frac{\partial \frac{1}{1+e^{-Z_{1i}}}}{\partial Z_{1i}} \\
&= -(1+e^{-Z_{1i}})^{-2}(-e^{-Z_{1i}}) \\
&= \frac{1}{1+e^{-Z_{1i}}} \cdot \frac{e^{-Z_{1i}}}{1+e^{-Z_{1i}}} \\
&= \frac{1}{1+e^{-Z_{1i}}} \cdot \frac{(1+e^{-Z_{1i}})-1}{1+e^{-Z_{1i}}} \\
&= \frac{1}{1+e^{-Z_{1i}}} \cdot (1 - \frac{1}{1+e^{-Z_{1i}}}) \\
&= A_{1i} \cdot (1 - A_{1i})
\end{aligned} \tag{11}$$

The derivative of Z_{1i} with respect to W_{1i}:

$$\frac{\partial Z_{1i}}{\partial W_{1i}} = X_i \tag{12}$$

Therefore Eq.(7) become:

$$\frac{\partial L}{\partial W_{1i}} = \frac{\partial L}{\partial A_{2i}} \cdot \frac{\partial A_{2i}}{\partial Z_{2i}} \cdot \frac{\partial Z_{2i}}{\partial A_{1i}} \cdot \frac{\partial A_{1i}}{\partial Z_{1i}} \cdot \frac{\partial Z_{1i}}{\partial W_{1i}} = (A_{2i} - Y_i) \cdot W_{2i} \cdot A_{1i} \cdot (1 - A_{1i}) \cdot X_i \tag{7}$$

NOTE: Directly bring these math notions into Python code is not going to work. You must be careful with the shape of your matrices when implementing your backward-pass. One way to see the big picture is writing down the chain rule and the shape of these matrices, you will see. Have fun while learning!

If you detect errors in this document, feel free to email me.