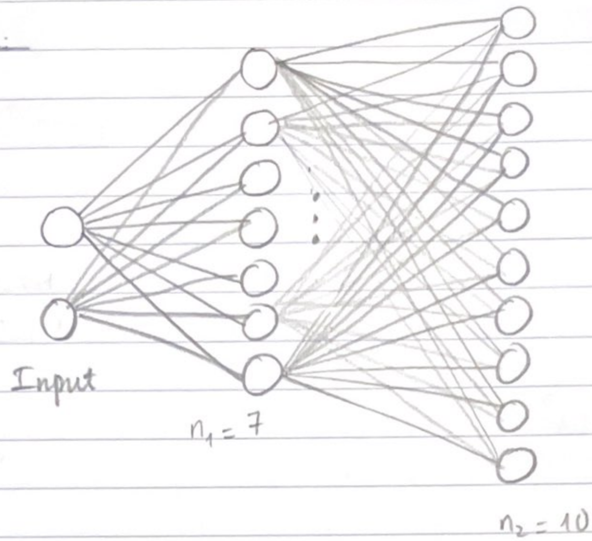
Memo No.__________
Date / /

CS-E4950 COMPUTER VISION
Exercise Round 9

Exercise 1:



(each neuron in Layer l is connected to all neurons in layer $l+1$)

1) We have $m = 1$:

$$E = \frac{1}{m}\sum_{j=1}^{m} -t_j \cdot \log(y_j) = -t.\log(y) = -t.\log(\sigma(Wx))$$
$$= -t.\log\left(\frac{1}{1+e^{-Wx}}\right)$$

Therefore, we can calculate:

$$\frac{\partial E}{\partial t} = -\log\left(\frac{1}{1+e^{-Wx}}\right)$$

$$\frac{\partial E}{\partial x} = -\frac{t}{\sigma(Wx)} \cdot \frac{\partial \sigma(Wx)}{\partial x} = -t(1+e^{-Wx})\left(1-\frac{1}{1+e^{-Wx}}\right)W\frac{1}{1+e^{-Wx}}$$
$$= -t.W.\frac{e^{-Wx}}{1+e^{-Wx}}$$

2) We have:

$$\frac{\partial E}{\partial z_i^{(2)}} = \sum_{j=1}^{n} \frac{\partial E_j}{\partial y_j^{(2)}} \cdot \frac{\partial y_j^{(2)}}{\partial z_i^{(2)}}$$ (From the chain rule)

$$\frac{\partial E_j}{\partial y_j^{(2)}} = \frac{\partial(-t_j \cdot \log y_i)}{\partial z_i^{(2)}} = -\frac{t_j}{y_i}$$

+) If $i = j$: $$\frac{\partial y_i^{(2)}}{\partial z_i^{(2)}} = \frac{\partial(\sigma(z_j^{(2)}))}{\partial z_i^{(2)}} = -y_i y_j$$

+) If $i \neq j$: $$\frac{\partial y_j^{(2)}}{\partial z_i^{(2)}} = \frac{\partial(\sigma(z_i)^{(2)})}{\partial z_i^{(2)}} = y_i(1-y_i)$$

HẢI TIẾN

Memo No.

Date / /

Assuming $\sum_{j=1}^{n} t_j = 1$, taking into account the previous equations:

$$\frac{\partial E}{\partial z_i^{(2)}} = \sum_{j=1}^{n} \frac{\partial E_j}{\partial y_j^{(2)}} \cdot \frac{\partial y_j^{(2)}}{\partial z_i^{(2)}} = \sum_{i \neq j} \frac{\partial E_j}{\partial y_j^{(2)}} \cdot \frac{\partial y_j^{(2)}}{\partial z_i^{(2)}} + \sum_{i=j} \frac{\partial E_j}{\partial y_j^{(2)}} \cdot \frac{\partial y_j^{(2)}}{\partial z_i^{(2)}}$$

$$= \sum_{i \neq j} \left(-\frac{t_j}{y_j} (-y_i y_j) \right) - \frac{t_i}{y_i} y_i (1 - y_i)$$

$$= \sum_{i \neq j} t_j y_i + t_i y_i - t_i$$

$$= y_i - t_i$$

Therefore,

$$\frac{\partial E}{\partial z^{(2)}} = (y^{(2)} - t)^T$$

3) We have:

$$\frac{\partial E}{\partial z^{(2)}} = (y^{(2)} - t)^T \Rightarrow \frac{\partial E}{\partial y^{(1)}} = \frac{\partial E}{\partial z^{(2)}} \cdot \frac{\partial z^{(2)}}{\partial y^{(1)}} = (y^{(2)} - t)^T W^{(2)}$$

4) We have:

$$\frac{\partial E}{\partial w_{uv}^{(1)}} = \frac{\partial E}{\partial z^{(2)}} \cdot \frac{\partial z^{(2)}}{\partial w_{uv}^{(2)}} = \left(\frac{\partial E}{\partial z^{(2)}} \right)_u y_v^{(2)} = (y_u^{(2)} - t_u) y_v^{(1)}$$

Hence, $\frac{\partial E}{\partial w_{uv}^{(2)}} = (y^{(2)} - t) y^{(1)T}$

5) We have:

$$\frac{\partial(\sigma(z))}{\partial z} = \frac{e^{-z}}{(1+e^{-z})^2} = \frac{1}{1+e^{-z}} \cdot \frac{e^{-z}}{(1+e^{-z})} = \sigma(z)(1-\sigma(z))$$

Therefore, $\frac{\partial y^{(1)}}{\partial z^{(1)}} = \frac{\partial(\sigma(z^{(1)}))}{\partial z^{(1)}} = y^{(1)}(1-y^{(1)}) = diag(y^{(1)} .* (1-y^{(1)}))$

6) We have:

$$\frac{\partial E}{\partial z^{(1)}} = \frac{\partial E}{\partial y^{(1)}} \cdot \frac{\partial y^{(1)}}{\partial z^{(1)}} = (y^{(2)} - t)^T W^{(2)} \cdot diag(y^{(1)} .* (1 - y^{(1)}))$$

HẢI TIẾN

Memo No.
Date / /

7) We have:

$$\frac{\partial E}{\partial w_{uv}^{(1)}} = \frac{\partial E}{\partial z^{(1)}} \cdot \frac{\partial z^{(1)}}{\partial w_{uv}^{(1)}} = \frac{\partial E}{\partial z_u^{(1)}} \cdot \frac{\partial z_u^{(1)}}{\partial w_{uv}^{(1)}} = \frac{\partial E}{\partial z_u^{(1)}} \cdot \frac{\partial (\sum_i w_{ui} x_i)}{\partial w_{uv}^{(1)}}$$

$$= \frac{\partial E}{\partial z_u^{(1)}} x_v$$

Hence, $\frac{\partial E}{\partial W^{(1)}} = \left(\frac{\partial E}{\partial z^{(1)}}\right)^T x^T$