week 9 theory question

Hong Pengfei 1002949

1. $$\frac{\partial E}{\partial n_4} = \frac{\partial E}{\partial n_6}\frac{\partial n_6}{\partial n_4}$$

$$= \sum_{i \in Data} 2(n_6 - y_6^{(i)}) \cdot W_{46}$$

2. $$\frac{\partial E}{\partial W_{2,5}} = \frac{\partial E}{\partial n_7} \cdot \frac{\partial n_7}{\partial n_5} \cdot \frac{\partial n_5}{\partial W_{2,5}}$$

$$= \sum_{i \in Data} 2(n_7 - y_6^{(i)}) \cdot W_{5,7} \cdot n_2$$

3. $$\frac{\partial E}{\partial (V_{1,1})_d} = \frac{\partial E}{\partial n_6}\frac{\partial n_6}{\partial n_3}\frac{\partial n_3}{\partial n_1}\frac{\partial n_1}{\partial (V_{1,1})_d} + \frac{\partial E}{\partial n_6}\frac{\partial n_6}{\partial n_4}\frac{\partial n_4}{\partial n_1} \cdot \frac{\partial n_1}{\partial (V_{1,1})_d}$$

$$= \sum_{i \in Data} 2(n_6 - y^{(i)}) \cdot (W_{3,6} \cdot W_{1,3} + W_{4,6} W_{1,4}) \cdot (X_1)_d$$

4. $$\frac{\partial E}{\partial (x_2)_d} = \left[\frac{\partial E}{\partial n_6}\frac{\partial n_6}{\partial n_4}\frac{\partial n_4}{\partial n_2}\frac{\partial n_2}{\partial (x_2)_d} + \frac{\partial E}{\partial n_7}\frac{\partial n_7}{\partial n_5}\frac{\partial n_5}{\partial n_2}\frac{\partial n_2}{\partial (x_2)_d}\right]$$

$$= \sum_{i \in Data} 2(n_6 - y) W_{4,6} W_{2,4} (V_{2,2})_d + 2(n_7 - y) W_{5,7} W_{2,5} (V_{2,2})_d$$

Fig2:

1. $$\frac{\partial E}{\partial (V_{2,2})_d} = \frac{\partial E}{\partial n_6}\frac{\partial n_6}{\partial n_4}\frac{\partial n_4}{\partial n_2}\frac{\partial n_2}{\partial (V_{2,2})_d} + \frac{\partial E}{\partial n_8}\frac{\partial n_8}{\partial n_4}\frac{\partial n_4}{\partial n_2}\frac{\partial n_2}{\partial (V_{2,2})_d}$$

$$= \sum_{i \in Data} 2(n_6 - y^{(i)}) W_{46} W_{24} (X_2)_d + 2(n_8 - y^{(i)}) W_{48} W_{24} (X_2)_d$$

2. $$\frac{\partial E}{\partial W_{2,4}} = \frac{\partial E}{\partial n_6}\frac{\partial n_6}{\partial n_4}\frac{\partial n_4}{\partial W_{2,4}} + \frac{\partial E}{\partial n_8}\frac{\partial n_8}{\partial n_4}\frac{\partial n_4}{\partial W_{2,4}}$$

$$= \sum_{i \in D} 2(n_6 - y^{(i)}) W_{46} n_2 + 2(n_8 - y^{(i)}) W_{48} n_2 .$$

3. $$\frac{\partial E}{\partial n_1} = \frac{\partial E}{\partial n_6}\frac{\partial n_6}{\partial n_3}\frac{\partial n_3}{\partial n_1} + \frac{\partial E}{\partial n_7}\frac{\partial n_7}{\partial n_5}\frac{\partial n_5}{\partial n_1} + \frac{\partial E}{\partial n_8}\frac{\partial n_8}{\partial n_3}\frac{\partial n_3}{\partial n_1}$$

$$= \sum_{i \in Data} 2(n_6 - y^{(i)}) w_{36} w_{1,3} + 2(n_7 - y^i) w_{57} w_{15} + 2(n_8 - y^i) w_{38} w_{13}$$

2.

There will be k terms in the product

Base case: $k=1$. The gradient has 1 term

$$\frac{\partial E}{\partial n_1} = \frac{\partial E}{\partial n_1}$$

There n_{k+1} is always one term more than n_k

because $$\frac{\partial E}{\partial n_{k+1}} = \frac{\partial E}{\partial n_1}\frac{\partial n_1}{\partial n_2} \cdots \frac{\partial n_{k-1}}{\partial n_k}\frac{\partial n_k}{\partial n_{k+1}}$$

There fore by induction $\frac{\partial E}{\partial n_k}$ will have $(k-1)$ layers more than basecase

since base case is 1. total term will be $(k-1)+1 = k$ terms