CSDS 440: Assignment 6

Shaochen (Henry) ZHONG, sxz517

Mingyang TIE, mxt497

Due and submitted on 10/16/2020
Fall 2020, Dr. Ray

Problem 24

For traditional backpropagation of n_k in k-th layer, we have $\frac{\partial L}{\partial n_k} = \sum_{k+1 \in \text{Downstream}(n_k)} \frac{\partial L}{\partial n_{k+1}} \cdot \frac{\partial n_{k+1}}{\partial n_k}$ assuming there is an edge from n_k to n_{k+1}. For the non-feedforward structure the question suggests, n_k will still calculate the above loss, but it will also consider other nodes in the k-th layer which it connects to n_k. We denotes these nodes as $n_{k_i} = \{n_{k_1}, n_{k_2}, n_{k_3}, ...\}$, where edges like $n_k \to n_{k_i}$ exist.

Since there is no cycle, these n_{k_i} nodes will not connect to any Downstream(n_k) nodes. So we can just do backpropagation layer by layer and once the above $\sum_{k+1 \in \text{Downstream}(n_k)} \frac{\partial L}{\partial n_{k+1}} \cdot \frac{\partial n_{k+1}}{\partial n_k}$ is calculated, we will calculate the loss of n_{k_i} nodes with respect to n_k and add to the loss of n_k. Which will give us:

$$
\begin{aligned}
\frac{\partial L}{\partial n_k} &= \sum_{k+1 \in \text{Downstream}(n_k)} \frac{\partial L}{\partial n_{k+1}} \cdot \frac{\partial n_{k+1}}{\partial n_k} + \sum_{n_{k_i}} \frac{\partial L}{\partial n_{k_i}} \cdot \frac{\partial n_{k_i}}{\partial n_k} \\
\implies \frac{\partial L}{\partial w_{(k-1)k}} &= \frac{\partial L}{\partial n_k} \frac{\partial n_j}{w_{(k-1)k}} \\
&= \frac{\partial L}{\partial n_k} \cdot x_{(k-1)k}
\end{aligned}
$$

Assume there is a $n_{k-1} \to n_k$ edge.