CMPUT 367

Roderick Lan

Contents

Lecture 13 - Mar 5

1 Recap

Foward Prop - know weights, compute output (recursive, bottom up)
Backward Prop - know output, update weights via gradients (recursive, top down)

2 Backprop

Initialization

L layers

$$\frac{\partial J}{\partial y^{(L)}}, \quad \frac{\partial J}{\partial z^{(L)}}$$

Depends on loss
The loss, J, should be **generally differentiable**
Derivatives must be informative

Recursion

Suppose $\dfrac{\partial J}{\partial y^{(l)}}, \dfrac{\partial J}{\partial z^{(l)}}$ is known.

Goal: Calculate

$$\frac{\partial J}{\partial y^{(l-1)}}, \quad \frac{\partial J}{\partial z^{(l-1)}}$$

$$\frac{\partial J}{\partial y_1} = \frac{\partial J}{\partial z_1} \cdot \frac{\partial z_1}{\partial y_1} + \frac{\partial J}{\partial z_2} \cdot \frac{\partial z_2}{\partial y_2} + \cdots$$

Termination

Terminate once you get derivative about input
ie. once

$$\frac{\partial J}{\partial y^{(1)}} \frac{\partial J}{\partial z^{(1)}}$$

are calculated