Introduction to Neural Network

Dihui Lai

dlai@wustl.edu

April 30, 2019

Neural Network: Forward

Each neuron at layer l recieves inputs from all neuron from the previous layer $l-1$

$$z_k^l = \sum_j w_{kj}^{l-1} a_j^{l-1}$$

The neuron tansfer the input signal z_k^l via a transfer function σ and send as input to to the next layer

$$a_k^l = \sigma(z_k^l)$$

The cost function of the neural network is dependeng on all the zs of neurons in all layers

$$C\left(z_1^l, z_2^l, \ldots z_k^l(z_1^{l-1}, z_2^{l-1}, z_3^{l-1}, \ldots), \ldots\right)$$

Neural Network: Backpropogation

The contribution to the cost function from a neuron in layer l can be cauculated iteratively as

$$
\begin{aligned}
\delta_k^l = \frac{\partial C}{\partial z_k^l} &= \sum_m \frac{\partial C}{\partial z_m^{l+1}} \frac{\partial z_m^{l+1}}{\partial z_k^l} \\
&= \left(\sum_m \frac{\partial C}{\partial z_m^{l+1}} \frac{\partial z_m^{l+1}}{\partial a_k^l} \right) \frac{\partial a_k^l}{\partial z_k^l} \\
&= \sum_m \delta_m^{l+1} w_{mk}^l \sigma'(z_k^l)
\end{aligned}
$$

The partial derivative of a cost function w.r.t the weight w_{kj}^{l-1} is

$$
\frac{\partial C}{\partial w_{kj}^{l-1}} = \frac{\partial C}{\partial z_k^l} \frac{\partial z_k^l}{\partial w_{kj}^{l-1}} = \delta_k^l a_j^{l-1}
$$