Exercise 9 - Martin Törnqvist

DD 2427

2015-05-09

Part 2

$$L(\bar{y}, \bar{z}^{(2)}) = \| \bar{y} - \bar{z}^{(2)} \|^2$$

$$\frac{\partial L(\bar{y}, \bar{z}^{(3)})}{\partial w_{k0}^{(1)}} = \sum_{s=1}^{K} \underbrace{\frac{\partial L(\bar{y}, \bar{z}^{(2)})}{\partial z_s^{(3)}} \cdot \frac{\partial z^{(3)}}{\partial a_s^{(3)}}}_{-(y_s - z_s^{(3)}) \cdot g'(a_s^{(3)})} \cdot \frac{\partial a_s^{(3)}}{\partial w_{k0}^{(1)}}$$

$g(x)$ activation function for output layer

$$\frac{\partial a_s^{(3)}}{\partial w_{k0}^{(1)}} = \sum_{r=1}^{M} \frac{\partial a_s^{(3)}}{\partial z_r^{(2)}} \cdot \frac{\partial z_r^{(2)}}{\partial a_r^{(2)}} \cdot \frac{\partial a_r^{(2)}}{\partial w_{k0}^{(1)}}$$

$$\frac{\partial a_s^{(3)}}{\partial z_r^{(2)}} = \frac{\partial \left(\sum_{i=1}^{M} w_{si}^{(2)} z_i^{(2)} + w_{s0}^{(2)} \right)}{\partial z_r^{(2)}} = w_{sr}^{(2)}$$

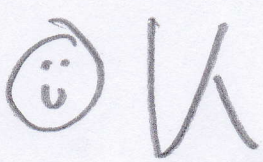
$$\frac{\partial z_r^{(2)}}{\partial a_r^{(2)}} = \frac{\partial f(a_r^{(2)})}{\partial a_r^{(2)}} = f'(a_r^{(2)})$$

$f(x)$ activation function for hidden layer

$$\frac{\partial a_r^{(2)}}{\partial w_{k0}^{(1)}} = \frac{\partial \left(\sum_{j=1}^{N} w_{rj}^{(1)} z_j^{(1)} + w_{r0}^{(1)} \right)}{\partial w_{k0}^{(1)}} = \delta_{rk} \Rightarrow$$

$$\Rightarrow \frac{\partial L(\bar{y}, z^{(3)})}{\partial w_{k0}^{(1)}} = \sum_{s=1}^{K} \underbrace{-(y_s - z_s^{(3)}) \cdot g'(a_s^{(3)})}_{\delta_s^{(3)}} \cdot \underbrace{\sum_{r=1}^{M} w_{sr}^{(2)} f'(a_r^{(2)}) \cdot \delta_{rk}}_{w_{sk}^{(2)} \cdot f'(a_k^{(2)})} =$$

$$= f'(a_k^{(2)}) \cdot \sum_{s=1}^{K} \delta_s^{(3)} w_{sk}^{(2)} = \delta_k^{(2)}$$



Note: K = no of output nodes

M = no of hidden nodes

N = no of inputs