# The MLP example

Nov. 5, 2019

Error function

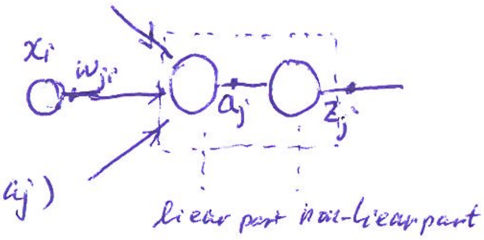
$$E_n = \frac{1}{2} \sum_{k=0}^{K} (y_k - t_k)^2$$

for the output layer, no non-linear function $y_k = a_k$

$$\frac{\partial E_n}{\partial a_k} = \frac{\partial E_n}{\partial y_k} = \frac{1}{2} \cdot 2 \cdot (y_k - t_k) = y_k - t_k$$

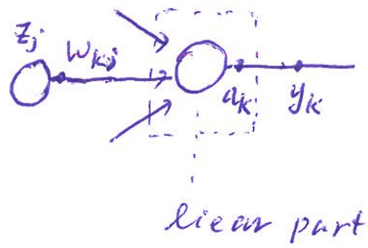
Step 1: Forward Pass

hidden layer $\quad a_j = \sum_{i=0}^{P} w_{ji}^{(1)} x_i$




$$z_j = \tanh(a_j) \equiv h(a_j)$$

output layer $\quad a_k = y_k = \sum_{j=0}^{M} w_{kj}^{(2)} z_j$




step 2: Gradient at the output layer

$$\frac{\partial E_n}{\partial w_{kj}^{(2)}} = \frac{\partial E_n}{\partial a_k} \cdot \frac{\partial a_k}{\partial w_{kj}^{(2)}}$$

$$= (y_k - t_k) \cdot \frac{\partial \left( \sum_{j=0}^{M} w_{kj}^{(2)} \cdot z_j \right)}{\partial w_{kj}}$$

$$= (y_k - t_k) \cdot z_j$$

let $\quad \delta_k \equiv \frac{\partial E_n}{\partial a_k} = y_k - t_k$

$$\delta_j \equiv \frac{\partial E_n}{\partial a_j}$$

Step3. Error back propagation, calculate Gradient of the hidden Layer

$$\frac{\partial E_n}{\partial w_{ji}^{(1)}} = \frac{\partial E_n}{\partial a_j} \cdot \frac{\partial a_j}{\partial w_{ji}^{(1)}} = \frac{\partial E_n}{\partial a_j} \cdot x_i = \delta_j x_i$$

$$\delta_j = \sum_{k=1}^{K} \left( \frac{\partial E_n}{\partial a_k} \cdot \frac{\partial a_k}{\partial a_j} \right) \quad \text{(error back propation)}$$

$$= \sum_{k=1}^{K} \left( \delta_k \cdot \frac{\partial a_k}{\partial a_j} \right)$$

$$\frac{\partial a_k}{\partial a_j} = \frac{\partial \left( \sum_{j=0}^{M} w_{kj}^{(2)} \cdot h(a_j) \right)}{\partial a_j} = w_{kj}^{(2)} \cdot \frac{\partial h(a_j)}{\partial a_j}$$

$$= w_{kj}^{(2)} \cdot (1 - h^2(a_j)) \qquad h(a) \equiv \tanh(a) = \frac{e^a - e^{-a}}{e^a + e^{-a}}$$

$$= w_{kj}^{(2)} \cdot (1 - z_j^2) \qquad h'(a) = 1 - h^2(a)$$

$$\delta_j = (1 - z_j^2) \sum_{k=0}^{K} w_{kj}^{(2)} \cdot \delta_k$$

$$\frac{\partial E_n}{\partial w_{ji}^{(1)}} = x_i (1 - z_j^2) \sum_{k=1}^{K} w_{kj}^{(2)} \delta_k$$

$$= x_i (1 - z_j^2) \sum_{k=1}^{K} w_{kj}^{(2)} \cdot (y_k - t_k)$$