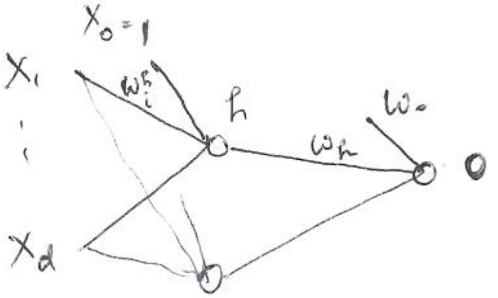
1- Hidden layer, 1- output



$$o = \sigma\left(w_0 + \sum_h w_h o_h\right) \equiv \sigma\left(\sum_h w_h o_h\right)$$

$$o_h = \sigma\left(w_0^h + \sum_i w_i^h x_i\right) \equiv \sigma\left(\sum_i w_i^h x_i\right)$$

$$E[W] = \frac{1}{2} \sum_{l \in D} (y^l - o^l)^2$$

set of all wts. from all layers

$$\frac{\partial E}{\partial w'} = \sum_{l \in D} -(y^l - o^l)\frac{\partial o^l}{\partial w'}$$

$w' = w_h$ $\quad \frac{\partial o^l}{\partial w_h} = \cancel{\frac{\partial \sigma(\sum_h w_h o_h^l)}{\partial w_h}} = \frac{\partial o^l}{\partial net^l} \cdot \frac{\partial net^l}{\partial w_h}$ where $net^l = \sum_h w_h o_h^l$

$$\frac{\partial o^l}{\partial net^l} = \frac{\partial \sigma(net^l)}{\partial net^l} = o^l(1-o^l)$$

$$\frac{\partial net^l}{\partial w_h} = o_h^l$$

$w' = w_i^h$ $\quad \frac{\partial o^l}{\partial w_i^h} = \frac{\partial o^l}{\partial net^l} \cdot \frac{\partial net^l}{\partial w_i^h}$

same

$$\frac{\partial net^l}{\partial w_i^h} = \frac{\partial}{\partial w_i^h}\left(\sum_{h'} w_{h'} \overbrace{\sigma\left(\underbrace{\sum_i w_i^{h'} x_i^l}_{net_{h'}^l}\right)}^{o_{h'}^l}\right)$$

$$= w_h \frac{\partial o_h^l}{\partial net_h^l} \cdot \frac{\partial net_h^l}{\partial w_i^h}$$

$$\frac{\partial o_h^l}{\partial net_h^l} = o_h^l(1-o_h^l)$$

$$\frac{\partial net_h^l}{\partial w_i^h} = x_i^l$$

Sum up:

$w' = w_h \qquad \dfrac{\partial E}{\partial w'} = \sum_{l \in D} \overbrace{-(y^l - o^l)\, o^l(1-o^l)}^{\delta^l}\, o_h^l$

$w' = w_i^h \qquad \dfrac{\partial E}{\partial w'} = \sum_{l \in D} \underbrace{\underbrace{-(y^l-o^l)\, o^l(1-o^l)}_{\delta^l}\, o_h^l(1-o_h^l)\, w_h}_{\delta_h^l = o_h^l(1-o_h^l) w_h \delta^l}\, x_i^l$

$w_{ij} \leftarrow w_{ij} + \Delta w_{ij}^l \qquad \Delta w_{ij}^l = -\eta \nabla E_l = \eta\, \delta_j^l\, o_i^l$

eg if $w_{ij} = w_i^h \qquad x_i \longrightarrow h$

$\Delta w_i^h = \eta\, \delta_h^l\, x_i^l \qquad o_i^l = x_i^l$

if $w_{ij} = w_h \qquad h \longrightarrow o$

$\Delta w_h = \eta\, \delta^l\, o_h^l \qquad o_i^l = o_h^l$