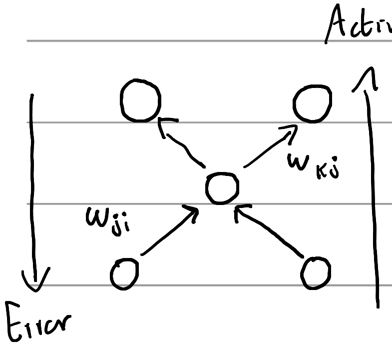
Neural Networks



Output k W weight matrix

Hidden j w_{kj} hidden → output

Input i w_{ji} input → hidden

Notation

subscript k output layer element

subscript j hidden layer element

subscript i input layer element

a - activation value of element within layer

t - target value of element within layer

net - net input of element within layer

Calculus Review:

$$\frac{d(e^u)}{dx} = e^u \frac{du}{dx} \qquad \frac{d(g+h)}{dx} = \frac{dg}{dx} + \frac{dh}{dx}$$

$$\frac{d(g^n)}{dx} = n g^{n-1} \frac{dg}{dx}$$

Total Network Error:

Sum squared error (magnitude of error without a sign)

$$E_T = \frac{1}{2} \sum_k (t_k - a_k)^2$$

only applies to activations / target values in the output layer

Goal of training is to adjust the networks weights to improve upon the error.

$$\Delta W \propto -\frac{\partial E}{\partial w}$$ ← partial

Back prop starts at the output layer

$$\Delta W_{Kj} \propto -\frac{\partial E}{\partial w_{Kj}}$$

we are interested in how the error changes with the output layer activations, how the output activations change with the net input to the output layer, & how the net inputs to the output layer change with weights connecting the hidden layer to the output layer.

$$\Delta W_{Kj} \propto \frac{\partial E}{\partial a_K} \frac{\partial a_K}{\partial net_K} \frac{\partial net_K}{\partial w_{Kj}}$$

$$= \varepsilon \frac{\partial E}{\partial a_K} \frac{\partial a_K}{\partial net_K} \frac{\partial net_K}{\partial w_{Kj}}$$

$$\frac{\partial E}{\partial a_K} = \frac{\partial}{\partial a_K}\left(\frac{1}{2}\sum_K (t_K - a_K)^2 \right)$$

can ignore summation

$$= \frac{1}{2}\frac{\partial}{\partial a_K}(t_K - a_K)^2 \quad *$$

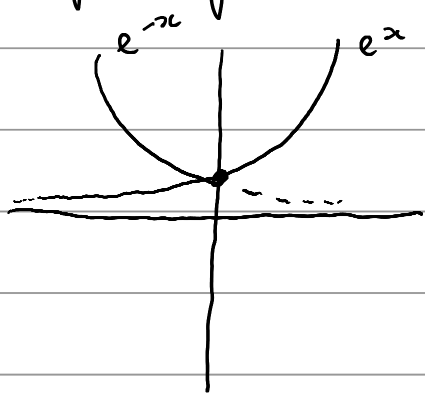
Think about it $\frac{\partial t_K}{\partial a_K} = 0$ as the target does not care about the activation value.

$$* = \frac{2}{2}(t_K - a_K) = t_K - a_K$$

Activation Function $\frac{1}{1+e^{-net_k}}$

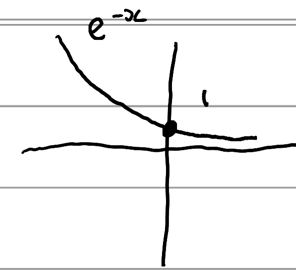
e^x	x
0	$-\infty$
1	0
∞	∞

e^{-x} reflect on y-axis



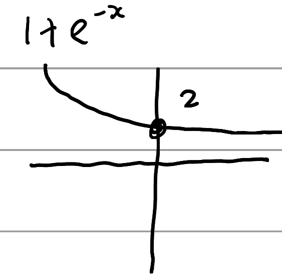
e^{-x}	x
$-\infty$	$-\infty$
1	0
0	∞

e^{-x}	x
$-\infty$	$-\infty$
1	0
0	∞



$1+e^{-x}$ (Shift up by 1)

e^{-x}	x
$-\infty$	$-\infty$
2	0
1	∞

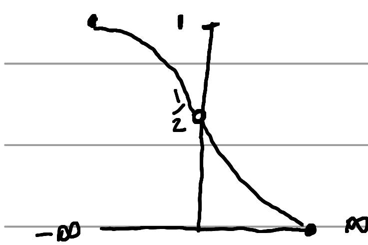


$f(x)\ \frac{1}{1+e^{-x}}$

$f(x)$	x
1	$-\infty$
0.5	0
0	∞

$f(\infty) = \frac{1}{1+\underbrace{e^{-\infty}}_{\approx 0}} = \frac{1}{1}$

$f(-\infty) = \frac{1}{1+\underbrace{e^{\infty}}_{\infty}} = \frac{1}{\infty} = 0$



$\frac{1}{1+e^{-x}}$ maps $x \in \{-\infty \ldots \infty\}$

$f(x) \in \{1 \ldots 0\}$

map unbounded range to bounded

$$a_k = \frac{1}{1+e^{-net_k}}$$

$$\frac{\partial a_k}{\partial net_k} = \frac{\partial}{\partial net_k} \frac{1}{1+e^{-net_k}} = \frac{\partial}{\partial net_k} (1+e^{-net_k})^{-1}$$

$$\boxed{\frac{d(g^n)}{dx} = n g^{n-1} \frac{dg}{dx}}$$

$$g^n = (1+e^{-net_k})$$

$$\frac{\partial a_k}{\partial net_k} = (-1)(1+e^{-net_k})^{-2} \frac{\partial}{\partial net_k} (1+e^{-net_k})$$

$$= -(1+e^{-net_k})^{-2} (-1) e^{-net_k}$$

$$= \frac{e^{-net_k}}{(1+e^{-net_k})^2} = \underbrace{\frac{1}{(1+e^{-net_k})}}_{a_k} \frac{e^{-net_k}}{(1+e^{-net_k})}$$

$$= a_k \frac{e^{-net_k}}{1+e^{-net_k}} \quad *$$

$$\frac{e^{-net_k}}{1+e^{-net_k}} = \frac{A}{1+e^{-net_k}} + \frac{B}{1+e^{-net_k}}$$

$$= \frac{A(1+e^{-net_k}) + B(1+e^{-net_k})}{1+e^{-net_k}}$$

$$A = 1$$

$$B = -\frac{1}{1+e^{-net_k}}$$

$$\frac{e^{-net_k}}{1+e^{-net_k}} = \frac{1+e^{-net_k}}{1+e^{-net_k}} - \frac{1}{1+e^{-net_k}}$$

$$= 1 - \underbrace{\frac{1}{1+e^{-net_k}}}_{a_k}$$

$$* \Rightarrow a_k (1 - a_k)$$

$$\frac{\partial}{\partial w_{kj}} net_k = \frac{\partial (w_{kj} a_j)}{\partial w_{kj}} = a_j$$

so what is Δw_{kj}

$$\Delta w_{kj} = \varepsilon (t_k - a_k) a_k (1 - a_k) a_j$$

$$\Delta w_{kj} = \varepsilon \frac{\partial E}{\partial a_k} \frac{\partial a_k}{\partial net_k} \frac{\partial net_k}{\partial w_{kj}}$$

$$\Delta w_{ji} \, \alpha \sum_k \left[\frac{\partial E}{\partial a_k} \frac{\partial a_k}{\partial net_k} \frac{\partial net_k}{\partial a_j} \right] \frac{\partial a_j}{\partial net_j} \frac{\partial net_j}{\partial w_{ji}}$$

$$= \varepsilon \left[\sum_k (t_k - a_k) a_k (1 - a_k) \right] a_j (1 - a_j) a_i$$

$$e^{-net_k} = A(1+e^{-net_k}) + B(1+e^{-net_k})$$

$$e^{-net_k} = A + Ae^{-net_k} + B + Be^{-net_k}$$

$$= (A+B) + e^{-net_k} \cdot (A+B)$$