NN week 3

Lisa Tostrams s4386167

2 Maart 2015

1 Math exercises

1.1

$\frac{\partial f}{\partial a}$ with $f(a) = \frac{1}{1+exp(-a)}$. $f(a)$ is composed of two functions:

$f(h) = \frac{1}{h}$ with derivative $\frac{\partial f}{\partial h} = -\frac{1}{h^2}$ and

$h(a) = 1 + exp(-a)$ with derivative $\frac{\partial h}{\partial a} = exp(-a)$

With the chain rule $\frac{\partial f}{\partial a} = \frac{\partial f}{\partial h} * \frac{\partial h}{\partial a}$ we get

$\frac{\partial f}{\partial a} = -\frac{1}{(1+exp(-a))^2} * exp(-a) = \frac{exp(-a)}{(1+exp(-a))^2}$

The given function $f(a)(1 - f(a)$ can be rewritten as:

$\frac{1}{1+exp(-a)} * \left(\frac{1+exp(-a)}{1+exp(-a)} - \frac{1}{1+exp(-a)}\right)$

and then

$\frac{1}{1+exp(-a)} * \left(\frac{exp(-a)}{1+exp(-a)}\right) = \frac{exp(-a)}{(1+exp(-a))^2}$

Which was the function that was derived from $\frac{\partial f}{\partial a}$.

1.2

We have the function

$\delta_j = \frac{\partial E^n}{\partial y_j} * \frac{\partial y_j}{\partial a_j}$

With $f(a_j) = y_j$ and $E^n(w) = \frac{1}{2} * (y_j - t_j)^2$ that is equivalent to

$\delta_j = \frac{\partial \frac{1}{2} * (y_j - t_j)^2}{\partial y_j} * \frac{\partial f(a_j)}{\partial a_j}$

The first part derives to

$$\frac{\partial \frac{1}{2} * (y_j - t_j)^2}{\partial y_j} = (y_j - t_j)$$

With the proof from exercise 1.1, the second part derives to

$$\frac{\partial f(a_j)}{\partial a_j} = f(a_j) * (1 - f(a_j))$$

Which, when we use $f(a_j) = y_j$, reduces to

$$f(a_j) * (1 - f(a_j)) = y_j * (1 - y_j)$$

So, finally:

$$\delta_j = \frac{\partial \frac{1}{2} * (y_j - t_j)^2}{\partial y_j} * \frac{\partial f(a_j)}{\partial a_j} = (y_j - t_j) * y_j * (1 - y_j)$$

1.3

With a linear activation function $f(a) = a$, the output for a datapoint

$$y^n = f(W^2 * f(W^1 * x^n))$$

can be rewritten as

$$y^n = W^2 * W^1 * x^n$$

Because W^2 and W^1 are both matrices, their multiplication evaluates to another matrix W.

$$y^n = W * x_n$$

Which is the same output function as the output function for a single layer network.

1.4

Same as 1.2.

1.5

2 Matlab exercises