Equations AET NN algorithm

Feedforward

$$z^l = w^l a^{l-1} + b^l$$

$$a^l = \sigma(z^l)$$

σ hidden layer: sigmoid
σ output layer: softmax

Quadratic cost function

$$C(w,b) = \frac{1}{2n} \sum_x ||y(x) - a||^2$$

Output error

$$\delta^L = \nabla_a C \circ \sigma'(z^l)$$

Backpropagate error

$$\delta^l = ((w^{l+1})^T \delta^{l+1} \circ \sigma'(z^l)$$

Gradient of cost function

$$\frac{\delta C}{\delta w^l_{jk}} = a^{l-1}_k \delta^l_j$$

$$\frac{\delta C}{\delta b^l_j} = \delta^l_j$$

Source: Neural Networks and Deep Learning Chapter 1