Forward-Feed, Multi-layer Artificial Neural Network — Part II

Layer 0 Layer 1 Layer 2

$(i = 0, 1)$ $(j = 0, 1)$

X

x_0 $\left(a_0^{(0)}\right)$

x_1 $\left(a_1^{(0)}\right)$

$w_{00}^{(0)}$ $w_{01}^{(0)}$ $w_{10}^{(0)}$ $w_{11}^{(0)}$

Σ g

$a_0^{(1)}$ $a_1^{(1)}$

$w_0^{(1)}$ $w_1^{(1)}$

z $\left(a_0^{(2)}\right)$

Note: eqn's (1) and (2) have the same structure:

$$(\text{delta} \otimes a) \cdot \alpha$$

Also note how the δ's are related.

Eqn (2) is more representative: generally Δw is a $m \times n$ matrix, the outer product of two vectors, δ (n-dim) and the input, a (m-dim).

$$\text{delta0}_j = \delta_j^{(0)} = \delta^{(1)} w_j^{(1)} \cdot g'(a_j^{(1)})$$

$$\Delta w_{ij}^{(0)} = \delta_j^{(0)} a_i^{(0)} \alpha \qquad (2)$$

output for layer 0

input for layer 0

$$\text{delta1} = \delta^{(1)} = (y - z) \cdot z(1 - z) = (y - z) \cdot g'(a_0^{(2)})$$

$$\Delta w_j^{(1)} = \delta^{(1)} \; a_j^{(1)} \alpha \qquad (1')$$

output for layer 1

input for layer 1