- The "generalized" Delta rule: $\Delta w_{ij} = \varepsilon\, \delta_i\, a_j$
  - For output nodes: $\delta_i = \sigma_i' \cdot (t_i - a_i)$
  - For hidden nodes: $\delta_i = \sigma_i' \cdot \sum_k \delta_k\, w_{ki}$
- The first derivative of the logistic function: $\sigma_i' = a_i\,(1 - a_i)$

Forward pass: For each node except the input nodes, calculate *net* (netinput) and then *a* (activation), until you reach the output nodes.

$$\begin{aligned} net_1 &= \textstyle\sum_j w_{ij} a_j \\ &= w_{1,0}\, a_0 + w_{1,i1}\, a_{i1} + w_{1,i2}\, a_{i2} \\ &= 0.75 \times 1 + \ldots \times 0 + \ldots \times 0 \\ &= 0.75 \end{aligned}$$

$a_1 = 0.7$ *(Lookup in table)*

$net_2 = -0.60$
$a_2 = 0.35$

$$\begin{aligned} net_3 &= -0.65 - 0.25 \times \overset{a_1}{0.7} + 0.45 \times \overset{a_2}{0.35} \\ &= -0.65 - 0.175 + 0.16 \\ &= -0.665 \end{aligned}$$

$a_3 = 0.35$

Backward pass: First compute all δ's layer by layer, then make weight changes (note that weight changes occur all at once at the end of the sweep).

$$\begin{aligned} \sigma_3' &= a_3\ (1 - a_3) \\ &= 0.35\ (1 - 0.35) \\ &= 0.2275 \end{aligned}$$

$$\begin{aligned} \delta_3 &= \sigma_3' \cdot (t_3 - a_3) \\ &= 0.2275 \times (0 - 0.35) \\ &= -0.08 \end{aligned}$$

$$\begin{aligned} \Delta w_{3,0} &= \varepsilon \cdot \delta_3 \cdot a_0 \\ &= 10 \times -0.08 \times 1 \\ &= -0.8 \end{aligned}$$

$$\begin{aligned} \Delta w_{3,1} &= 10 \times -0.08 \times 0.7 \\ &= -0.56 \end{aligned}$$

$$\begin{aligned} \Delta w_{3,2} &= 10 \times -0.08 \times 0.35 \\ &= -0.28 \end{aligned}$$

$$\begin{aligned} \sigma_1' &= 0.7\ (1 - 0.7) \\ &= 0.21 \end{aligned}$$

$$\begin{aligned} \delta_1 &= \sigma_1' \cdot \delta_3 \cdot w_{3,1} \\ &= 0.21 \times -0.08 \times -0.25 \\ &= 0.0042 \end{aligned}$$

$$\begin{aligned} \Delta w_{1,0} &= 10 \times 0.0042 \times 1 \\ &= 0.042 \end{aligned}$$

$\Delta w_{1,i1} = 10 \times 0.0042 \times 0 = 0$
$\Delta w_{1,i2} = 0$

$\sigma_2' = 0.2275\ (= \sigma_3')$

$$\begin{aligned} \delta_2 &= \sigma_2' \cdot \delta_3 \cdot w_{3,2} \\ &= 0.2275 \times -0.08 \times 0.45 \\ &= -0.008 \end{aligned}$$

$\Delta w_{2,0} = 10 \times -0.008 \times 1 = -0.08$
$\Delta w_{2,i1} = 0$
$\Delta w_{2,i2} = 0$