0.1 BP Calculus

Forward Pass:

$$net = \sum_i w_i x_i$$

$$y = \sigma(net) = \frac{1}{1 + e^{-net}}$$

Loss Function:

$$L = \frac{1}{2}(t - y)^2$$

Compute Gradient of Loss w.r.t. Weight:

$$\frac{\partial L}{\partial w_i} = \frac{\partial L}{\partial y} \cdot \frac{\partial y}{\partial net} \cdot \frac{\partial net}{\partial w_i}$$

where:

$$\frac{\partial L}{\partial y} = -(t - y)$$

$$\frac{\partial y}{\partial net} = y(1 - y)$$

$$\frac{\partial net}{\partial w_i} = x_i$$

Therefore:

$$\frac{\partial L}{\partial w_i} = -(t - y)\, y(1 - y)\, x_i$$

Gradient Descent Update Rule:

$$w_i^{\text{new}} = w_i^{\text{old}} - \eta \frac{\partial L}{\partial w_i}$$

Substitute the gradient:

$$w_i^{\text{new}} = w_i^{\text{old}} + \eta (t - y)\, y(1 - y)\, x_i$$

Define:

$$\delta = (t - y)\, y(1 - y)$$

Then the update becomes:

$$w_i^{\text{new}} = w_i^{\text{old}} + \eta\, \delta\, x_i$$

$$
\begin{aligned}
y &= \sigma\left(\sum_i w_i x_i\right) \\
L &= \tfrac{1}{2}(t-y)^2 \\
\frac{\partial L}{\partial w_i} &= -(t-y)y(1-y)x_i \\
w_i^{\text{new}} &= w_i^{\text{old}} - \eta \frac{\partial L}{\partial w_i} \\
&= w_i^{\text{old}} + \eta(t-y)y(1-y)x_i
\end{aligned}
$$