Network Architecture:

$x_1 \rightarrow In_1 \rightarrow P_3 \rightarrow P_5 \rightarrow a_5$

$x_2 \rightarrow In_2 \rightarrow P_4 \rightarrow P_6 \rightarrow a_6$

0 **1** **2**

Example Matrix Translations:

$\vec{x} = [x_1 \quad x_2]$

$a^o = [a_1 \quad a_2]$

$\vec{y} = [y_5 \quad y_6]$ (expected out)

$w^1 = \begin{bmatrix} w_{1,3} & w_{1,4} \\ w_{2,3} & w_{2,4} \end{bmatrix}$

$b^1 = [b_3 \quad b_4]$

Sample Network:

$\vec{x} = [2 \quad 3]$

$\vec{y} = [.8 \quad 1]$

$w^1 = \begin{bmatrix} -1 & -.5 \\ 1 & .5 \end{bmatrix}$ $\quad b^1 = [1 \quad -1]$

$w^2 = \begin{bmatrix} 1 & 2 \\ -1 & -2 \end{bmatrix}$ $\quad b^2 = [-.5 \quad .5]$ $\quad A(x) = \dfrac{1}{1+e^{-x}}$

	Layer 0:	Layer 1:	Layer 2:	Calculating Error:
Individual Variable Notation:	In_1: $a_1 = x_1$	P_3: $dot_3 = w_{1,3} \cdot a_1 + w_{2,3} \cdot a_2 + b_3$ $a_3 = A(dot_3)$	P_5: $dot_5 = w_{3,5} \cdot a_3 + w_{4,5} \cdot a_4 + b_5$ $a_5 = A(dot_5)$	$E = \frac{1}{2}\sum (y_i - a_i)^2 =$ $\frac{1}{2}((y_5 - a_5)^2 + (y_6 - a_6)^2)$
	In_2: $a_2 = x_2$	P_4: $dot_4 = w_{1,4} \cdot a_1 + w_{2,4} \cdot a_2 + b_4$ $a_4 = A(dot_4)$	P_6: $dot_6 = w_{3,6} \cdot a_3 + w_{4,6} \cdot a_4 + b_6$ $a_6 = A(dot_6)$	Superscript! Squared!
Matrix Notation:	$a^0 = \vec{x}$	$dot^1 = a^0 \cdot w^1 + b^1$ $a^1 = A(dot^1)$	$dot^2 = a^1 \cdot w^2 + b^2$ $a^2 = A(dot^2)$	$E = \frac{1}{2} mag(\vec{y} - a^2)^2$

Partial Derivatives:

$$-\begin{bmatrix} \frac{\partial E}{\partial w_{1,3}} & \frac{\partial E}{\partial w_{1,4}} \\ \frac{\partial E}{\partial w_{2,3}} & \frac{\partial E}{\partial w_{2,4}} \end{bmatrix} = \begin{bmatrix} \Delta_3 a_1 & \Delta_4 a_1 \\ \Delta_3 a_2 & \Delta_4 a_2 \end{bmatrix}$$

$$-\begin{bmatrix} \frac{\partial E}{\partial b_3} & \frac{\partial E}{\partial b_4} \end{bmatrix} = [\Delta_3 \quad \Delta_4]$$

$$-\begin{bmatrix} \frac{\partial E}{\partial w_{3,5}} & \frac{\partial E}{\partial w_{3,6}} \\ \frac{\partial E}{\partial w_{4,5}} & \frac{\partial E}{\partial w_{4,6}} \end{bmatrix} = \begin{bmatrix} \Delta_5 a_3 & \Delta_6 a_3 \\ \Delta_5 a_4 & \Delta_6 a_4 \end{bmatrix}$$

$$-\begin{bmatrix} \frac{\partial E}{\partial b_5} & \frac{\partial E}{\partial b_6} \end{bmatrix} = [\Delta_5 \quad \Delta_6]$$

Delta Definitions:

$\Delta_3 = A'(dot_3) \cdot (\Delta_5 w_{3,5} + \Delta_6 w_{3,6})$
$\Delta_4 = A'(dot_4) \cdot (\Delta_5 w_{4,5} + \Delta_6 w_{4,6})$
$\Delta_5 = (y_5 - a_5) \cdot A'(dot_5)$
$\Delta_6 = (y_6 - a_6) \cdot A'(dot_6)$

And in matrices:

$\Delta^1 = [\Delta_3 \quad \Delta_4]$
$\Delta^2 = [\Delta_5 \quad \Delta_6]$

Some PD Calculations:

$$-\frac{\partial E}{\partial b_5} = (y_5 - a_5) \cdot A'(dot_5) \cdot 1 = \Delta_5 \cdot 1$$

$$-\frac{\partial E}{\partial w_{3,5}} = (y_5 - a_5) \cdot A'(dot_5) \cdot a_3 = \Delta_5 \cdot a_3$$

$$-\frac{\partial E}{\partial b_3} = \Delta_5 w_{3,5} \cdot A'(dot_3) \cdot 1 + \Delta_6 w_{3,6} \cdot A'(dot_3) \cdot 1$$

$$= A'(dot_3) \cdot (\Delta_5 w_{3,5} + \Delta_6 w_{3,6}) \cdot 1$$
$$= \Delta_3 \cdot 1$$

$$-\frac{\partial E}{\partial w_{1,3}} = \Delta_3 a_1$$

Some Matrix Calcs:

Matrix of PDs of w^1:

$\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \cdot [\Delta_3 \quad \Delta_4]$

$= (a^0)^T \cdot \Delta^1$

Calculating Δ^1:

This:

$[\Delta_5 \quad \Delta_6] \cdot \begin{bmatrix} w_{3,5} & w_{4,5} \\ w_{3,6} & w_{4,6} \end{bmatrix}$

gives the parenthetical sum, then pairwise by the activation derivative:

$= A'(dot^1) \otimes \Delta^2 \cdot (w^2)^T$

Final Algorithm, As Matrices:

Calculate deltas back from N:

$\Delta^N = A'(dot^N) \otimes (\vec{y} - a^N)$
$\Delta^L = A'(dot^L) \otimes \Delta^{L+1} \cdot (w^{L+1})^T$

Update weights and biases:

$b^L = b^L + \lambda \cdot \Delta^L$
$w^L = w^L + \lambda \cdot (a^{L-1})^T \cdot \Delta^L$

Fun fact:

$$
\begin{aligned}
A'(x) &= \frac{e^{-x}}{(1+e^{-x})^2} \\
&= \frac{1}{1+e^{-x}} \cdot \frac{e^{-x}}{1+e^{-x}} \\
&= \frac{1}{1+e^{-x}} \cdot \frac{1+e^{-x}-1}{1+e^{-x}} \\
&= \frac{1}{1+e^{-x}} \cdot \left(\frac{1+e^{-x}}{1+e^{-x}} - \frac{1}{1+e^{-x}}\right) \\
&= \frac{1}{1+e^{-x}} \cdot \left(1 - \frac{1}{1+e^{-x}}\right) \\
&= A(x) \cdot (1 - A(x))
\end{aligned}
$$