1 Training MLP

1.1 Network Topology

Figure 1 contains a network topology consisting of 2 inputs x_1 and x_2 with a weight vector $\{w_1 ... w_2\}$:

w_1	w_2	w_3	w_4	w_5	w_6	w_7	w_8	w_9
-0.1	-0.3	-0.2	0.3	0.1	-0.1	0.2	-0.1	0.2

Table 1: Training Weights

The *Incremental Backpropagation Algorithm* is to be performed with $\eta = 0.2$

x_1	x_2	y_{out}
0.0	1.0	1.0
1.0	0.0	1.0

Table 2: Training Vectors

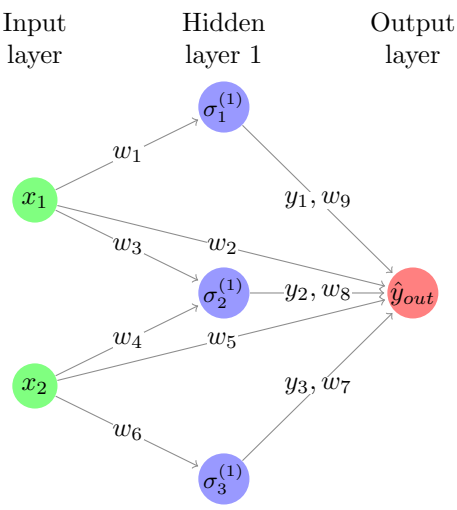

Figure 1: Multilayer Perceptron Topology Example

1.2 Incremental Backpropagation Sigmoid

Since the algorithm is incremental the training vector is not processed as a batch each example contains a forward pass, backward pass to compute errors and compute weight updates.

1.3 First Training Example

1.3.1 Forward Pass

$$
\begin{aligned}
y_1 &= \sigma(w_1 \cdot x_1) \\
&= \sigma(-0.1 \cdot 0.0) \\
&= \sigma(0.0) \\
&= 0.5
\end{aligned}
$$

$$
\begin{aligned}
y_2 &= \sigma(w_3 \cdot x_1 + w_4 \cdot x_2) \\
&= \sigma(-0.2 \cdot 0.0 + 0.3 \cdot 1.0) \\
&= \sigma(0.3) \\
&\approx 0.5744
\end{aligned}
$$

$$
\begin{aligned}
y_3 &= \sigma(w_6 \cdot x_2) \\
&= \sigma(-0.1 \cdot 1.0) \\
&= \sigma(-0.1) \\
&\approx 0.4750
\end{aligned}
$$

$$
\begin{aligned}
\hat{y}_{out} &= w_9 \cdot y_1 + w_2 \cdot x_1 + w_8 \cdot y_2 + w_5 \cdot x_2 + w_7 \cdot y_3 \\
&= 0.2 \cdot 0.5 - 0.3 \cdot 0.0 - 0.1 \cdot 0.5744 + 0.1 \cdot 1.0 + 0.2 \cdot 0.4750 \\
&\approx 0.2376
\end{aligned}
$$

1.3.2 Backwards Pass

$$
\begin{aligned}
\beta_{out} &= y_{out} - \hat{y}_{out} \\
&= 1 - 0.2376 \\
&= 0.7624
\end{aligned}
$$

$$
\begin{aligned}
\beta_1 &= y_1 \cdot (1 - y_1) \cdot \beta_{out} \cdot w_9 \\
&= 0.5 \cdot (1 - 0.5) \cdot 0.7624 \cdot 0.2 \\
&\approx 0.0381
\end{aligned}
$$

$$
\begin{aligned}
\beta_2 &= y_2 \cdot (1 - y_2) \cdot \beta_{out} \cdot w_8 \\
&= 0.5744 \cdot (1 - 0.5744) \cdot 0.7624 \cdot (-0.1) \\
&\approx -0.0186
\end{aligned}
$$

$$
\begin{aligned}
\beta_3 &= y_3 \cdot (1 - y_3) \cdot \beta_{out} \cdot w_7 \\
&= 0.4750 \cdot (1 - 0.4750) \cdot 0.7624 \cdot 0.2 \\
&\approx 0.0380
\end{aligned}
$$

1.3.3 Weight Changes

1. Calculating δ_n:

$$\begin{aligned}\delta_9 &= \eta \cdot \beta_{out} \cdot y_1 \\ &= 0.2 \cdot 0.7624 \cdot 0.5 \\ &\approx 0.0762\end{aligned}$$

$$\begin{aligned}\delta_2 &= \eta \cdot \beta_{out} \cdot x_1 \\ &= 0.2 \cdot 0.7624 \cdot 0 \\ &= 0\end{aligned}$$

$$\begin{aligned}\delta_8 &= \eta \cdot \beta_{out} \cdot y_2 \\ &= 0.2 \cdot 0.7624 \cdot 0.5744 \\ &\approx 0.0876\end{aligned}$$

$$\begin{aligned}\delta_5 &= \eta \cdot \beta_{out} \cdot x_2 \\ &= 0.2 \cdot 0.7624 \cdot 1 \\ &\approx 0.1525\end{aligned}$$

$$\begin{aligned}\delta_7 &= \eta \cdot \beta_{out} \cdot y_3 \\ &= 0.2 \cdot 0.7624 \cdot 0.4750 \\ &\approx 0.0724\end{aligned}$$

$$\begin{aligned}\delta_1 &= \eta \cdot \beta_1 \cdot x_1 \\ &= 0.2 \cdot 0.0381 \cdot 0 \\ &= 0\end{aligned}$$

$$\begin{aligned}\delta_3 &= \eta \cdot \beta_2 \cdot x_1 \\ &= 0.2 \cdot -0.0186 \cdot 0 \\ &= 0\end{aligned}$$

$$\begin{aligned}\delta_4 &= \eta \cdot \beta_2 \cdot x_2 \\ &= 0.2 \cdot -0.0186 \cdot 1 \\ &\approx -0.0037\end{aligned}$$

$$\begin{aligned}\delta_6 &= \eta \cdot \beta_3 \cdot x_2 \\ &= 0.2 \cdot 0.0380 \cdot 1 \\ &\approx 0.0076\end{aligned}$$

2. Updating Weights:

$$\begin{aligned}w_9 &= w_9 + \delta_9 \\ &= 0.2 + 0.0762 \\ &= 0.2762\end{aligned}$$

$$\begin{aligned}w_2 &= w_2 + \delta_2 \\ &= -0.3 + 0 \\ &= -0.3\end{aligned}$$

$$\begin{aligned}w_8 &= w_8 + \delta_8 \\ &= -0.1 + 0.0876 \\ &= -0.0124\end{aligned}$$

$$\begin{aligned}w_5 &= w_5 + \delta_5 \\ &= 0.1 + 0.1525 \\ &= 0.2525\end{aligned}$$

$$\begin{aligned}w_7 &= w_7 + \delta_7 \\ &= 0.2 + 0.0724 \\ &= 0.2724\end{aligned}$$

$$\begin{aligned}w_1 &= w_1 + \delta_1 \\ &= -0.1 + 0 \\ &= -0.1\end{aligned}$$

$$\begin{aligned}w_3 &= w_3 + \delta_3 \\ &= -0.2 + 0 \\ &= -0.2\end{aligned}$$

$$\begin{aligned}w_4 &= w_4 + \delta_4 \\ &= 0.3 - 0.0037 \\ &= 0.2963\end{aligned}$$

$$\begin{aligned}w_6 &= w_6 + \delta_6 \\ &= -0.1 + 0.0076 \\ &= -0.0924\end{aligned}$$

3. Computed Weights:

w_1	w_2	w_3	w_4	w_5	w_6	w_7	w_8	w_9
-0.1	-0.3	-0.2	0.2963	0.2525	-0.0924	0.2724	-0.0124	0.2762

Table 3: Updated Weights after First Training Example

1.4 Second Training Example

Using the updated weights the second training example is processed.

1.4.1 Forward Pass

$$
\begin{aligned}
y_1 &= \sigma(w_1 \cdot x_1) \\
&= \sigma(-0.1 \cdot 1.0) \\
&= \sigma(-0.1) \\
&\approx 0.4750
\end{aligned}
$$

$$
\begin{aligned}
y_2 &= \sigma(w_3 \cdot x_1 + w_4 \cdot x_2) \\
&= \sigma(-0.2 \cdot 1.0 + 0.2963 \cdot 0.0) \\
&= \sigma(-0.2) \\
&\approx 0.4502
\end{aligned}
$$

$$
\begin{aligned}
y_3 &= \sigma(w_6 \cdot x_2) \\
&= \sigma(-0.0924 \cdot 0.0) \\
&= \sigma(0) \\
&= 0.5
\end{aligned}
$$

$$
\begin{aligned}
\hat{y}_{out} &= w_9 \cdot y_1 + w_2 \cdot x_1 + w_8 \cdot y_2 + w_5 \cdot x_2 + w_7 \cdot y_3 \\
&= 0.2762 \cdot 0.4750 - 0.3 \cdot 1 - 0.0124 \cdot 0.4502 \\
&\quad + 0.2525 \cdot 0.0 + 0.2724 \cdot 0.5 \\
&\approx -0.0382
\end{aligned}
$$

1.4.2 Backwards Pass

$$
\begin{aligned}
\beta_{out} &= y_{out} - \hat{y}_{out} \\
&= 1 - (-0.0382) \\
&= 1.0382
\end{aligned}
$$

$$
\begin{aligned}
\beta_1 &= y_1 \cdot (1 - y_1) \cdot \beta_{out} \cdot w_9 \\
&= 0.4750 \cdot (1 - 0.4750) \cdot 1.0382 \cdot 0.2762 \\
&\approx 0.0715
\end{aligned}
$$

$$
\begin{aligned}
\beta_2 &= y_2 \cdot (1 - y_2) \cdot \beta_{out} \cdot w_8 \\
&= 0.4502 \cdot (1 - 0.4502) \cdot 1.0382 \cdot (-0.0124) \\
&\approx -0.0032
\end{aligned}
$$

$$
\begin{aligned}
\beta_3 &= y_3 \cdot (1 - y_3) \cdot \beta_{out} \cdot w_7 \\
&= 0.5 \cdot (1 - 0.5) \cdot 1.0382 \cdot 0.2724 \\
&\approx 0.0707
\end{aligned}
$$

1.4.3 Weight Changes

1. Calculating δ_n:

$$\begin{aligned}\delta_9 &= \eta \cdot \beta_{out} \cdot y_1 \\ &= 0.2 \cdot 1.0382 \cdot 0.4750 \\ &\approx 0.0986\end{aligned}$$

$$\begin{aligned}\delta_2 &= \eta \cdot \beta_{out} \cdot x_1 \\ &= 0.2 \cdot 1.0382 \cdot 1 \\ &\approx 0.2076\end{aligned}$$

$$\begin{aligned}\delta_8 &= \eta \cdot \beta_{out} \cdot y_2 \\ &= 0.2 \cdot 1.0382 \cdot 0.4502 \\ &\approx 0.0935\end{aligned}$$

$$\begin{aligned}\delta_5 &= \eta \cdot \beta_{out} \cdot x_2 \\ &= 0.2 \cdot 1.0382 \cdot 0 \\ &= 0\end{aligned}$$

$$\begin{aligned}\delta_7 &= \eta \cdot \beta_{out} \cdot y_3 \\ &= 0.2 \cdot 1.0382 \cdot 0.5 \\ &\approx 0.1038\end{aligned}$$

$$\begin{aligned}\delta_1 &= \eta \cdot \beta_1 \cdot x_1 \\ &= 0.2 \cdot 0.0715 \cdot 1 \\ &\approx 0.0143\end{aligned}$$

$$\begin{aligned}\delta_3 &= \eta \cdot \beta_2 \cdot x_1 \\ &= 0.2 \cdot -0.0032 \cdot 1 \\ &\approx -0.0006\end{aligned}$$

$$\begin{aligned}\delta_4 &= \eta \cdot \beta_2 \cdot x_2 \\ &= 0.2 \cdot -0.0032 \cdot 0 \\ &= 0\end{aligned}$$

$$\begin{aligned}\delta_6 &= \eta \cdot \beta_3 \cdot x_2 \\ &= 0.2 \cdot 0.0380 \cdot 0 \\ &= 0\end{aligned}$$

2. Updating Weights:

$$\begin{aligned}w_9 &= w_9 + \delta_9 \\ &= 0.2762 + 0.0986 \\ &= 0.3748\end{aligned}$$

$$\begin{aligned}w_2 &= w_2 + \delta_2 \\ &= -0.3 + 0.2076 \\ &= -0.0924\end{aligned}$$

$$\begin{aligned}w_8 &= w_8 + \delta_8 \\ &= -0.0124 + 0.0935 \\ &= 0.0811\end{aligned}$$

$$\begin{aligned}w_5 &= w_5 + \delta_5 \\ &= 0.2525 + 0 \\ &= 0.2525\end{aligned}$$

$$\begin{aligned}w_7 &= w_7 + \delta_7 \\ &= 0.2724 + 0.1038 \\ &= 0.3762\end{aligned}$$

$$\begin{aligned}w_1 &= w_1 + \delta_1 \\ &= -0.1 + 0.0143 \\ &= -0.0857\end{aligned}$$

$$\begin{aligned}w_3 &= w_3 + \delta_3 \\ &= -0.2 + (-0.0006) \\ &= -0.2006\end{aligned}$$

$$\begin{aligned}w_4 &= w_4 + \delta_4 \\ &= 0.2963 + 0 \\ &= 0.2963\end{aligned}$$

$$\begin{aligned}w_6 &= w_6 + \delta_6 \\ &= -0.0924 + 0 \\ &= -0.0924\end{aligned}$$

3. Computed Weights:

w_1	w_2	w_3	w_4	w_5	w_6	w_7	w_8	w_9
-0.0857	-0.0924	-0.2006	0.2963	0.2525	-0.0924	0.3762	0.0811	0.3748

Table 4: Updated Weights after Second Training Example