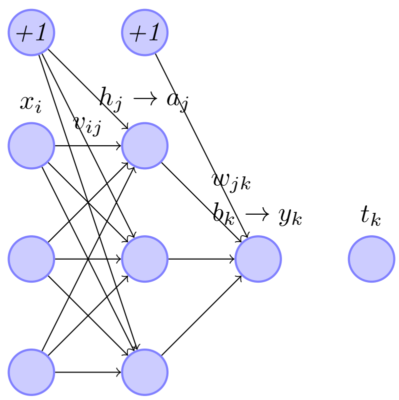
MLP

Wayne Wang

1 Principles of Backpropagation



Suppost we have a simple binary classification problem that needs to be solved by MLP. The structure of the MLP is shown as above which has 3 input nodes for input layer, 3 hidden nodes for one hidden layer and 1 output node for output layer.

The activation function is sigmoid function: $g(x) = \frac{1}{1+e^{-x}}$ and $g'(x) = g(x)(1 - g(x))$.

Step 1: **Going forward**

$$h_j = \sum_i x_i \cdot v_{ij} \tag{1}$$

$$a_j = g(h_j) \tag{2}$$

$$b_k = \sum_j a_j \cdot w_{jk} \tag{3}$$

$$y_k = g(b_k) \tag{4}$$

Step 2: **Compute Error**

$$
\begin{aligned}
E(\mathbf{v}, \mathbf{w}) &= \frac{1}{2}\sum_k (t_k - y_k)^2 \\
&= \frac{1}{2}\sum_k (t_k - g(b_k)^2 \\
&= \frac{1}{2}\sum_k (t_k - g(\sum_j a_j \cdot w_{jk}))^2 \\
&= \frac{1}{2}\sum_k (t_k - g(\sum_j g(h_j) \cdot w_{jk})^2 \\
&= \frac{1}{2}\sum_k (t_k - g(\sum_j g(\sum_i x_i \cdot v_{ij}) \cdot w_{jk})^2
\end{aligned}
\tag{5}
$$

Step 3: Compute the derivative of Error function respect to **w** and **v**

$$
\begin{aligned}
\frac{\partial E}{\partial w_{jk}} &= \frac{\partial}{\partial w_{jk}}(\frac{1}{2}\sum_k (t_k - y_k)^2) \\
&= \sum_k (t_k - y_k) \cdot \frac{\partial}{\partial w_{jk}}(t_k - y_k) \\
&= \sum_k (t_k - y_k) \cdot \frac{\partial}{\partial w_{jk}}(-g(b_k)) \\
&= \sum_k (t_k - y_k) \cdot -g'(b_k) \cdot \frac{\partial}{\partial w_{jk}}(\sum_j a_j \cdot w_{jk}) \\
&= \sum_k (y_k - t_k) \cdot y_k \cdot (1 - y_k) \cdot a_j
\end{aligned}
\tag{6}
$$

$$
\begin{aligned}
\frac{\partial E}{\partial v_{ij}} &= \frac{\partial E}{\partial a_j} \cdot \frac{\partial a_j}{\partial v_{ij}} \\
&= \frac{\partial}{\partial a_j}(\frac{1}{2}\sum_k (t_k - g(\sum_j a_j \cdot w_{jk}))^2) \cdot \frac{\partial}{\partial v_{ij}}(g(\sum_i x_i \cdot v_{ij})) \\
&= (\sum_k (y_k - t_k) \cdot y_k \cdot (1 - y_k) \cdot w_{jk}) \cdot a_j \cdot (1 - a_j) \cdot x_i
\end{aligned}
\tag{7}
$$

Step 4: Update weight **w** and **v**

$$
w_j k = w_j k - \eta \frac{\partial E}{\partial w_{jk}} \tag{8}
$$

$$
v_i j = v_i j - \eta \frac{\partial E}{\partial v_{ij}} \tag{9}
$$