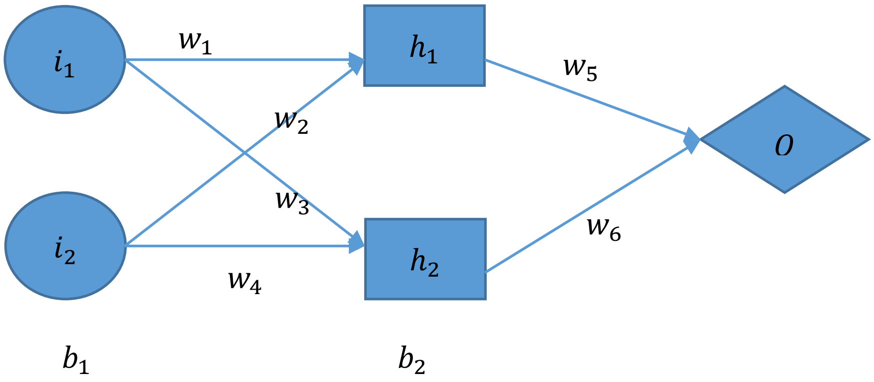
以带 1 层隐藏层的前馈型神经网络为例解释反向传播法（Back Propagation,BP）



其中，i_1,i_2是输入层神经元，h_1,h_2是隐藏层神经元，O 是输出层神经元，b_1,b_2分别是输入层和隐藏层的偏置。除了输入层外，其他两层的神经元接受上层输入的数据，经激活函数变换后输出至下一层（输出层神经元的输出是模型最终输出）。假设隐藏层和输出层的激活函数均为 Sigmoid，即

$$f(x) = \frac{1}{1+e^{-x}}$$

Sigmoid 函数的导函数：

$$f'(x) = f(x)(1 - f(x))$$

h_1,h_2的输入是i_1,i_2的加权再加上偏置b_1，即：

$$net_{h_1} = i_1 w_1 + i_2 w_2 + b_1$$
$$net_{h_2} = i_1 w_3 + i_2 w_4 + b_1$$

h_1,h_2的输出是激活函数作用在输入上：

$$out_{h_1} = f\left(net_{h_1}\right)$$

$$out_{h_2} = f\left(net_{h_2}\right)$$

同样的，O 的输入来自h_1,h_2的输出的加权再加上偏置b_2，即

$$net_O = out_{h_1} w_5 + out_{h_2} w_6 + b_2$$

O 的输出是激活函数作用在输入上：

$$out_O = f(net_O)$$

再假设损失函数是均方误差：

$$E = \frac{1}{2}(target - output)^2 = \frac{1}{2}(target - out_O)^2$$

则根据链式法则可求出 E 对w_5的梯度：

$$\frac{\partial E}{\partial w_5} = \frac{\partial E}{\partial out_O}\frac{\partial out_O}{\partial net_O}\frac{\partial net_O}{\partial w_5}$$

由于

$$\frac{\partial E}{\partial out_O} = -(target - out_O)$$

$$\frac{\partial out_O}{\partial net_O} = f'(net_O) = f(net_O)\left(1 - f(net_O)\right) = out_O(1 - out_O)$$

$$\frac{\partial net_O}{\partial w_5}=\frac{\partial(out_{h_1}w_5+out_{h_2}w_6+b_2)}{\partial w_5}=out_{h_1}$$

故

$$\frac{\partial E}{\partial w_5}=-(target-out_O)out_O(1-out_O)out_{h_1}$$

类似可求出$\frac{\partial E}{\partial w_6}$和$\frac{\partial E}{\partial b_2}$

E 对w_1的梯度是：

$$\frac{\partial E}{\partial w_1}=\frac{\partial E}{\partial out_O}\frac{\partial out_O}{\partial net_O}\frac{\partial net_O}{\partial out_{h_1}}\frac{\partial out_{h_1}}{\partial net_{h_1}}\frac{\partial net_{h_1}}{\partial w_1}$$

其中$\frac{\partial E}{\partial out_O}$、$\frac{\partial out_O}{\partial net_O}$已经求出，而

$$\frac{\partial net_O}{\partial out_{h_1}}=\frac{\partial(out_{h_1}w_5+out_{h_2}w_6+b_2)}{\partial out_{h_1}}=w_5$$

$$\frac{\partial out_{h_1}}{\partial net_{h_1}}=f'\left(net_{h_1}\right)=f\left(net_{h_1}\right)\left(1-f\left(net_{h_1}\right)\right)=out_{h_1}(1-out_{h_1})$$

$$\frac{\partial net_{h_1}}{\partial w_1}=\frac{\partial(i_1w_1+i_2w_2+b_1)}{\partial w_1}=i_1$$

故

$$\frac{\partial E}{\partial w_1}=-(target-out_O)out_O(1-out_O)w_5out_{h_1}(1-out_{h_1})i_1$$

结合梯度下降法，w_1和w_5的更新公式为：

$$w_1^+=w_1-\eta\frac{\partial E}{\partial w_1}$$

$$w_5^+=w_5-\eta\frac{\partial E}{\partial w_5}$$

同理可求出其他参数的更新公式。