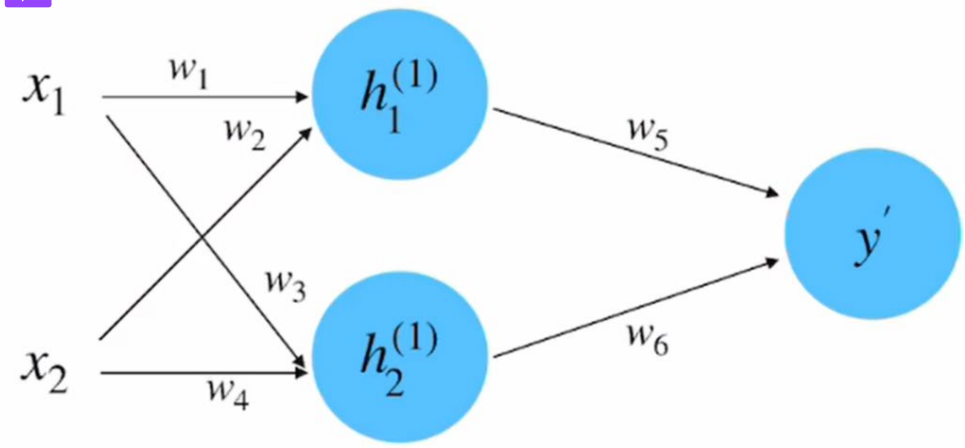
构建一个只有一层的神经网络



假设这个样本的特征有两个维度，
每个维度的值为

$$x_1 = 0.5$$
$$x_2 = 1.0$$

假设这个样本的真实值为：

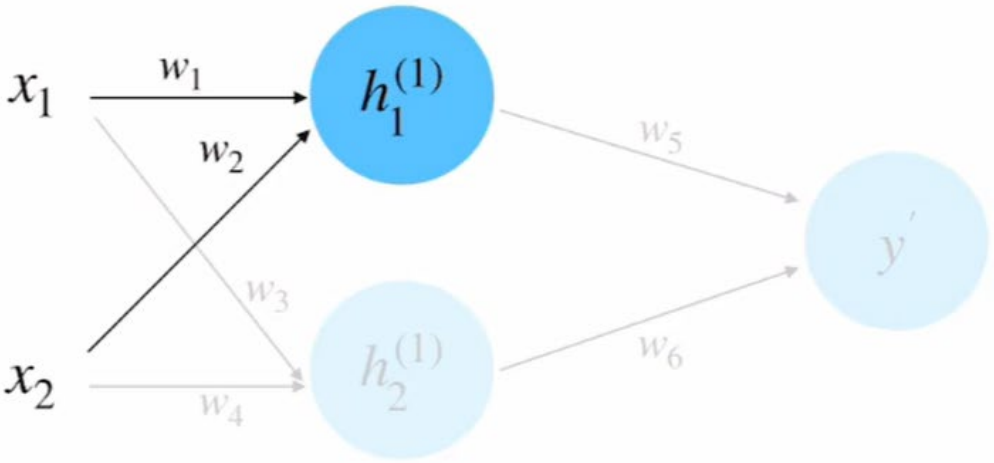
$$y = 0.8$$

神经网络所有参数
随机初始化，且恰好如下

$$w_1 = 1.0$$
$$w_2 = 0.5$$
$$w_3 = 0.5$$
$$w_4 = 0.7$$
$$w_5 = 1.0$$
$$w_6 = 2.0$$

正向传播：计算第一个神经元的输出

$$
\begin{aligned}
h_1^{(1)} &= w_1 \cdot x_1 + w_2 \cdot x_2 \\
&= 1.0 \cdot 0.5 + 0.5 \cdot 1.0 \\
&= 1.0
\end{aligned}
$$



假设这个样本的特征有两个维度，
每个维度的值为

$$
\begin{aligned}
x_1 &= 0.5 \\
x_2 &= 1.0
\end{aligned}
$$

假设这个样本的真实值为：

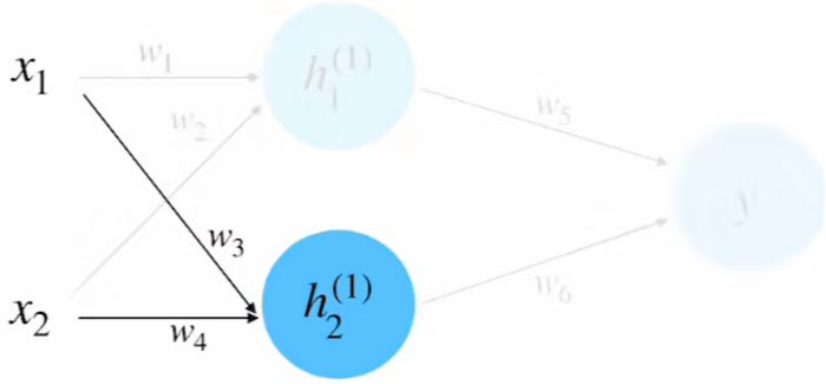
$$y = 0.8$$

神经网络所有参数
随机初始化，且恰好如下

$$
\begin{aligned}
w_1 &= 1.0 \\
w_2 &= 0.5 \\
w_3 &= 0.5 \\
w_4 &= 0.7 \\
w_5 &= 1.0 \\
w_6 &= 2.0
\end{aligned}
$$

正向传播：计算第二个神经元的输出

$$
\begin{aligned}
h_2^{(1)} &= w_3 \cdot x_1 + w_4 \cdot x_2 \\
&= 0.5 \cdot 0.5 + 0.7 \cdot 1.0 \\
&= 0.95
\end{aligned}
$$



假设这个样本的特征有两个维度，
每个维度的值为

$$
\begin{aligned}
x_1 &= 0.5 \\
x_2 &= 1.0
\end{aligned}
$$

假设这个样本的真实值为：

$$y = 0.8$$

神经网络所有参数
随机初始化，且恰好如下

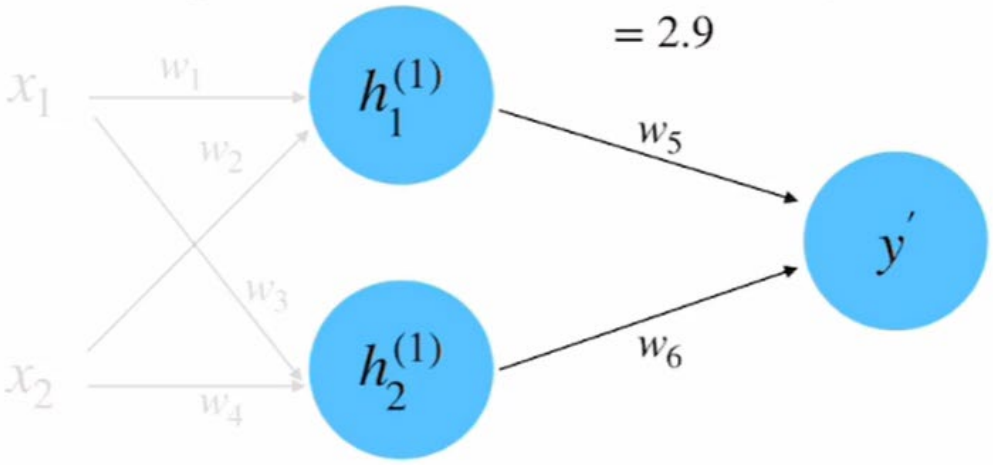
$$
\begin{aligned}
w_1 &= 1.0 \\
w_2 &= 0.5 \\
w_3 &= 0.5 \\
w_4 &= 0.7 \\
w_5 &= 1.0 \\
w_6 &= 2.0
\end{aligned}
$$

正向传播：计算最后一个神经元的输出

$$h_1^{(1)} = 1.0$$

$$h_2^{(1)} = 0.95$$

$$\begin{aligned} y' &= w_5 \cdot h_1^{(1)} + w_6 \cdot h_2^{(1)} \\ &= 1.0 \cdot 1.0 + 2.0 \cdot 0.95 \\ &= 2.9 \end{aligned}$$



假设这个样本的特征有两个维度，
每个维度的值为

$$x_1 = 0.5$$

$$x_2 = 1.0$$

假设这个样本的真实值为：

$$y = 0.8$$

神经网络所有参数
随机初始化，且恰好如下

$$w_1 = 1.0$$

$$w_2 = 0.5$$

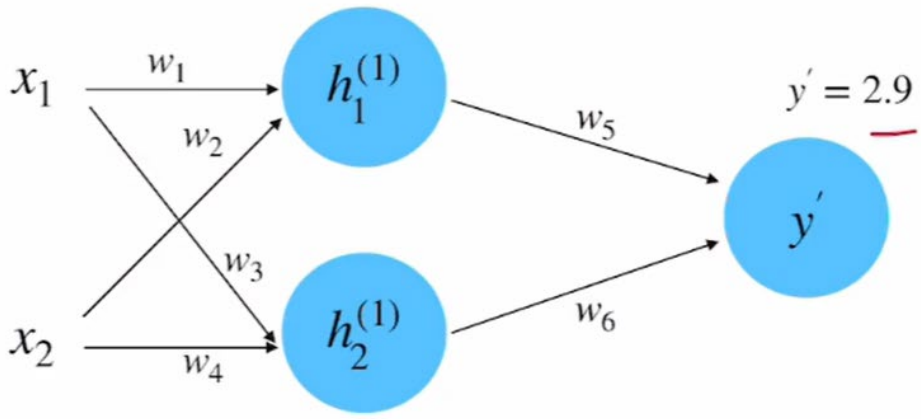
$$w_3 = 0.5$$

$$w_4 = 0.7$$

$$w_5 = 1.0$$

$$w_6 = 2.0$$

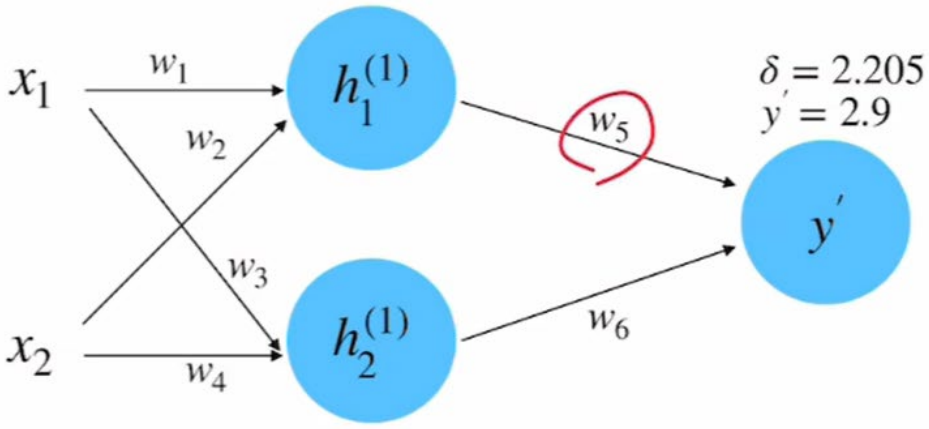
正向传播完成



这一轮正向传播后
误差如右所示：
(这个公式是损失函数的计算，
一般不同任务，损失函数会有所不同，
这里暂不详细说损失函数的问题)

$$\delta = \frac{1}{2}(y - y')^2$$
$$= 0.5(0.8 - 2.9)^2$$
$$= 2.205$$

backward propagation反向传播



以计算w_5为例，w_6同理

$$\frac{\partial \delta}{\partial w_5} = \frac{\partial \delta}{\partial y'} \cdot \frac{\partial y'}{\partial w_5}$$

$$\delta = \frac{1}{2}(y - y')^2$$

$$\frac{\partial \delta}{\partial y'} = 2 \cdot \frac{1}{2} \cdot (y - y')(-1)$$

$$= y' - y$$

$$= 2.9 - 0.8 = 2.1$$

$$y' = w_5 \cdot h_1^{(1)} + w_6 \cdot h_2^{(1)}$$

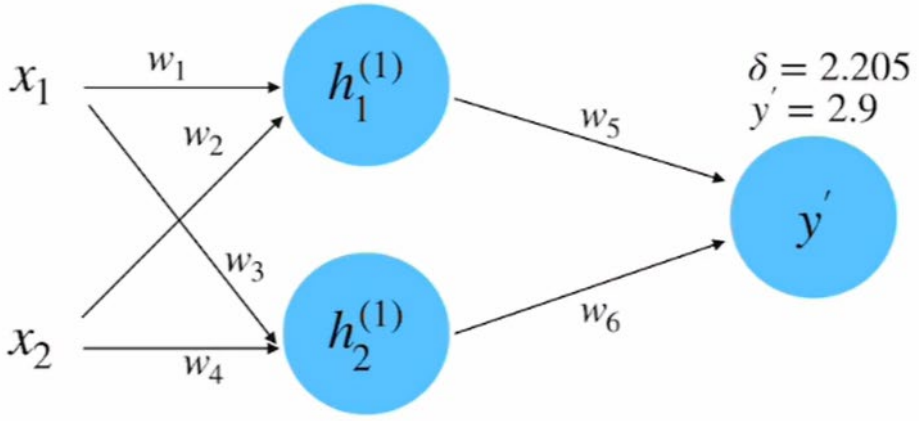
$$\frac{\partial y'}{\partial w_5} = h_1^{(1)} + 0$$

$$= 1.0$$

全选

$$\frac{\partial \delta}{\partial w_5} = 2.1 \cdot 1.0 = 2.1$$

backward propagation反向传播



由前一页可知

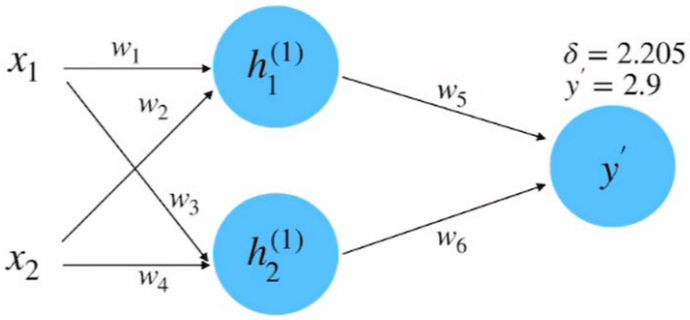
$$\frac{\partial \delta}{\partial w_5} = 2.1 \cdot 1.0 = 2.1$$

那么便可更新w5的值

学习率，属于超参数，自行设定

$$\begin{aligned} w_5^{(update)} &= w_5 - \eta \cdot \frac{\partial \delta}{\partial w_5} \\ &= 1.0 - 0.1 \cdot 2.1 \\ &= 0.79 \end{aligned}$$

backward propagation反向传播



以计算w_1为例，
w_2, w_3, w_4同理

$$\frac{\partial \delta}{\partial w_1} = \frac{\partial \delta}{\partial y'} \cdot \frac{\partial y'}{\partial h_1^{(1)}} \cdot \frac{\partial h_1^{(1)}}{\partial w_1}$$

$$\delta = \frac{1}{2}(y - y')^2$$

$$\frac{\partial \delta}{\partial y'} = 2 \cdot \frac{1}{2} \cdot (y - y')(-1)$$
$$= y' - y$$
$$= 2.9 - 0.8$$
$$= 2.1$$

$$y' = w_5 \cdot h_1^{(1)} + w_6 \cdot h_2^{(1)}$$

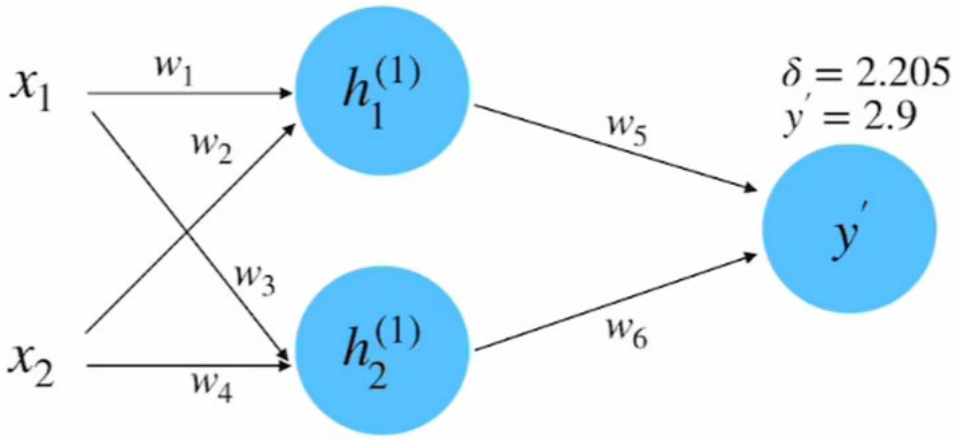
$$\frac{\partial y'}{\partial h_1^{(1)}} = w_5 + 0$$
$$= 1.0$$

$$h_1^{(1)} = w_1 \cdot x_1 + w_2 \cdot x_2$$

$$\frac{\partial h_1^{(1)}}{\partial w_1} = x_1 + 0$$
$$= 0.5$$

$$\frac{\partial \delta}{\partial w_1} = 2.1 \cdot 1.0 \cdot 0.5 = 1.05$$

backward propagation反向传播



由前一页可知

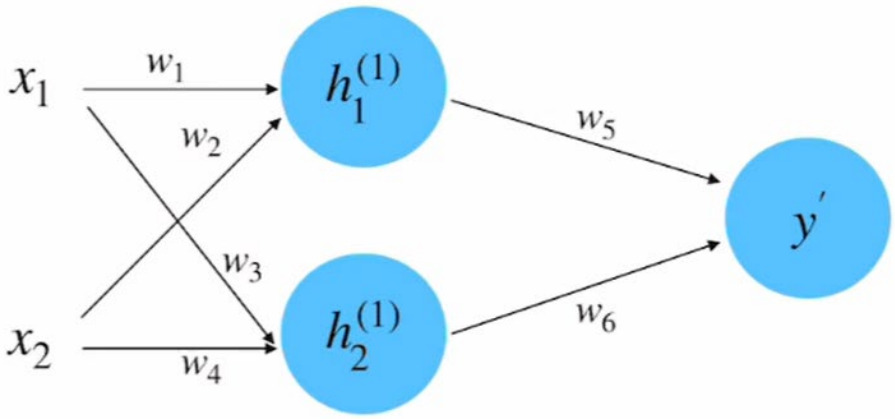
$$\frac{\partial \delta}{\partial w_1} = 2.1 \cdot 1.0 \cdot 0.5 = 1.05$$

学习率，属于超参数，自行设定

那么便可更新w_1的值

$$\begin{aligned} w_1^{(update)} &= w_1 - \eta \cdot \frac{\partial \delta}{\partial w_1} \\ &= 1.0 - 0.1 \cdot 1.05 \\ &= 0.895 \end{aligned}$$

backward propagation反向传播



假设这个样本的特征有两个维度，每个维度的值为

$$x_1 = 0.5$$
$$x_2 = 1.0$$

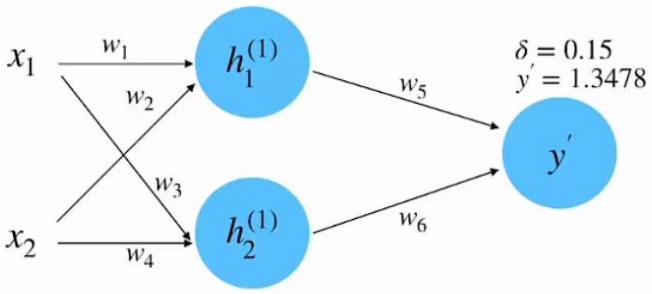
假设这个样本的真实值为：

$$y = 0.8$$

一轮正向反向传播后更新得到的参数：

$$w_1 = 0.895$$
$$w_2 = 0.29$$
$$w_3 = 0.29$$
$$w_4 = 0.28$$
$$w_5 = 0.79$$
$$w_6 = 1.8005$$

第二轮正向传播得到的误差



$$h_1^{(1)} = w_1 \cdot x_1 + w_2 \cdot x_2 = 0.895 \cdot 0.5 + 0.29 \cdot 1.0 = 0.7375$$

$$h_2^{(1)} = w_3 \cdot x_1 + w_4 \cdot x_2 = 0.29 \cdot 0.5 + 0.28 \cdot 1.0 = 0.425$$

$$y' = w_5 \cdot h_1^{(1)} + w_6 \cdot h_2^{(1)} = 0.79 \cdot 0.7375 + 1.8005 \cdot 0.425 = 1.3478$$

误差比第一轮减小了 $\delta = \frac{1}{2}(y - y')^2$