神经网络图：
x_1 —w_1—→ h_1 —w_5—→ y
x_2 —w_2, w_3—→ h_2 —w_6—→ y
w_4

Q1: 为什么要求梯度？

Q2: 求关于谁的梯度？

$$y = f_w(x)$$

$$\Downarrow$$

$$\min_{w} \sum_{x} \| f_w(x) - y \|^2$$

E：误差，求得一组 w 使得 E 最小 (放用 SGD)

随机初始化 W，然后使用下面公式进行更新 W，使得 $f_w(x)$
更可能逼近 y：

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

真实值：
令 $x_1 = 1$
$x_2 = 0.5$

w_1 w_2 w_3 w_4 真实值分别为 1, 2, 3, 4

w_5 w_6 真实值分别 0.5，0.6

y 真实值 计算可得 4

模拟BP:

已知 $x_1 = 1$, $x_2 = 0.5$ $t = 4$, w_1, w_2 w_3 w_4 w_5 w_6 未知

随机初始化 $w_1 = 0.5$ $w_2 = 1.5$ $w_3 = 2.3$ $w_4 = 3$ $w_5 = 1$ $w_6 = 1$

① 前馈运算:

计算 h_1, h_2, y 和 误差 E:

$h_1 = w_1 x_1 + w_2 x_2 = 1.25$

$h_2 = w_3 x_1 + w_4 x_2 = 3.8$

$y = w_5 h_1 + w_6 h_2 = 5.05$

$E = \frac{1}{2}(y-t)^2 = 0.55125$

② 反向传播

$$\boxed{\frac{dE}{dw_5}} = \frac{dE}{dy} \cdot \frac{dy}{dw_5} = 1.05 \cdot h_1 = 1.05 \cdot 1.25 = 1.3125$$

$$\frac{dE}{dy} = \left[\frac{1}{2}(y-t)^2\right]' = 2 \cdot \frac{1}{2}(y-t) \cdot -1 = 1.05$$

$$\boxed{w_5^+} = w_5 - \eta \cdot \frac{dE}{dw_5} \quad (令 \eta = 0.1)$$

$$= 1 - 0.1 \times 1.312J$$

$$= 0.8687J$$

$$\boxed{\frac{\partial E}{\partial W_6}} = \frac{\partial E}{\partial y} \cdot \frac{\partial y}{\partial W_6} = 1.05 \cdot h_2 = 1.05 \times 3.8 = 3.99$$

$$\boxed{W_6^+} = W_6 - \eta \cdot \frac{\partial E}{\partial W_6}$$

$$= 1 - 0.1 \times 3.99$$

$$= 0.601$$

$$\boxed{\frac{\partial E}{\partial W_1}} = \frac{\partial E}{\partial y} \cdot \frac{\partial y}{\partial h_1} \cdot \frac{\partial h_1}{\partial W_1}$$

没有更新之前 W_5，不是 W_5^+

$$= 1.05 \cdot \boxed{W_5} \cdot x_1$$

$$= 1.05 \cdot 1 \cdot 1 = 1.05$$

$$\boxed{W_1^+} = W_1 - \eta \cdot \frac{\partial E}{\partial W_1}$$

$$= 0.5 - 0.1 \times 1.05$$

$$= 0.395$$

同样方法可以求得:

$$W_2^+ = 1.4475 \qquad W_3^+ = 2.195 \qquad W_4^+ = 2.9475$$

③ 前馈运算:

$$h_1 = W_1^+ \cdot x_1 + W_2^+ x_2 = 1.11875$$

$$h_2 = W_3^+ \cdot x_1 + W_4^+ \cdot x_4 = 3.66875$$

$$y = W_5^+ h_1 + W_6^+ h_2 = 3.1768$$

$$E = \frac{1}{2}(y-t)^2 = \boxed{0.3388}$$

这次误差比上一次前馈误差

0.55125 小