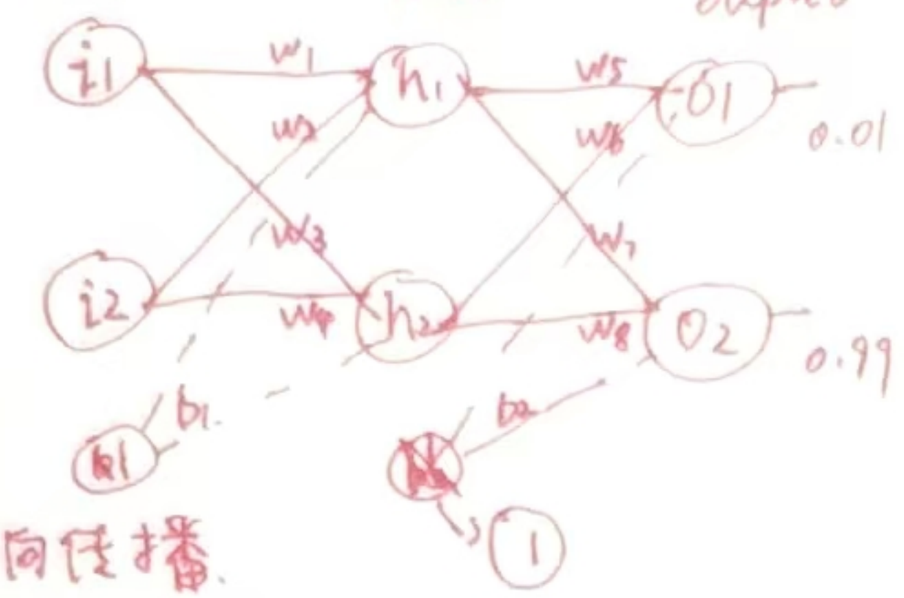
手推神经网络训练过程

以课上的例子为例.



$i_1 = 0.05$ $i_2 = 0.1$

$w_1 = 0.15$ $w_2 = 0.20$ $w_3 = 0.25$ $w_4 = 0.30$ $w_5 = 0.40$ $w_6 = 0.45$ $w_7 = 0.50$ $w_8 = 0.55$

$b_1 = 0.35$ $b_2 = 0.60$

step 1. 前向传播.

input → hide:

Z: 神经元的加权输入和

中间经过激活函数 sigmoid.

a: 神经元的加权输出和

$Z_{h_1} = i_1 \times w_1 + i_2 \times w_2 + b_1 \times 1$

$= 0.05 \times w_1 + 0.1 \times 0.2 + 0.35$

$= 0.3775$

↓ sigmoid →

$a_{h_1} = \frac{1}{1+e^{-z_{h_1}}} = \frac{1}{1+e^{-0.3775}}$

$= 0.593269992$

$Z_{h_2} = i_1 \times w_3 + i_2 \times w_4 + b_1 \times 1$

$= 0.3925$

$a_{h_2} = \frac{1}{1+e^{-z_{h_2}}} = 0.596884378$

hide → output:

$Z_{o_1} = a_{h_1} \times w_5 + a_{h_2} \times w_6 + b_2 \times 1$

$= 1.105905967$

$a_{o_1} = \frac{1}{1+e^{-z_{o_1}}} = 0.751365069$

$Z_{o_2} = a_{h_1} \times w_7 + a_{h_2} \times w_8 + b_2 \times 1$

$= 1.224921404$

$a_{o_2} = \frac{1}{1+e^{-z_{o_2}}} = 0.772928465$

step 1 正向传播结束.

输出结果.	0.75136…, 0.77292…	} 差很多.
正确答案	0.01 , 0.99.	

↓

反向传播. 更新权重. 重新输出

step 2. 反向传播.

1. 使用MSE损失函数.

$$E_{total} = \sum \frac{1}{n} (target - output)^2$$

$\frac{1}{n}$ ↓ 输出值个数.

分别计算两个值的 E，再相加.

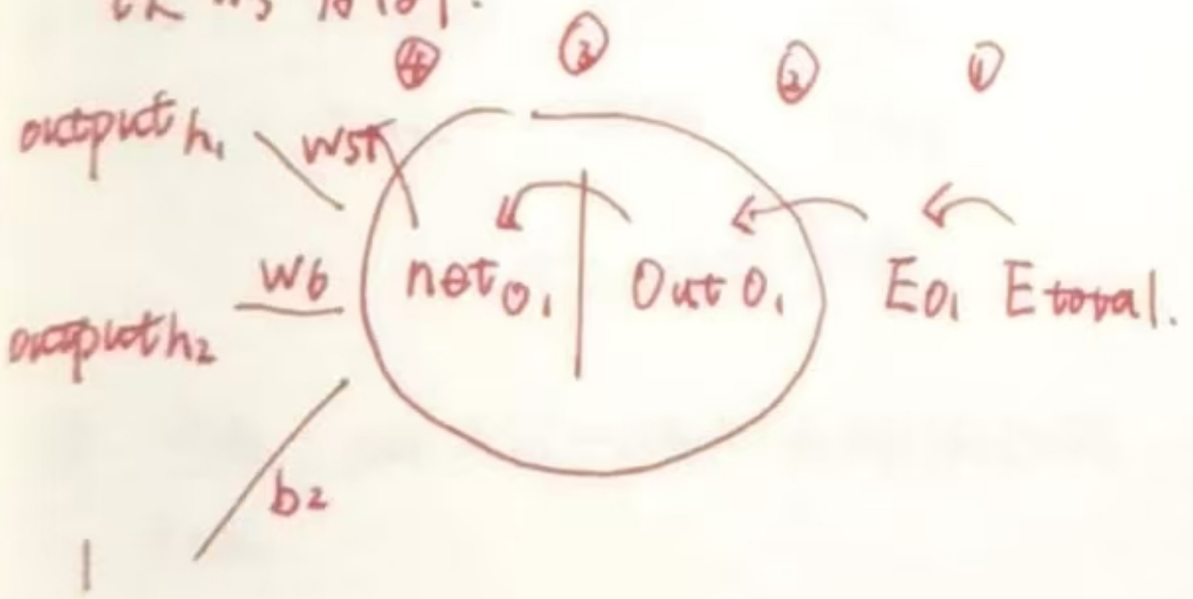
$$E_{O_1} = \frac{1}{2}(0.01 - 0.751365069)^2 = 0.274811083.$$

$$E_{O_2} = \frac{1}{2}(0.99 - 0.772928465)^2 = 0.023560026$$

$$E_{total} = E_{O_1} + E_{O_2} = 0.298371109.$$

2. 隐藏层 → 输出层 权值更新

以 w_5 为例：



↓ 求偏导

$$\frac{\partial E_{total}}{\partial w_5} = \frac{\partial E_{total}}{\partial out_{O_1}} \times \frac{\partial out_{O_1}}{\partial net_{O_1}} \times \frac{\partial net_{O_1}}{\partial w_5}$$

↓ 拆开算

② $E_{total} = \frac{1}{2}(target_{O_1} - a_{O_1})^2 + \frac{1}{2}(target_{O_2} - a_{O_2})^2$

$$\frac{\partial E_{total}}{\partial a_{O_1}} = 2 \times \frac{1}{2}(target_{O_1} - a_{O_1}) \times (-1) = 0.741365069$$

③ $a_{O_1} = \frac{1}{e^{-z_{O_1}} + 1}$

$$\frac{\partial a_{O_1}}{\partial z_{O_1}} = a_{O_1} \times (1 - a_{O_1}) = 0.186815062$$

④ $z_{O_1} = w_5 \times a_{h_1} + w_6 \times a_{h_2} + b_2 \times 1$ $\quad \frac{\partial z_{O_1}}{\partial w_5} = a_{h_1} = 0.093269992$

$$\therefore \frac{\partial E_{total}}{\partial w_5} = ② \times ③ \times ④ = 0.082167041.$$

↓ 更新 w_5.

$$w_5^+ = w_5 - \underbrace{\eta}_{\text{学习率}} \times \frac{\partial E_{total}}{\partial w_5}$$

$$= 0.4 - \underbrace{0.5}_{\text{经验值}} \times 0.082167041 = 0.35891648$$

同理：

$$w_6^+ = 0.408666186$$

$$w_7^+ = 0.511301270$$

$$w_8^+ = 0.561370121$$

3. 隐藏层 → 输隐藏层

以 w_1 为例.

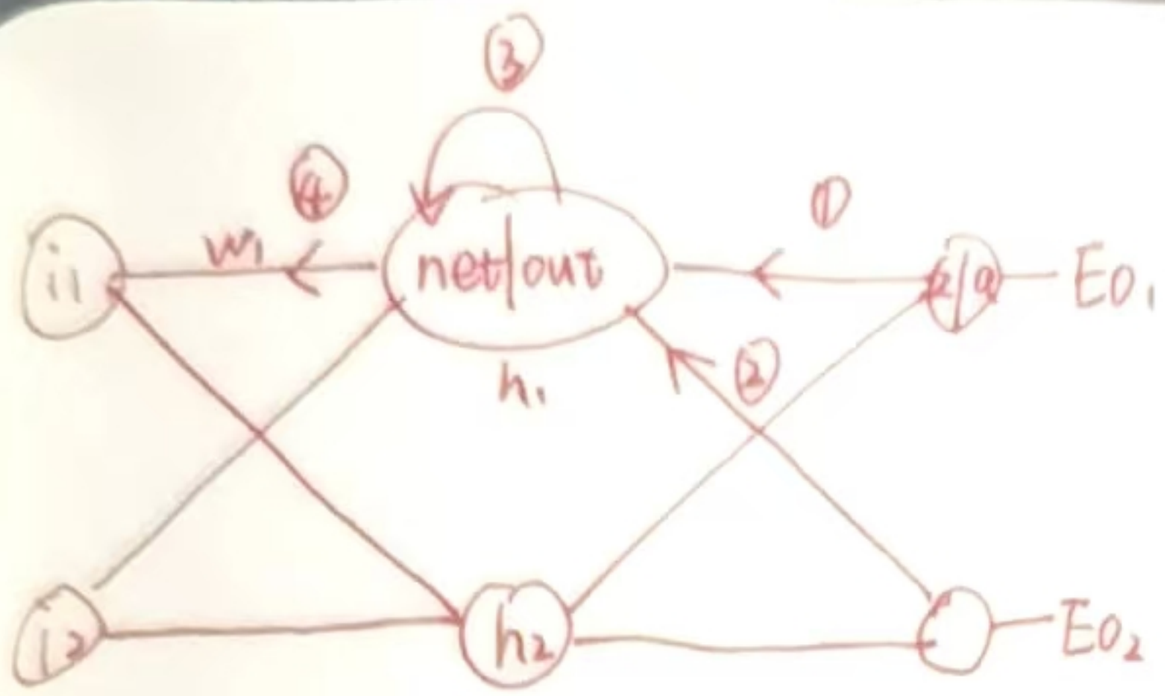
先看左 → 再看右

$$\frac{\partial E_{total}}{\partial w_1} = \frac{\partial E_{total}}{\partial out_{h_1}} \times \frac{\partial out_{h_1}}{\partial net_{h_1}} \times \frac{\partial net_{h_1}}{\partial w_1}$$

↓

$$\frac{\partial E_{total}}{\partial out_{h_1}} = \frac{\partial E_{O_1}}{\partial out_{h_1}} + \frac{\partial E_{O_2}}{\partial out_{h_1}}$$

($\because out_{h_1}$ 同时 influ E_{O_1}、E_{O_2})



$E_{total} = E_{o1} + E_{o2}$.

$$①: \frac{\partial E_{o1}}{\partial a_{h1}} = \frac{\partial E_{o1}}{\partial a_{o1}} \times \frac{\partial a_{o1}}{\partial z_{o1}} \times \frac{\partial z_{o1}}{\partial a_{h1}}$$

$$= 0.085539425.$$

同理②: $\frac{\partial E_{o2}}{\partial a_{h1}} = -0.\dot{1}9049119$

$$①+②: \frac{\partial E_{total}}{\partial a_{h1}} = \frac{\partial E_{o1}}{\partial a_{h1}} + \frac{\partial E_{o2}}{\partial a_{h1}}$$

$$= 0.03635306$$

$$③ \frac{\partial a_{h1}}{\partial z_{h1}} = a_{h1} \times (1 - a_{h1}) = 0.241300706$$

$$④ \frac{\partial z_{h1}}{\partial w_1} = i_1 = 0.05.$$

↓

$$\frac{\partial E_{total}}{\partial w_1} = (①+②) \times ③ \times ④ = 0.000438568$$

↓

更新 w_1 权值.

$$w_1^+ = w_1 - \eta \times \frac{\partial E_{total}}{\partial w_1} = 0.15 - 0.5 \times$$

$$0.000438568 = 0.149780716$$

同理. $w_2^+ = 0.19956143$

$w_3^+ = 0.24975114$

$w_4^+ = 0.29950229$

一次正向、反向完成. $w_1 \sim w_3$ 都更新了.

↓

再次迭代.

↓

直到误差满足要求，训练完成