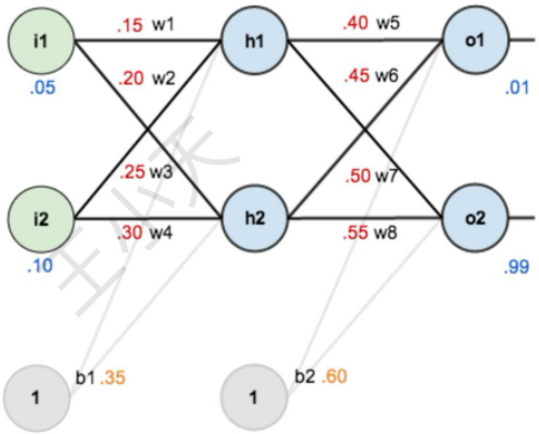
神经网络训练过程

- z：神经元的加权和
- a：神经元的输出
- 激活函数用 sigmoid



— step1 前向传播

① 输入层 → 隐藏层

$z_{h_1} = i_1 \cdot w_1 + i_2 \cdot w_2 + b_1 \cdot 1 = 0.3775$

$z_{h_2} = i_1 \cdot w_3 + i_2 \cdot w_4 + b_1 \cdot 1$

$a_{h_1} = \frac{1}{1+e^{-z_{h_1}}} = 0.593269992$

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

② 隐藏层 → 输出层

$z_{o1} = a_{h_1} \cdot w_5 + a_{h_2} \cdot w_7 + b_2 \cdot 1 = 1.105905967$

$a_{o1} = \frac{1}{1+e^{-z_{o1}}} = 0.751365069$

同理可得 $a_{o2} = 0.772928465$

— step2 反向传播

① 计算损失函数 $E_{total} = \sum \frac{1}{2}(target - output)^2$

$E_{o1} = \frac{1}{2} \cdot (0.01 - a_{o1})^2 = 0.274811083$

$E_{o2} = \frac{1}{2} \cdot (0.99 - a_{o2})^2 = 0.023560026$

$E_{total} = E_{o1} + E_{o2} = 0.298371109$

② 隐藏层 → 输出层的权值更新

$\frac{\partial E_{total}}{\partial w_5} = \frac{\partial E_{total}}{\partial a_{o1}} \cdot \frac{\partial a_{o1}}{\partial z_{o1}} \cdot \frac{\partial z_{o1}}{\partial w_5}$ ①

- 计算 $\frac{\partial E_{total}}{\partial a_{o1}}$：$E_{total} = \frac{1}{2} \cdot (target_{o1} - a_{o1})^2 + \frac{1}{2}(target_{o2} - a_{o2})^2$

  $\frac{\partial E_{total}}{\partial a_{o1}} = (target_{o1} - a_{o1}) \cdot (-1)$

- 计算 $\frac{\partial a_{o1}}{\partial z_{o1}}$：$a_{o1} = \frac{1}{1+e^{-z_{o1}}}$

  $\frac{\partial a_{o1}}{\partial z_{o1}} = a_{o1} \cdot (1 - a_{o1})$

- 计算 $\frac{\partial z_{o1}}{\partial w_5}$：$z_{o1} = a_{h_1} \cdot w_5 + a_{h_2} \cdot w_7 + b_2 \cdot 1$

  $\frac{\partial z_{o1}}{\partial w_5} = a_{h_1}$

故 $\frac{\partial E_{total}}{\partial w_5} = -(target_{o1} - a_{o1}) \cdot a_{o1} \cdot (1 - a_{o1}) \cdot a_{h_1} = 0.082167041$

更新 w_5 的值：$w_5^+ = w_5 - \eta \cdot \frac{\partial E_{total}}{\partial w_5} = 0.35819608$，$\eta$ 为学习率，

同理可更新：$w_6^+ = 0.40866186$

$w_7^+ = 0.51130120$

$w_8^+ = 0.56137021$

③ 隐藏层/输入层 → 隐藏层的权值更新

$$\frac{\partial E_{total}}{\partial w_1} = \frac{\partial E_{total}}{\partial a_{h1}} \cdot \frac{\partial a_{h1}}{\partial z_{h1}} \cdot \frac{\partial z_{h1}}{\partial w_1}$$

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

$$\frac{\partial E_{o1}}{\partial a_{h1}} = \frac{\partial E_{o1}}{\partial a_{o1}} \cdot \frac{\partial a_{o1}}{\partial z_{o1}} \cdot \frac{\partial z_{o1}}{\partial a_{h1}}$$

其中 $\frac{\partial E_{o1}}{\partial a_{o1}} = (Target_{o1} - a_{o1}) \cdot (-1)$

$\frac{\partial a_{o1}}{\partial z_{o1}} = a_{o1} \cdot (1 - a_{o1})$

$\frac{\partial z_{o1}}{\partial a_{h1}} = w_5$

$\therefore \frac{\partial E_{o1}}{\partial a_{h1}} = -(Target_{o1} - a_{o1}) \cdot a_{o1} \cdot (1 - a_{o1}) \cdot w_5$

$= -(0.01 - 0.75136507) \cdot 0.75136507 \cdot (1 - 0.75136507) \cdot 0.4$

$= 0.05539425$

同理可得 $\frac{\partial E_{o2}}{\partial a_{h1}} = -0.01904919$

$\frac{\partial E_{total}}{\partial a_{h1}} = \frac{\partial E_{o1}}{\partial a_{h1}} + \frac{\partial E_{o2}}{\partial a_{h1}} = 0.03635030$

$\frac{\partial a_{h1}}{\partial z_{h1}} = a_{h1} \cdot (1 - a_{h1})$

$\frac{\partial z_{h1}}{\partial w_1} = i_1$

$\frac{\partial E_{total}}{\partial w_1} = \frac{\partial E_{total}}{\partial a_{h1}} \cdot \frac{\partial a_{h1}}{\partial z_{h1}} \cdot \frac{\partial z_{h1}}{\partial w_1}$

$= \frac{\partial E_{total}}{\partial a_{h1}} \cdot a_{h1} \cdot (1 - a_{h1}) \cdot i_1$

$= 0.00049385 68$

更新 w_1 的权值：$w_1^+ = w_1 - \eta \cdot \frac{\partial E_{total}}{\partial w_1}$

$= 0.15 - 0.5 \cdot 0.00049385 68$

$= 0.14978071\,6$

同理，更新 w_2, w_3, w_0 权值：

$w_2^+ = 0.19956143$

$w_3^+ = 0.24975114$

$w_0^+ = 0.29950229$