手推训练过程

设输入层有2个神经元 $z_1=0.5$, $z_2=0.1$, 一层隐藏层, 2个隐藏神经元 h_1, h_2, 输出2个 O_1和$O_2=0.99$

一前向传播 输入层到隐藏层

1. 随机初始化参数 $w_1=0.2$, $w_2=0.3$, $w_3=0.1$, $w_4=0.4$
   $w_5=0.45$, $w_6=0.35$, $w_7=0.5$, $w_8=0.55$

神经元 h_1 的输入加权和:
$$z_{h_1} = w_1 z_1 + w_2 z_2 = 0.2 \times 0.5 + 0.3 \times 0.1 = 0.13$$

使用 Sigmoid 激活函数 h_1 的输出 a_{h_1}:
$$a_{h_1} = \frac{1}{1+e^{-z_{h_1}}} = \frac{1}{1+e^{-0.13}} = 0.532454$$

同理: $z_{h_2} = w_3 z_1 + w_4 z_2 = 0.1 \times 0.5 + 0.4 \times 0.1 = 0.09$
$$a_{h_2} = \frac{1}{1+e^{-z_{h_2}}} = \frac{1}{1+e^{-0.09}} = 0.522485$$

2. 隐藏层 → 输出层.

$z_{O_1} = w_5 \times a_{h_1} + w_6 \times a_{h_2} =$
$$= 0.45 \times 0.532454 + 0.35 \times 0.522485 = 0.422474$$
$$a_{O_1} = \frac{1}{1+e^{-z_{O_1}}} = \frac{1}{1+e^{-0.422474}} = 0.604075$$

$z_{O_2} = w_7 \times a_{h_1} + w_8 \times a_{h_2}$
$$= 0.5 * 0.532454 + 0.55 \times 0.522485 = 0.553594$$
$$a_{O_2} = \frac{1}{1+e^{-z_{O_2}}} = \frac{1}{1+e^{-0.553594}} = 0.634969$$

二、反向传播.

1. 隐藏层到输出层权值更新.

$$E_{total} = \frac{1}{2}(a_{o1} - O_1)^2 + \frac{1}{2}(a_{o2} - O_2)^2$$

以更新 w_5 为例：

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

$$\frac{\partial E_{total}}{\partial a_{o1}} = a_{o1} - O_1 = 0.604075 - 0.01 = \cancel{0.604} \quad 0.594075$$

$$\frac{\partial a_{o1}}{\partial z_{o1}} = a_{o1}(1 - a_{o1}) = 0.239168$$

$$\frac{\partial z_{o1}}{\partial w_5} = a_{h1} = 0.532454$$

$$\frac{\partial E_{total}}{\partial w_5} = 0.594075 \times 0.239168 \times 0.532454$$
$$= 0.075653$$

设学习率 $\eta = 0.5$

$$w_5^+ = w_5 - \eta * \frac{\partial E_{total}}{\partial w_5} = 0.45 - 0.5 \times 0.075653$$
$$= 0.412174$$

2. 隐藏层 → 隐藏层权值更新，以 w_1 为例)

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

$$\frac{\partial Z_{total}}{\partial a_{h1}} = \frac{\partial Z_{total}}{\partial a_{o1}} \cdot \frac{\partial a_{o1}}{\partial z_{o1}} \cdot w_5 +$$

$$\frac{\partial Z_{total}}{\partial a_{o2}} \cdot \frac{\partial a_{o2}}{\partial z_{o2}} \cdot w_7$$

$$= 0.594075 \times 0.239168 \times 0.45 + (-0.355031) \times 0.231783 \times 0.5$$

$$= 0.02276$$

$$w_1^+ = w_1 - \eta \times \frac{\partial Z_{total}}{}$$

$$\frac{\partial Z_{total}}{\partial w_1} = 0.02276 \times 0.532454 (1 - 0.532454) \times 0.5$$

$$= 0.002833$$

$$w_1^+ = w_1 - \eta \times \frac{\partial Z_{total}}{\partial w_1}$$

$$= 0.2 - 0.5 \times 0.002833 = 0.198583$$