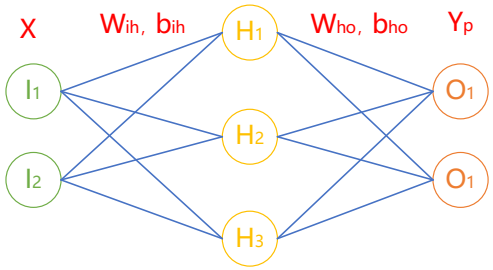
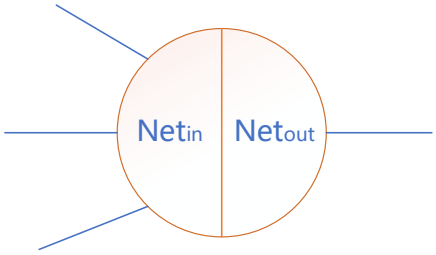
神经网络反向传播算法公式推导（by：GiffordY）





Netin = z = WX+b

Netout = a = f(z)

学习率：α

损失函数：$Loss = MSE = \sum_{i=1}^{n} \frac{1}{n}(y_{ti} - y_{pi})^2 = \sum_{i=1}^{n} \frac{1}{n}(y_{ti} - a_{oi})^2$

损失函数对 a_o 求导：$\frac{\partial Loss}{\partial a_o} = -\frac{2}{n}[(y_{t1} - a_{o1}) + ... + (y_{tn} - a_{on})] = -\frac{2}{n}(Y_t - a_o)$

激活函数：$f(z) = sigmoid(z) = \frac{1}{1+e^{-z}}$

激活函数求导：$f'(z) = \frac{\partial a}{\partial z} = (-1) \times \frac{1}{(1+e^{-z})^2} \times e^{-z}(-1) = f^2 \times (\frac{1}{f} - 1) = a(1-a)$

1、 输出层—>隐藏层：

（1）损失函数对 W_{ho} 的偏导：$\frac{\partial Loss}{\partial W_{ho}} = \frac{\partial Loss}{\partial a_o} \times \frac{\partial a_o}{\partial z_o} \bullet \times \frac{\partial z_o}{\partial W_{ho}}$

代入，得：$\frac{\partial Loss}{\partial W_{ho}} = -\frac{2}{n}(Y_t - a_o) \times a_o(1-a_o) \bullet \times a_h$

损失函数对 b_{ho} 的偏导：$\frac{\partial Loss}{\partial b_{ho}} = \frac{\partial Loss}{\partial a_o} \times \frac{\partial a_o}{\partial z_o} \times \frac{\partial z_o}{\partial b_{ho}}$

代入，得：$\frac{\partial Loss}{\partial b_{ho}} = -\frac{2}{n}(Y_t - a_o) \times a_o(1-a_o)$

（2）更新权重 W_{ho}：$W_{ho} = W_{ho} - \alpha \times \frac{\partial Loss}{\partial W_{ho}}$

b_{ho}：$b_{ho} = b_{ho} - \alpha \times \frac{\partial Loss}{\partial W_{ho}}$

2、 隐藏层—>输入层：

（1） 损失函数对 W_{ih} 的偏导：$\frac{\partial Loss}{\partial W_{ih}} = \frac{\partial Loss}{\partial a_h} \times \frac{\partial a_h}{\partial z_h} \bullet \times \frac{\partial z_h}{\partial W_{ih}}$

其中：$\frac{\partial Loss}{\partial a_h}=\frac{\partial Loss}{\partial a_o}\times\frac{\partial a_o}{\partial z_o}\times\frac{\partial z_o}{\partial a_h}=-\frac{2}{n}(Y_t-a_o)\times a_o(1-a_o)\bullet\times W_{ho}$

所以：$\frac{\partial Loss}{\partial W_{ih}}=-\frac{2}{n}(Y_t-a_o)\times a_o(1-a_o)\bullet\times W_{ho}\times a_h(1-a_h)\bullet\times X$

损失函数对b_{ih}的偏导：$\frac{\partial Loss}{\partial b_{ih}}=\frac{\partial Loss}{\partial a_h}\times\frac{\partial a_h}{\partial z_h}\times\frac{\partial z_h}{\partial b_{ih}}$

代入，得：$\frac{\partial Loss}{\partial b_{ih}}=-\frac{2}{n}(Y_t-a_o)\times a_o(1-a_o).\times W_{ho}\times a_h(1-a_h)=\frac{\partial Loss}{\partial a_h}\times a_h(1-a_h)$

（2） 更新权重W_{ih}：$W_{ih}=W_{ih}-\alpha\times\frac{\partial Loss}{\partial W_{ih}}$

b_{ih}：$b_{ih}=b_{ih}-\alpha\times\frac{\partial Loss}{\partial b_{ih}}$

注：公式中“.x”表示矩阵乘法，“.x”后面的变量需要先变为它的转置矩阵，左乘还是右乘要先推一下矩阵的 shape。