



已知条件：

激活函数  $f$  为 SIGMOD 函数；该函数的导数  $f'(x)=f(x)(1-f(x))$  （公式 1）

初始权重： $v_{11}=0.1, v_{12}=0.3, v_{21}=0.3, v_{22}=0.4; w_{11}=0.4, w_{12}=0.6$

初始输入数据： $X=[0.2 \ 0.6]$ , 该数据的真实标签  $y=0.7$

梯度下降的学习率  $\eta$  设置为 10

前向传播过程：

$$net_1 = v_{11} * X_1 + v_{12} * X_2 = 0.2$$

$$net_2 = v_{21} * X_1 + v_{22} * X_2 = 0.3$$

$$z_1 = f(net_1) = 0.5498$$

$$z_2 = f(net_2) = 0.5744$$

$$y_1 = f(w_{11} * z_1 + w_{12} * z_2) = f(0.5646) = 0.6375$$

因此，输出误差为：

$$\delta_1 = (y_d - y_1) f'(net) = (y_d - y_1) y_1 (1 - y_1) = 0.01444$$

正向传播阶段计算完成

反向传播过程：

$$w_{11} = w_{11} - (-\eta \delta_1 z_1) = 0.4794$$

$$w_{12} = w_{12} + \eta \delta_1 z_2 = 0.6830$$

$$v_{11} = v_{11} + \eta (\delta_1 * w_1) f'(net_1) X_1 = v_{11} + \eta (\delta_1 * w_1) f(net_1) (1 - f(net_1)) X_1 = v_{11} + \eta (\delta_1 * w_1) z_1 (1 - z_1) X_1 = 0.1029$$

（此处：只计算了一次导数  $f'$ ，且计算过程用到了前向传播过程中的中间结果）

类似地，

$$v_{12} = v_{12} + \eta (\delta_1 * w_1) f'(net_1) X_2 = 0.3086$$

$$v_{21} = v_{21} + \eta (\delta_1 * w_1) f'(net_2) X_1 = 0.3042$$

$$v_{21}=v_{22}+\eta(\delta_1*w_1)f'(net_2)X_2=0.4127$$

到此完成了反向传播的一次迭代。更新后的  $y_1$  为 0.6588，更接近实际标签 0.7