Lec 9 课程基本知识体系：神经网络习题

如图 2 所示多层感知机模型，第一层是输入层，包含两个神经元：x1=0.08，x2=0.12 和偏置 b1；第二层是隐藏层，包含两个神经元：h1，h2 和偏置项 b2；第三层是输出：y1，y2。每条线上标的 $w_{i,j}$ 是第 i 层第 j 个权重参数，激活函数是 sigmoid 函数（h 神经元之后），Loss 函数使用 MSE（均方误差）函数，真实标签 $Label_1 = 0.05, Label_2 = 0.95$，学习率 $\alpha = 0.5$，求在经过一次反向传播后所有权重参数和偏置项参数的值（写出计算过程，最后结果保留四位小数）。

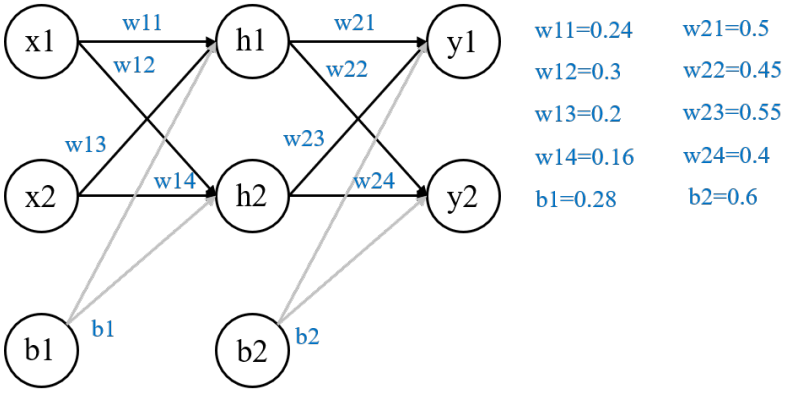

Fig. 1: 多层感知机模型

答：设 $sig(x) = \frac{1}{1+e^{-x}}$，根据已知条件可求得：

$$
\begin{aligned}
h_1 &= x_1 \cdot w_{11} + x_2 \cdot w_{13} + b_1 \\
&= 0.08 \times 0.24 + 0.12 \times 0.2 + 0.28 \\
&= 0.0192 + 0.024 + 0.28 \\
&= 0.3232 \\
sig(h_1) &= \frac{1}{1+e^{-h_1}} \approx 0.58 \\
h_2 &= x_1 \cdot w_{12} + x_2 \cdot w_{14} + b_1 \\
&= 0.08 \times 0.3 + 0.12 \times 0.16 + 0.28 \\
&= 0.024 + 0.0192 + 0.28 \\
&= 0.3232 \\
sig(h_2) &= \frac{1}{1+e^{-h_2}} \approx 0.58 \\
y_1 &= sig(h_1) \cdot w_{21} + sig(h_2) \cdot w_{23} + b_2 \\
&= 0.58 \times 0.5 + 0.58 \times 0.55 + 0.6 \\
&= 0.29 + 0.319 + 0.6 \\
&= 1.209
\end{aligned}
$$

$$\begin{aligned} y_2 &= sig(h_1) \cdot w_{22} + sig(h_2) \cdot w_{24} + b_2 \\ &= 0.58 \times 0.45 + 0.58 \times 0.4 + 0.6 \\ &= 0.493 + 0.6 \\ &= 1.093 \end{aligned}$$

又因为

$$sig'(x) = \frac{e^{-x}}{(1+e^{-x})^2}$$

可求得

$$sig'(h_1) = \frac{e^{-h_1}}{(1+e^{-h_1})^2} \approx 0.2435$$
$$sig'(h_2) = \frac{e^{-h_2}}{(1+e^{-h_2})^2} \approx 0.2435$$

则Loss为

$$Loss_1 = \frac{1}{2}(y_1 - Label_1)^2 = \frac{1}{2}(y_1^2 - 2y_1 Label_1 + Label_1^2)$$
$$Loss_2 = \frac{1}{2}(y_2 - Label_2)^2 = \frac{1}{2}(y_2^2 - 2y_2 Label_2 + Label_2^2)$$
$$\text{Loss}s = Loss_1 + Loss_2$$

因此可用Loss对每一个参数求偏导：

$$\frac{\partial Loss_1}{\partial y_1} = y_1 - Label_1$$
$$\frac{\partial Loss_2}{\partial y_2} = y_2 - Label_2$$
$$y_1 - Label_1 = 1.209 - 0.05 = 1.159$$
$$y_2 - Label_2 = 1.093 - 0.95 = 0.143$$
$$\begin{aligned} \frac{\partial Loss_1}{\partial w_{21}} &= \frac{\partial Loss_1}{\partial y_1} \cdot \frac{\partial y_1}{\partial w_{21}} \\ &= (y_1 - Label_1) \cdot sig(h_1) \\ &= 1.159 \times 0.58 \\ &= 0.672 \end{aligned}$$
$$\begin{aligned} \frac{\partial Loss_1}{\partial w_{23}} &= \frac{\partial Loss_1}{\partial y_1} \cdot \frac{\partial y_1}{\partial w_{23}} \\ &= (y_1 - Label_1) \cdot sig(h_2) \\ &= 1.159 \times 0.58 \\ &= 0.672 \end{aligned}$$

$$\begin{aligned}\frac{\partial Loss_2}{\partial w_{22}} &= \frac{\partial Loss_2}{\partial y_2} \cdot \frac{\partial y_2}{\partial w_{22}} \\ &= (y_2 - Label_2) \cdot sig(h_1) \\ &= 0.143 \times 0.58 \\ &= 0.083\end{aligned}$$

$$\begin{aligned}\frac{\partial Loss_2}{\partial w_{24}} &= \frac{\partial Loss_2}{\partial y_2} \cdot \frac{\partial y_2}{\partial w_{24}} \\ &= (y_2 - Label_2) \cdot sig(h_1) \\ &= 0.143 \times 0.58 \\ &= 0.083\end{aligned}$$

$$\begin{aligned}\frac{\partial Loss}{\partial b_2} &= \frac{\partial Loss_1}{\partial b_2} + \frac{\partial Loss_2}{\partial b_2} \\ &= \frac{\partial Loss_1}{\partial y_1} \cdot \frac{\partial y_1}{\partial b_2} + \frac{\partial Loss_2}{\partial y_2} \cdot \frac{\partial y_2}{\partial 2_2} \\ &= (y_1 - Label_1) + (y_2 - Label_2) \\ &= 1.159 + 0.143 \\ &= 1.302\end{aligned}$$

$$\begin{aligned}&\frac{\partial Loss}{\partial w_{11}} = \frac{\partial Loss_1}{\partial w_{11}} + \frac{\partial Loss_2}{\partial w_{11}} \\ &= \frac{\partial Loss_1}{\partial y_1} \cdot \frac{\partial y_1}{\partial sig(h_1)} \cdot \frac{\mathrm{d}sig(h_1)}{\mathrm{d}x} \cdot \frac{\partial h_1}{\partial w_{11}} + \frac{\partial Loss_2}{\partial y_2} \cdot \frac{\partial y_2}{\partial sig(h_1)} \cdot \frac{\mathrm{d}sig(h_1)}{\mathrm{d}x} \cdot \frac{\partial h_1}{\partial w_{11}} \\ &= (y_1 - Label_1) \cdot w_{21} \cdot sig^{'}(h_1) \cdot x_1 + (y_2 - Label_2) \cdot w_{22} \cdot sig^{'}(h_1) \cdot x_1 \\ &= [(y_1 - Label_1) \cdot w_{21} + (y_2 - Label_2) \cdot w_{22}] \cdot sig^{'}(h_1) \cdot x_1 \\ &= (1.159 \times 0.5 + 0.143 \times 0.45) \times 0.2435 \times 0.08 \\ &= 0.0125\end{aligned}$$

$$\begin{aligned}\frac{\partial Loss}{\partial w_{12}} &= \frac{\partial Loss_1}{\partial w_{12}} + \frac{\partial Loss_2}{\partial w_{12}} \\ &= \frac{\partial Loss_1}{\partial y_1} \cdot \frac{\partial y_1}{\partial sig(h_2)} \cdot sig^{'}(h_2) \cdot \frac{\partial h_2}{\partial w_{12}} + \frac{\partial Loss_2}{\partial y_2} \cdot \frac{\partial y_2}{\partial sig(h_2)} \cdot sig^{'}(h_2) \cdot \frac{\partial h_2}{\partial w_{12}} \\ &= [(y_1 - Label_1) \cdot w_{23} + (y_2 - Label_2) \cdot w_{24}] \cdot sig^{'}(h_2) \cdot x_1 \\ &= (1.159 \times 0.55 + 0.143 \times 0.4) \times 0.2435 \times 0.08 \\ &= 0.0135\end{aligned}$$

$$\begin{aligned}\frac{\partial Loss}{\partial w_{13}} &= \frac{\partial Loss_1}{\partial w_{13}} + \frac{\partial Loss_2}{\partial w_{13}} \\ &= \frac{\partial Loss_1}{\partial y_1} \cdot \frac{\partial y_1}{\partial sig(h_1)} \cdot sig^{'}(h_1) \cdot \frac{\partial h_1}{\partial w_{13}} + \frac{\partial Loss_2}{\partial y_2} \cdot \frac{\partial y_2}{\partial sig(h_1)} \cdot sig^{'}(h_1) \cdot \frac{\partial h_1}{\partial w_{13}} \\ &= [(y_1 - Label_1) \cdot w_{21} + (y_2 - Label_2) \cdot w_{22}] \cdot sig'(h_1) \cdot x_2 \\ &= (1.159 \times 0.5 + 0.143 \times 0.45) \times 0.2435 \times 0.12 \\ &= 0.0188\end{aligned}$$

$$\begin{aligned}\frac{\partial Loss}{\partial w_{14}} &= \frac{\partial Loss_1}{\partial w_{14}} + \frac{\partial Loss_2}{\partial w_{14}} \\ &= \frac{\partial Loss_1}{\partial y_1} \cdot \frac{\partial y_1}{\partial sig(h_2)} \cdot sig^{'}(h_2) \cdot \frac{\partial h_2}{\partial w_{14}} + \frac{\partial Loss_2}{\partial y_2} \cdot \frac{\partial y_2}{\partial sig(h_2)} \cdot sig^{'}(h_2) \cdot \frac{\partial h_2}{\partial w_{14}} \\ &= [(y_1 - Label_1) \cdot w_{23} + (y_2 - Label_2) \cdot w_{24}] \cdot sig^{'}(h_2) \cdot x_2 \\ &= (1.159 \times 0.55 + 0.143 \times 0.4) \times 0.2435 \times 0.12 \\ &= 0.0202\end{aligned}$$

$$\begin{aligned}\frac{\partial Loss}{\partial b_1} &= \frac{\partial Loss_1}{\partial b_1} + \frac{\partial Loss_2}{\partial b_1} \\ &= \frac{\partial Loss_1}{\partial y_1} \cdot [\frac{\partial y_1}{\partial sig(h_1)} \cdot sig^{'}(h_1) \cdot \frac{\partial h_1}{\partial b_1} + \frac{\partial y_1}{\partial sig(h_2)} \cdot sig^{'}(h_2) \cdot \frac{\partial h_2}{\partial b_1}] \\ &+ \frac{\partial Loss_2}{\partial y_2} \cdot [\frac{\partial y_2}{\partial sig(h_1)} \cdot sig^{'}(h_1) \cdot \frac{\partial h_1}{\partial b_1} + \frac{\partial y_2}{\partial sig(h_2)} \cdot sig^{'}(h_2) \cdot \frac{\partial h_2}{\partial b_1}] \\ &= (y_1 - Label_1) \cdot [w_{21} \cdot sig^{'}(h_1) + w_{23} \cdot sig^{'}(h_2)] + (y_2 - Label_2) \cdot [w_{22} \cdot sig^{'}(h_1) + w_{24} \cdot sig^{'}(h_2)] \\ &= 1.159 \times (0.5 \times 0.2435 + 0.55 \times 0.2435) + 0.143 \times (0.45 \times 0.2435 + 0.4 \times 0.2435) \\ &= 0.326\end{aligned}$$

根据公式

$$w = w - \alpha \cdot \frac{\partial Loss}{\partial w}$$

可得

$$\begin{aligned}w_{21}^{+} &= w_{21} - \alpha \cdot \frac{\partial Loss_1}{\partial w_{21}} = 0.5 - 0.5 \times 0.672 = 0.164 \\ w_{23}^{+} &= w_{23} - \alpha \cdot \frac{\partial Loss_1}{\partial w_{23}} = 0.55 - 0.5 \times 0.672 = 0.214 \\ w_{22}^{+} &= w_{22} - \alpha \cdot \frac{\partial Loss_2}{\partial w_{22}} = 0.45 - 0.5 \times 0.083 = 0.4085 \\ w_{24}^{+} &= w_{24} - \alpha \cdot \frac{\partial Loss_2}{\partial w_{24}} = 0.4 - 0.5 \times 0.083 = 0.3585 \\ b_2^{+} &= b_2 - \alpha \cdot \frac{\partial Loss}{\partial b_2} = 0.6 - 0.5 \times 1.302 = -0.051 \\ w_{11}^{+} &= w_{11} - \alpha \cdot \frac{\partial Loss}{\partial w_{11}} = 0.24 - 0.5 \times 0.0125 = 0.23375 \\ w_{12}^{+} &= w_{12} - \alpha \cdot \frac{\partial Loss}{\partial w_{12}} = 0.3 - 0.5 \times 0.0135 = 0.29325 \\ w_{13}^{+} &= w_{13} - \alpha \cdot \frac{\partial Loss}{\partial w_{13}} = 0.2 - 0.5 \times 0.0188 = 0.1906 \\ w_{14}^{+} &= w_{14} - \alpha \cdot \frac{\partial Loss}{\partial w_{14}} = 0.16 - 0.5 \times 0.0202 = 0.1499 \\ b_1^{+} &= b_1 - \alpha \cdot \frac{\partial Loss}{\partial b_1} = 0.28 - 0.5 \times 0.326 = 0.117\end{aligned}$$

综上所述，更新后的参数为：

$$\begin{array}{ccccc} w_{11} = 0.2338 & w_{12} = 0.2933 & w_{13} = 0.1906 & w_{14} = 0.1499 & b_1 = 0.117 \\ w_{21} = 0.164 & w_{22} = 0.4085 & w_{23} = 0.214 & w_{24} = 0.3585 & b_2 = -0.051 \end{array}$$