

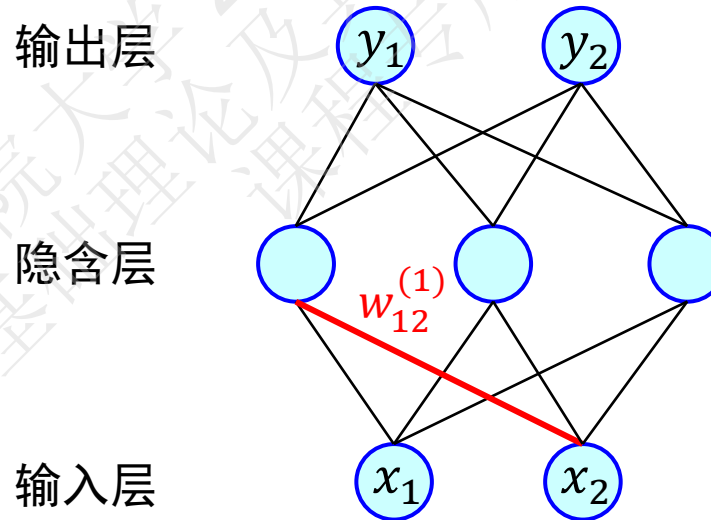
# 小作业4



1. 如果用 BP 算法对下图所示的神经网络（2 输入 2 输出）进行训练：（4分）

- 1) 试推导图中权重  $w_{12}^{(1)}$  (红色连接) 的更新公式；
- 2) 试回答该网络中有多少参数（包括权重和偏置）需要确定。

假设：只有一个样本  $(x, y)$  用于训练，损失函数为  $E = \frac{1}{2}((\hat{y}_1 - y_1)^2 + (\hat{y}_2 - y_2)^2)$ ，所有神经元均采用 sigmoid 函数作为激活函数。



# 小作业4

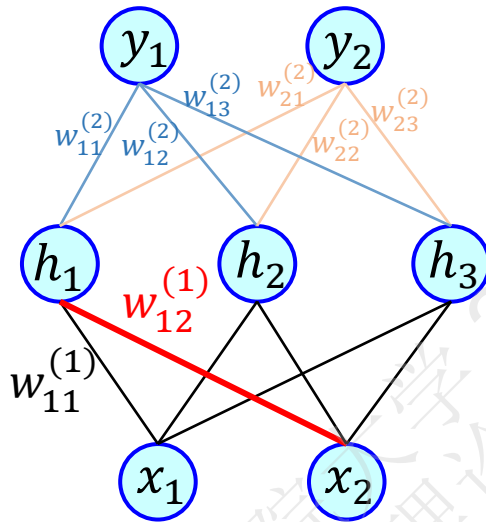


1) 设 $f$ 为sigmoid 函数, 则 $f'(x) = f(x)[1 - f(x)]$

输出层

隐含层

输入层



$$E = \frac{1}{2}[(\hat{y}_1 - y_1)^2 + (\hat{y}_2 - y_2)^2]$$

$$y_1 = f(z_1^{(3)}), y_2 = f(z_2^{(3)})$$

$$z_1^{(3)} = w_{11}^{(2)}h_1 + w_{12}^{(2)}h_2 + w_{13}^{(2)}h_3$$

$$z_2^{(3)} = w_{21}^{(2)}h_1 + w_{22}^{(2)}h_2 + w_{23}^{(2)}h_3$$

$$h_1 = f(z_1^{(2)})$$

$$z_1^{(2)} = w_{11}^{(1)}x_1 + w_{12}^{(1)}x_2$$

链式法则

$$w_{12}^{(1)} \rightarrow z_1^{(2)} \rightarrow h_1 \rightarrow \begin{cases} z_1^{(3)} \rightarrow y_1 \rightarrow E \\ z_2^{(3)} \rightarrow y_2 \end{cases}$$

# 小作业4



$$w_{12}^{(1)} \rightarrow w_{12}^{(1)} + \Delta w_{12}^{(1)}$$

$$\text{链式法则 } w_{12}^{(1)} \rightarrow z_1^{(2)} \rightarrow h_1 \rightarrow \begin{cases} z_1^{(3)} \rightarrow y_1 \\ z_2^{(3)} \rightarrow y_2 \end{cases} \rightarrow E$$

$$\Delta w_{12}^{(1)} = -\eta \frac{\partial E}{\partial w_{12}^{(1)}} = -\eta \frac{\partial E}{\partial h_1} \frac{\partial h_1}{\partial z_1^{(2)}} \frac{\partial z_1^{(2)}}{\partial w_{12}^{(1)}} \quad E = \frac{1}{2} [(\hat{y}_1 - y_1)^2 + (\hat{y}_2 - y_2)^2]$$

$$= -\eta \left[ \frac{\partial E}{\partial y_1} \frac{\partial y_1}{\partial z_1^{(3)}} \frac{\partial z_1^{(3)}}{\partial h_1} + \frac{\partial E}{\partial y_2} \frac{\partial y_2}{\partial z_2^{(3)}} \frac{\partial z_2^{(3)}}{\partial h_1} \right] \frac{\partial h_1}{\partial z_1^{(2)}} \frac{\partial z_1^{(2)}}{\partial w_{12}^{(1)}}$$

$$= \eta \left[ (\hat{y}_1 - y_1) f' \left( z_1^{(3)} \right) w_{11}^{(2)} + (\hat{y}_2 - y_2) f' \left( z_2^{(3)} \right) w_{21}^{(2)} \right] f' \left( z_1^{(2)} \right) x_2$$

$$= \eta \left[ \delta_1^{(3)} w_{11}^{(2)} + \delta_2^{(3)} w_{21}^{(2)} \right] f' \left( z_1^{(2)} \right) x_2$$

$$= \eta \delta_1^{(2)} x_2$$

2) 该网络需确定的参数共有： $(2+1) \times 3 + (3+1) \times 2 = 17$ 个

其中权重参数12个，偏置参数5个。

# 小作业4



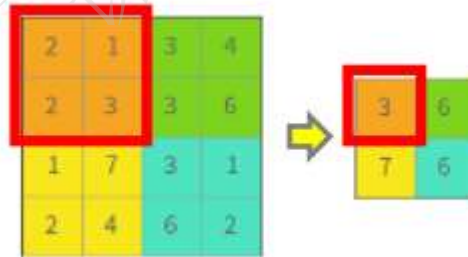
2. 有一个  $44 \times 44 \times 16$  的输入，用10个  $5 \times 5 \times 16$  的卷积核进行卷积，步长为1，无填充，输出是多少？考虑偏置，学习的参数为多少个？然后再进行步长为2，采样窗口为  $2 \times 2$  的最大池化，输出又是多少？（3分）

$44 - 5 + 1 = 40$ ，卷积后输出： $40 \times 40 \times 10$ （10个卷积核对应10个特征层）

学习参数个数： $(5 \times 5 \times 16 + 1) \times 10 = 4010$ 个（卷积核上的权重4000个，偏置参数10个）

$40 / 2 = 20$ ，池化后输出： $20 \times 20 \times 10$

例：



3. 两个任务：机器翻译、从一张合影照片中找到特定的人。CNN和RNN，请选择分别用何种神经网络完成更适合？为什么？（2分）

机器翻译更适合用RNN。因为RNN的主要用途是处理和预测序列信息，而在翻译中，前一个单词对于后一个单词的预测有很大影响。

第二个任务更适合用CNN。CNN的局部连接和参数共享机制可以降低图像处理中参数数目，节省训练开销。且不同的卷积核可以提取图片中不同的特征。