

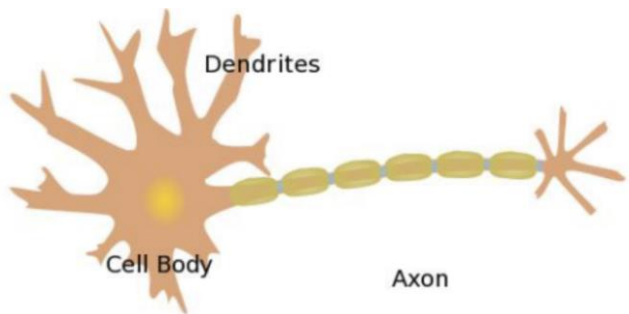


机器学习工程师直通车

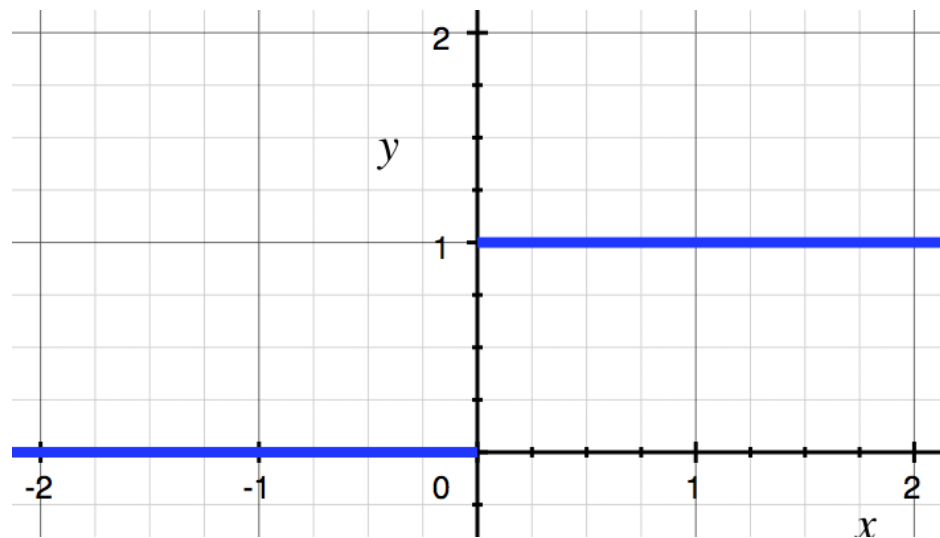
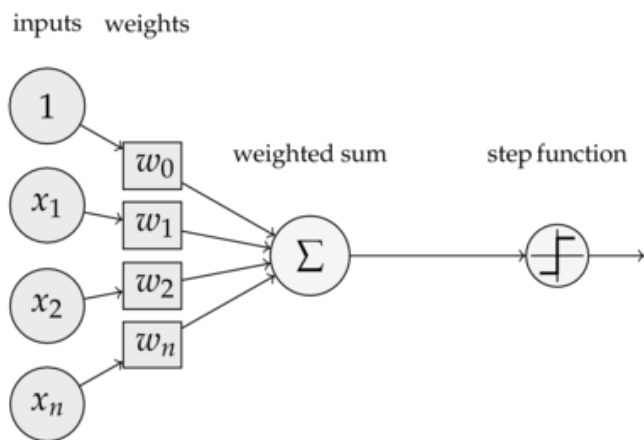
—— 深度学习部分

讲师：智亮

1. 感知器的前向计算
2. 神经网络的前向计算
3. 损失函数与梯度
4. 神经网络的训练
5. 神经网络的反向传播



$$y = f(w_1x_1 + w_2x_2 + w_3x_3 + bias)$$
$$= f(w_1x_1 + w_2x_2 + w_3x_3 + w_0x_0), x_0 = 1$$
$$f(x) = \begin{cases} 1 & x > 0 \\ 0 & x \leq 0 \end{cases}$$



dot product:

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \cdot \begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{bmatrix} = \begin{bmatrix} 1 * 0.1 + 2 * 0.4 & 1 * 0.2 + 2 * 0.5 & 1 * 0.3 + 2 * 0.6 \\ 3 * 0.1 + 4 * 0.4 & 3 * 0.2 + 4 * 0.5 & 3 * 0.3 + 4 * 0.6 \\ 5 * 0.1 + 6 * 0.2 & 5 * 0.2 + 6 * 0.5 & 5 * 0.3 + 6 * 0.6 \end{bmatrix} = \begin{bmatrix} 0.9 & 1.2 & 1.5 \\ 1.9 & 2.6 & 3.3 \\ 2.9 & 4.0 & 5.1 \end{bmatrix}$$

$$\mathbb{R}[m, n] \cdot \mathbb{R}[n, p] \rightarrow \mathbb{R}[m, p]$$

element-wise product / Hadamard product :

$$\begin{bmatrix} 1 & 3 & 2 \\ 1 & 0 & 0 \\ 1 & 2 & 2 \end{bmatrix} \odot \begin{bmatrix} 0 & 0 & 2 \\ 7 & 5 & 0 \\ 2 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 * 0 & 3 * 0 & 2 * 2 \\ 1 * 7 & 0 * 5 & 0 * 0 \\ 1 * 2 & 2 * 1 & 2 * 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 4 \\ 7 & 0 & 0 \\ 2 & 2 & 2 \end{bmatrix}$$

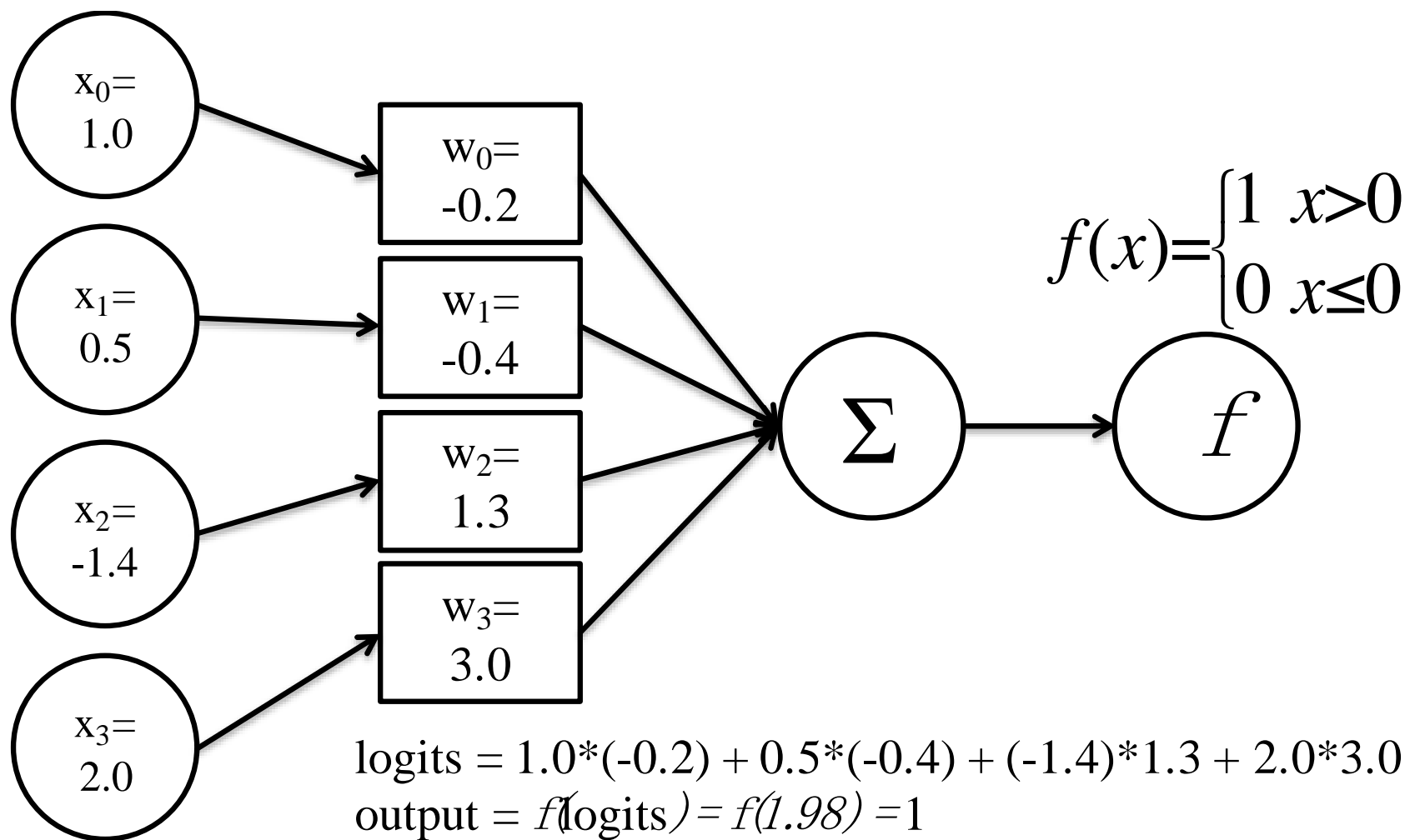
$$\text{logit} = w_0 x_0 + w_1 x_1 + w_2 x_2 + w_3 x_3 + \dots + w_n x_n$$

$$w_0 = b \text{ (bias, 偏置)}, \quad x_0 = 1$$

$$w = [w_0, w_1, w_2, \dots, w_n], \quad x = [x_0, x_1, x_2, \dots, x_n]$$

$$\text{则 } \text{logit} = w \cdot x$$

$$\text{output} = f(\text{logit}),$$



- 向量化

例如

$$x_1 = [-1.0, 3.0, 2.0] \quad w = [4.0, -3.0, 5.0]$$

$$x_2 = [2.0, -1.0, 5.0] \quad b = 2.0$$

$$x_3 = [-2.0, 0.0, 3.0]$$

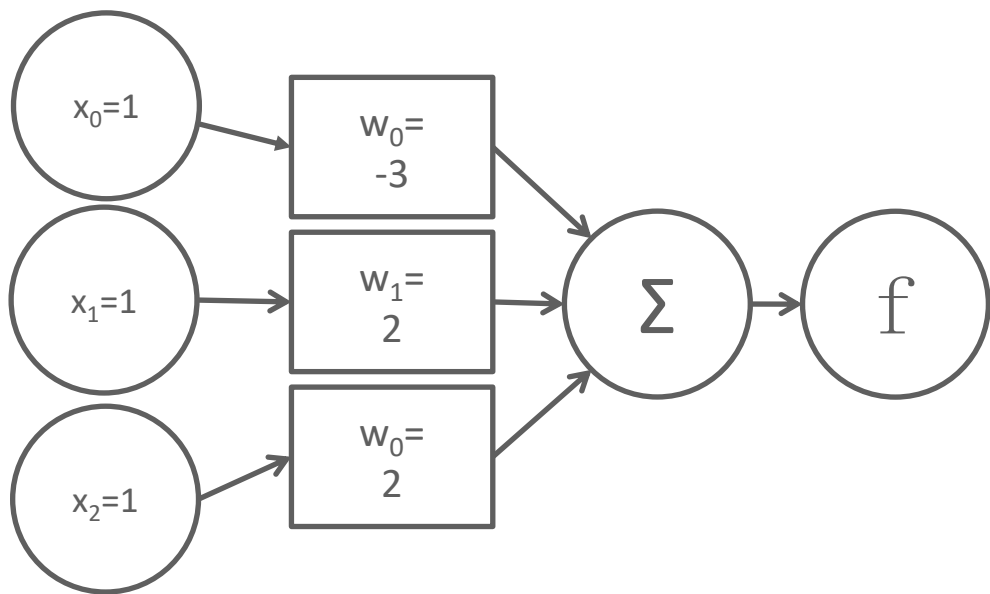
$$x_4 = [4.0, 1.0, 6.0]$$

- 则 $X = \begin{bmatrix} -1.0 & 3.0 & 2.0 \\ 2.0 & -1.0 & 5.0 \\ -2.0 & 0.0 & 3.0 \\ 4.0 & 1.0 & 6.0 \end{bmatrix}$ $logits = \begin{bmatrix} -1.0 & 3.0 & 2.0 \\ 2.0 & -1.0 & 5.0 \\ -2.0 & 0.0 & 3.0 \\ 4.0 & 1.0 & 6.0 \end{bmatrix} \cdot \begin{bmatrix} 4.0 \\ -3.0 \\ 5.0 \end{bmatrix} + 2.0$
 $= [-1.0 \quad 38.0 \quad 9.0 \quad 45.0]^T$

- 则 $output = f(x) = [0 \quad 1 \quad 1 \quad 1]$

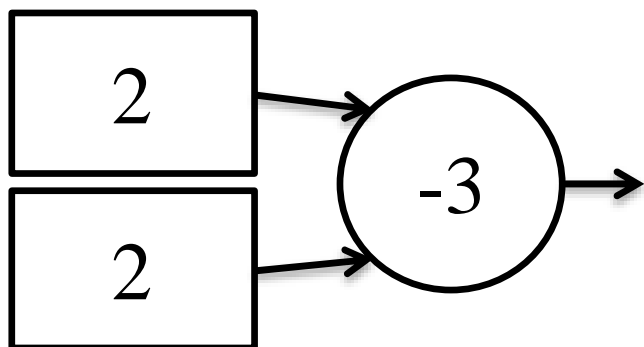
- 使用感知器可以完成一些基础逻辑操作

- 例如：逻辑与

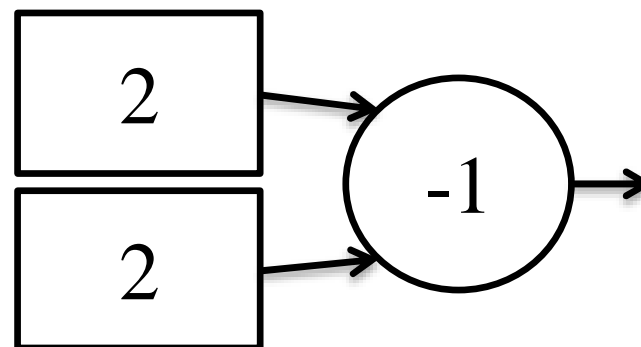


x1	x2	output
1	1	1
1	0	0
0	1	0
0	0	0

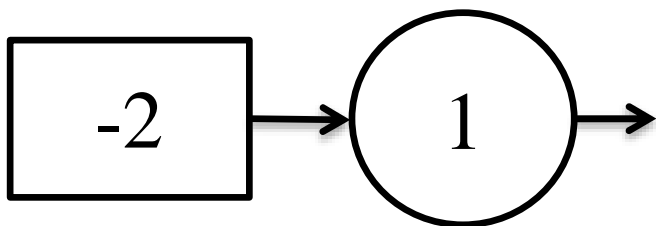
逻辑与



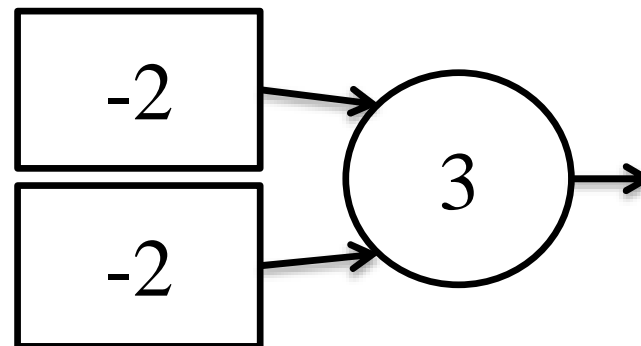
逻辑或

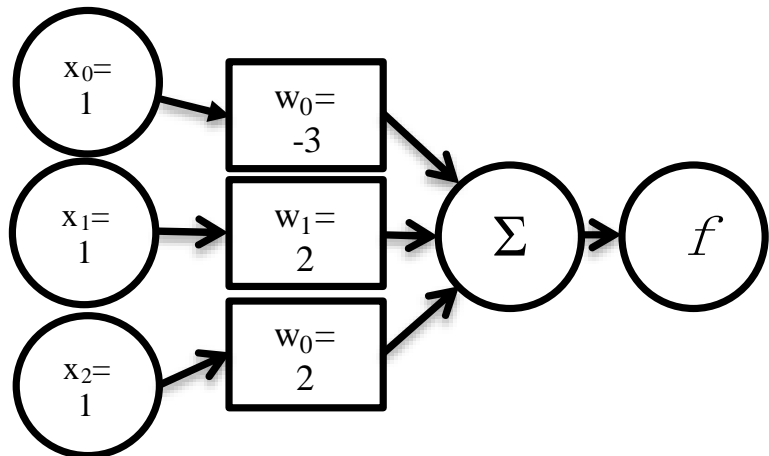


逻辑非



逻辑与非





x1	x2	output
1	1	1
1	0	0
0	1	0
0	0	0

$$w_1x_1 + w_2x_2 + b = 0$$

特征 (feature) 真值 (ground truth)

--- $x_1 + x_2 - 1 = 0$

— $2x_1 + 2x_2 - 3 = 0$

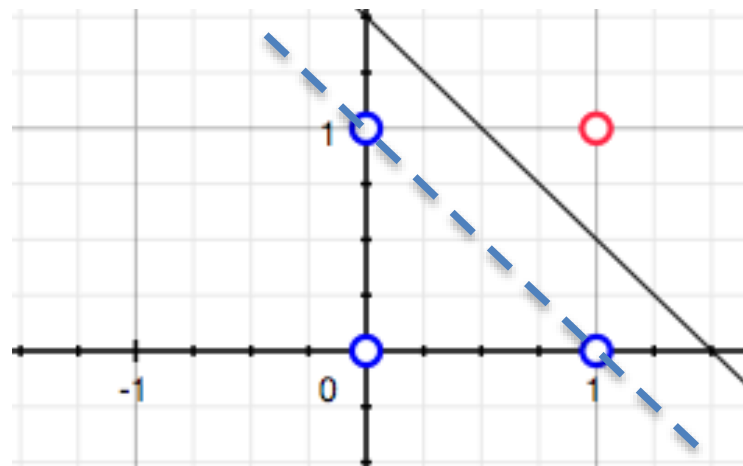
通过真值表求解

可能的一些解:

$$\begin{cases} 1 \times w_1 + 1 \times w_2 + b > 0 \\ 1 \times w_1 + 0 \times w_2 + b \leq 0 \\ 0 \times w_1 + 1 \times w_2 + b \leq 0 \\ 0 \times w_1 + 0 \times w_2 + b \leq 0 \end{cases}$$



$$\begin{cases} w_1 = 2 \\ w_2 = 2 \\ b = -3 \end{cases} \text{ 或 } \begin{cases} w_1 = 1 \\ w_2 = 1 \\ b = -1 \end{cases}$$



$$X = \begin{bmatrix} 1 & 1 \\ -1 & 1 \\ -1 & -0.5 \\ -1 & 1.5 \end{bmatrix}$$

$$Y = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

$$\begin{cases} w_1 + w_2 + b \leq 0 \\ -w_1 + w_2 + b > 0 \\ -w_1 - 0.5w_2 + b \leq 0 \\ -w_1 + 1.5w_2 + b > 0 \end{cases}$$

直接进行数值求解
一组可能的解：

$$\begin{cases} w_1 = -1 \\ w_2 = 4 \\ b = -3 \end{cases}$$

