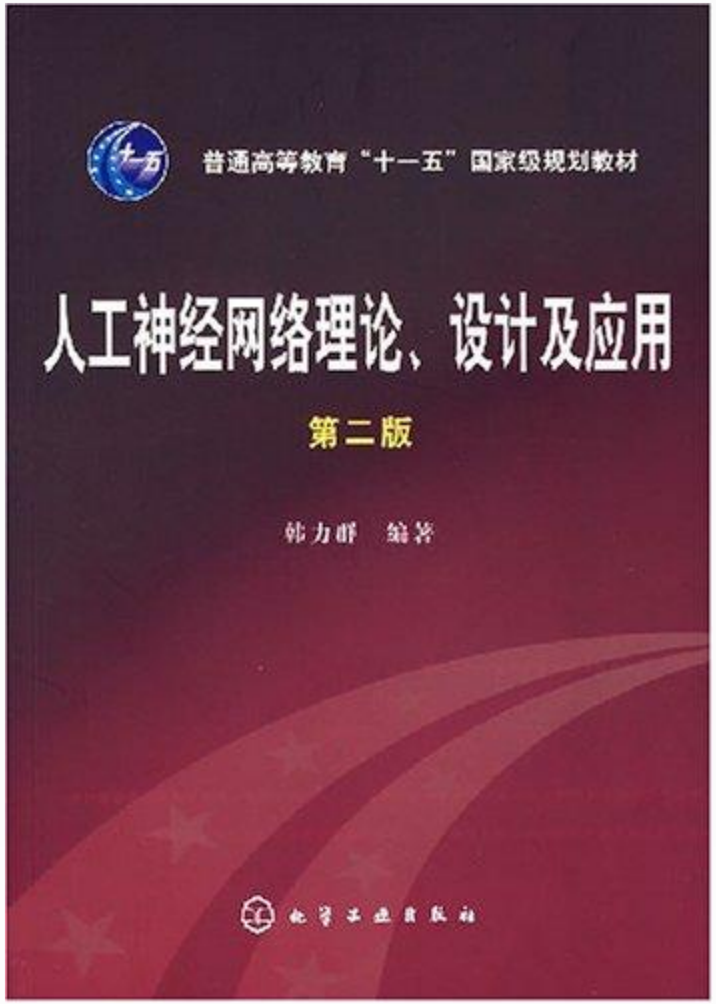
第三课——从0到1-单层感知器

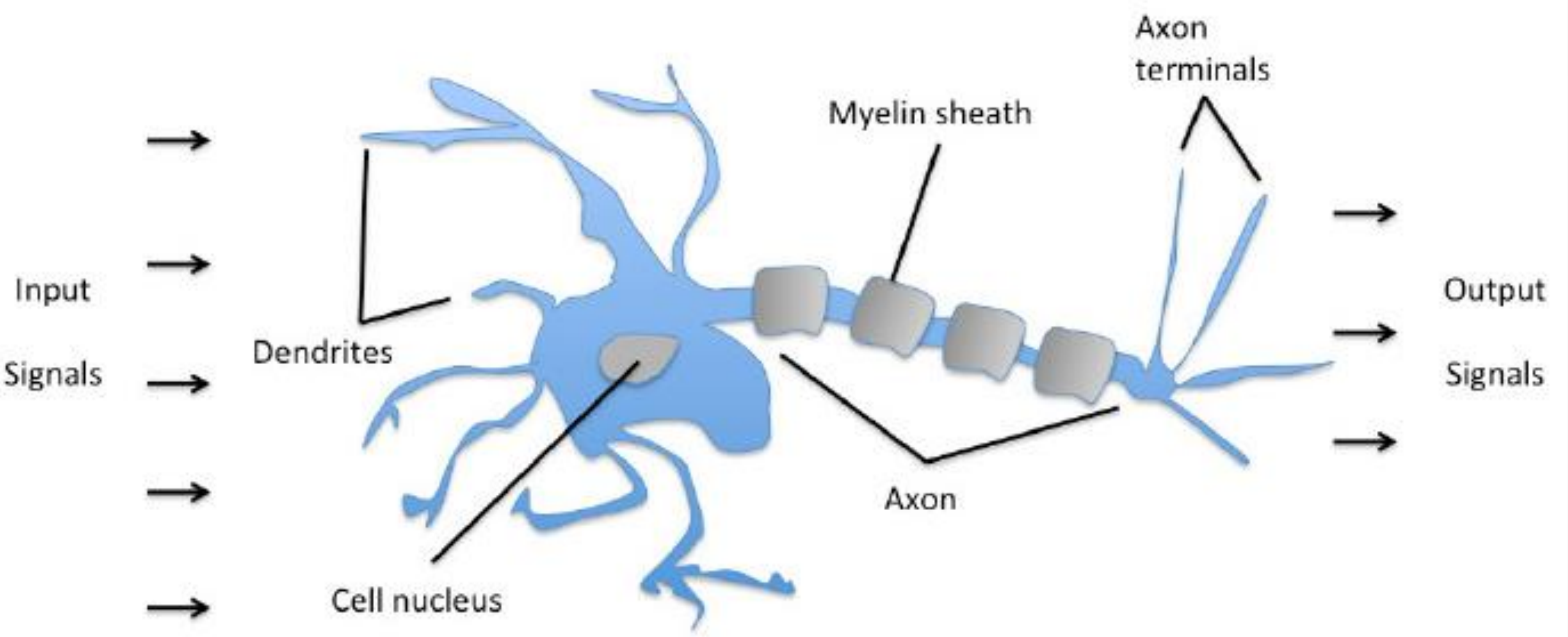
51CTO学院

edu.51cto.com

参考书



神经元



Schematic of a biological neuron.

单层感知器

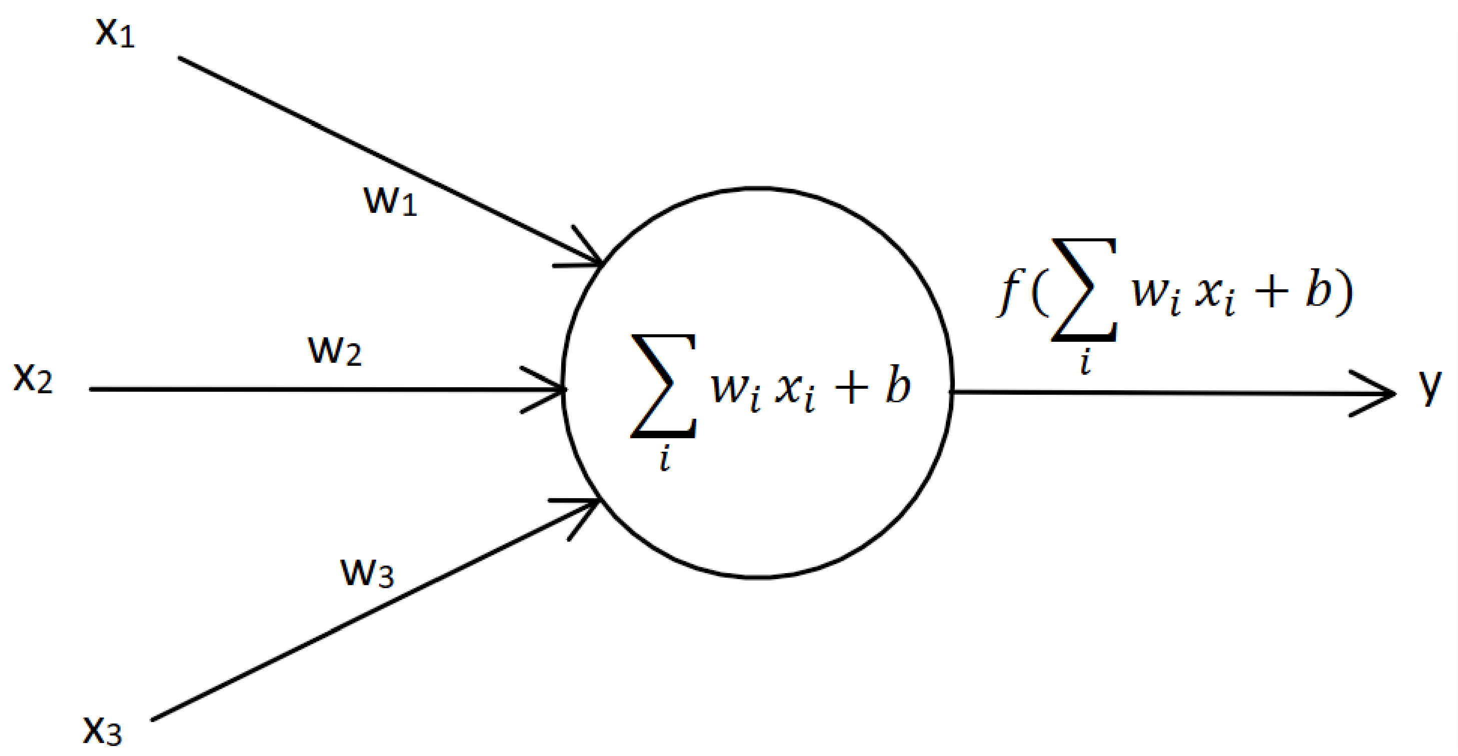
输入节点：x_1,x_2,x_3

输出节点：y

权向量：w_1,w_2,w_3

偏置因子：b

激活函数：$sign(x)=\begin{cases} 1 & X>0 \\ 0 & X=0 \\ -1 & X<0 \end{cases}$

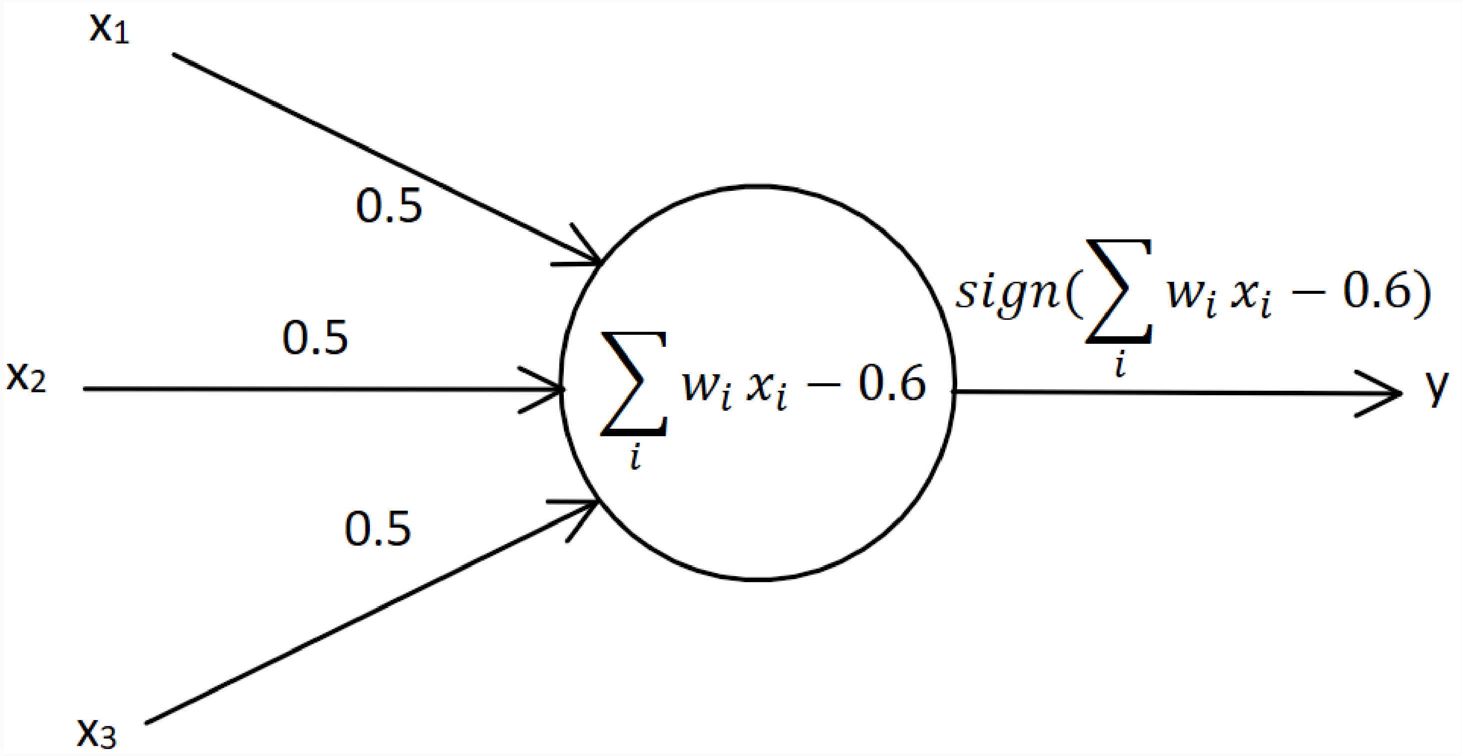


单层感知器举例

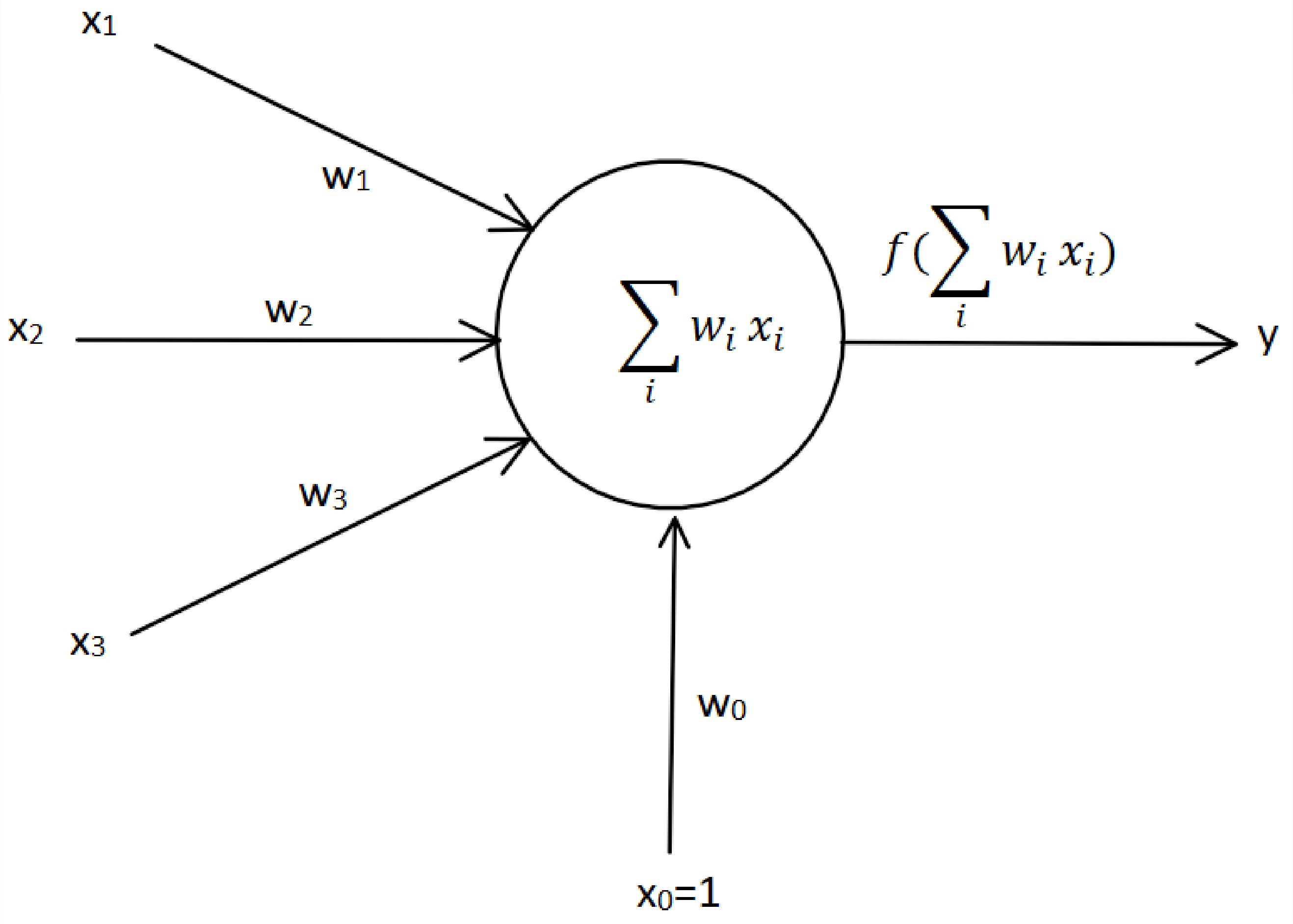
b=-0.6

x_1	x_2	x_3	Y
0	0	0	-1
0	0	1	-1
0	1	0	-1
0	1	1	1
1	0	0	-1
1	0	1	1
1	1	0	1
1	1	1	1

$$y=\begin{cases} 1 & (0.5x_1+0.5x_2+0.5x_3-0.6>0) \\ -1 & (0.5x_1+0.5x_2+0.5x_3-0.6<0) \end{cases}$$



把偏置当做特殊权值



感知器学习规则

1958 年，美国学者 Frank Rosenblatt 首次定义了一个具有单层计算单元的神经网络结构，称为 Perceptron（感知器）。感知器的学习规则规定，学习信号等于神经元期望输出（教师信号）与实际输出之差：

$$r = d_j - o_j \tag{2.17}$$

式中，d_j 为期望的输出，$o_j = f(\boldsymbol{W}_j{}^{\mathrm{T}}\boldsymbol{X})$。感知器采用了符号函数作为转移函数，其表达为：

$$f(\boldsymbol{W}_j^{\mathrm{T}}\boldsymbol{X}) = \mathrm{sgn}(\boldsymbol{W}_j^{\mathrm{T}}\boldsymbol{X}) = \begin{cases} 1, & \boldsymbol{W}_j^{\mathrm{T}}\boldsymbol{X} \geqslant 0 \\ -1, & \boldsymbol{W}_j^{\mathrm{T}}\boldsymbol{X} < 0 \end{cases} \tag{2.18}$$

因此，权值调整公式应为：

$$\Delta \boldsymbol{W}_j = \eta[d_j - \mathrm{sgn}(\boldsymbol{W}_j^{\mathrm{T}}\boldsymbol{X})]\boldsymbol{X} \tag{2.19a}$$

$$\Delta w_{ij} = \eta[d_j - \mathrm{sgn}(\boldsymbol{W}_j^{\mathrm{T}}\boldsymbol{X})]x_i \qquad i = 0, 1, \cdots, n \tag{2.19b}$$

式中，当实际输出与期望值相同时，权值不需要调整；在有误差存在情况下，由于 d_j 和 $\mathrm{sgn}(\boldsymbol{W}_j^{\mathrm{T}}\boldsymbol{X}) \in \{-1, 1\}$，权值调整公式可简化为：

$$\Delta \boldsymbol{W}_j = \pm 2\eta \boldsymbol{X} \tag{2.19c}$$

感知器学习规则只适用于二进制神经元，初始权值可取任意值。

感知器学习规则代表一种有导师学习。由于感知器理论是研究其他神经网络的基础，该规则对于神经网络的有导师学习具有极为重要的意义。

学习率

- η学习率(0<η≤1)。
- 学习率太大，容易造成权值调整不稳定。
- 学习率太小，权值调整太慢，迭代次数太多。

收敛条件

收敛条件通常可以是：

- 误差小于某个预先设定的较小的值。
- 两次迭代之间的权值变化已经很小。
- 设定最大迭代次数，当迭代超过最大次数就停止。

微信公众号：深度学习与神经网络

QQ群：616043628

51CTO学院

Thank You！

为梦想增值！