

# 多层感知器与反向传播算法

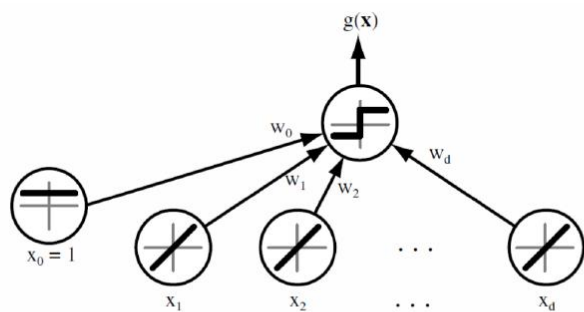
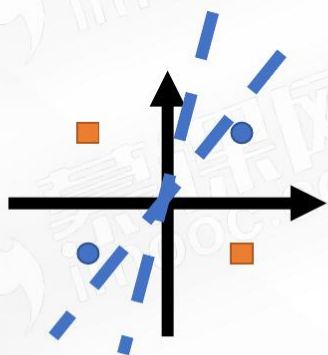
## 目录

- ◆ 多层感知器
- ◆ 反向传播算法

## 多层感知器

### 单层感知器的缺陷

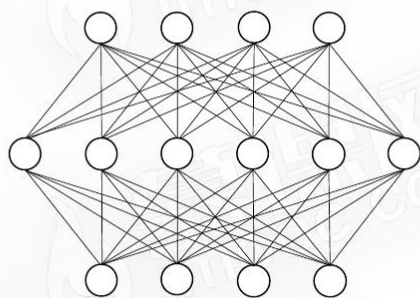
- ◆ 马文·明斯基在1969年出版《感知器》书中证明单层感知器无法解决异或问题



Duda et al. Pattern Classification 2000

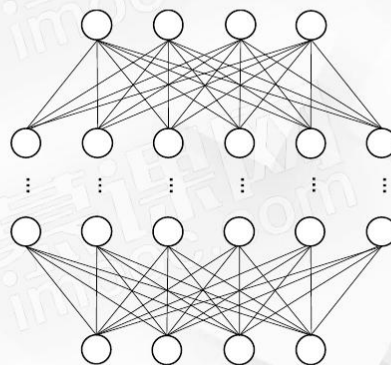
# 多层感知器Multi Layer Perceptron

- ◆ 简称MLP：在单层神经网络基础上引入一个或多个隐藏层，使网络有多个网络层，被称为多层感知器，或前馈神经网络。



单隐层前馈网络

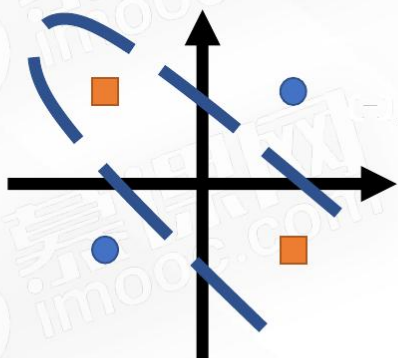
非线性激活函数



多隐层前馈网络

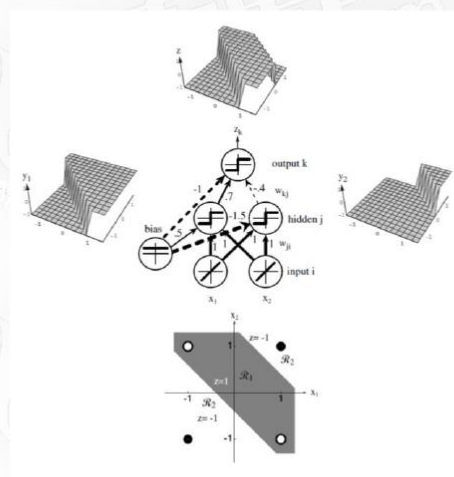
## 多层感知器表达能力

- ◆ 多层感知机解决异或问题



线性分界面：  
非线性分界面：

+ = +



Duda et al. Pattern Classification 2000

# 多层感知器表达能力

## ◆ 万能逼近定理

- George Cybenko在1989年首次严格给出并论证了用以Sigmoid函数作为激活函数的神经网络的逼近能力：只要有一个隐藏层，前馈型神经网络就能一致逼近至任意连续函数。
- Cybenko G. Approximation by superpositions of a sigmoidal function[J]. Mathematics of control, signals and systems, 1989, 2(4): 303-314.
- Hornik7在1991年证明多层前馈神经网络的结构本身给神经网络提供了泛逼近性 (Universal Approximators)，即泛逼近定理 (Universal Approximation Theorem)
- K.Hornil, "Multilayer Feedforward Networks Are Universal Approximators," Neural Networks, pp. 359-366, 1989.
- K.Hornil, "Approximation Capabilities of Multilayer Feedforward Networks," Neural Networks, pp. 251-257, 1991.

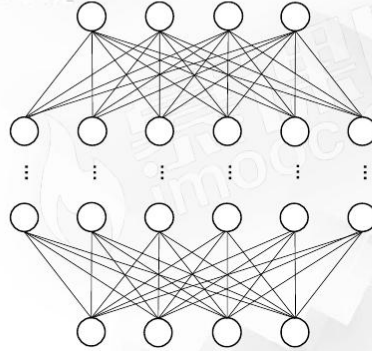
## 反向传播算法

# 什么是反向传播算法

- ◆ 1986年, Rummelhart、McClelland、Hinton等人改进了反向传播算法 (Backpropagation, 缩写为BP) 用于多层神经网络学习

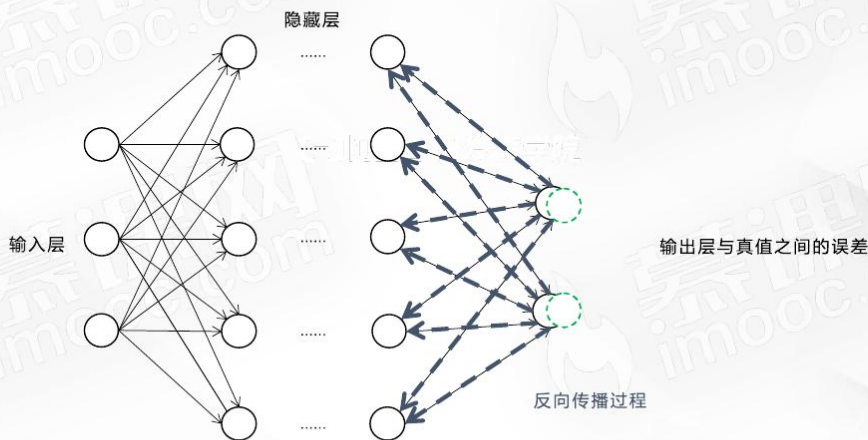
令  $w_{ij}$  表示第  $i$  层的第  $j$  个神经元到第  $i+1$  层的第  $j$  个神经元的连接权值,  $x_i$  表示第  $i$  层第  $j$  个神经元的输入,  $y_i$  表示第  $i$  层第  $j$  个神经元的输出,  $b_i$  表示第  $i$  层第  $j$  个神经元的偏置,  $\sigma$  表示激活函数,  $C$  代表损失函数/代价函数 (Cost Function)

当前神经元的输入是上一层神经元输出的线性组合, 当前神经元的输出经过激活函数映射



# 什么是反向传播算法

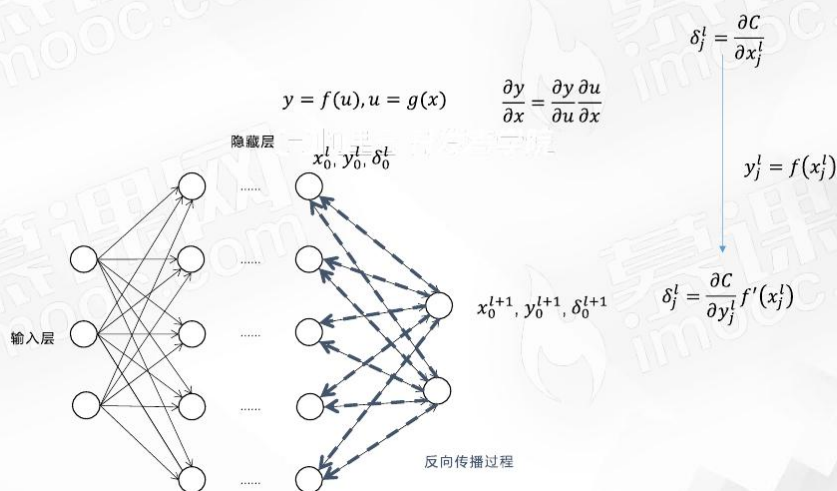
- ◆ 反向传播: 损失函数开始从后向前, 梯度逐步传递至第一层





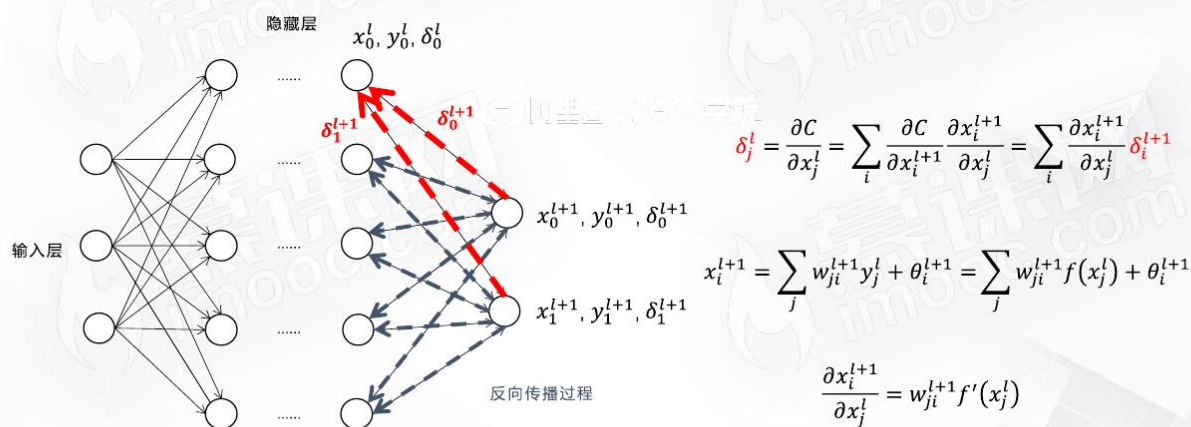
# 误差反向传播算法原理

- ◆ 反向传播中的误差，令  $\delta$  表示第  $l$  层第  $j$  个神经元上的误差



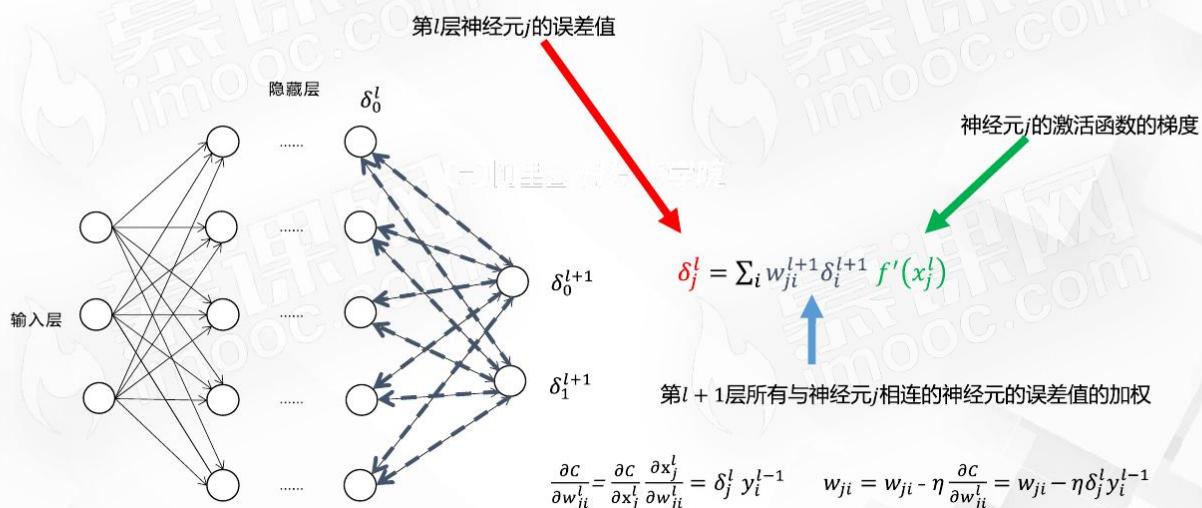
# 误差反向传播算法原理

- ◆ 前向计算时，神经元的输入是上一层神经元输出的线性组合，类似地，反向计算时，可通过下层神经元的误差来计算当前层的误差



# 误差反向传播算法原理

## ◆ 基于反向传播算法的参数更新



下次预告：多层神经网络案例实践