

11 人工神经网络

11 人工神经网络

11.1 本章内容简介

11.2 动物神经系统的基本原理

11.3 人工神经网络的基本思想

11.4 神经元的原理

11.5 sigmoid函数

11.6 神经网络的结构

11.7 神经网络每一层完成的变换

11.8 正向传播算法

11.9 神经网络的本质

11.10 怎样用于实际问题

手写数字识别的例子

11.1 本章内容简介

简写ANN

1536684525091

11.2 动物神经系统的基本原理

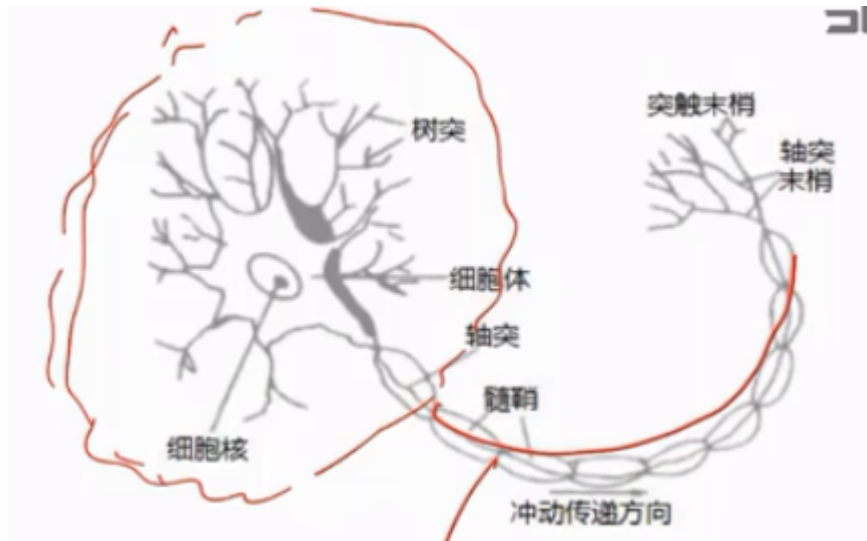
动物的神经元结构

人类的大脑由大约800亿个神经元组成

这些神经元由突触与其他神经元相互连接，交换电信号和化学信号

大脑通过神经元之间的协作完成各种功能

神经元之间的连接关系是通过进化、生长发育和后天刺激形成的



整个大脑最重要的是整个神经元，神经元之间的连接关系。

11.3 人工神经网络的基本思想

人工神经网络
受动物神经系统的启发，是一种仿生的方法
是感知器模型的进一步发展

感知器模型



感知器模型是存在很大的局限性的。

11.4 神经元的原理

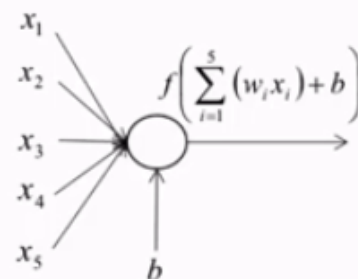
神经元结构

输入值 $(x_1, x_2, x_3, x_4, x_5)$

连接权重 $(w_1, w_2, w_3, w_4, w_5)$

偏置项 b

输出值 $y = f\left(\sum_{i=1}^5 w_i x_i + b\right)$

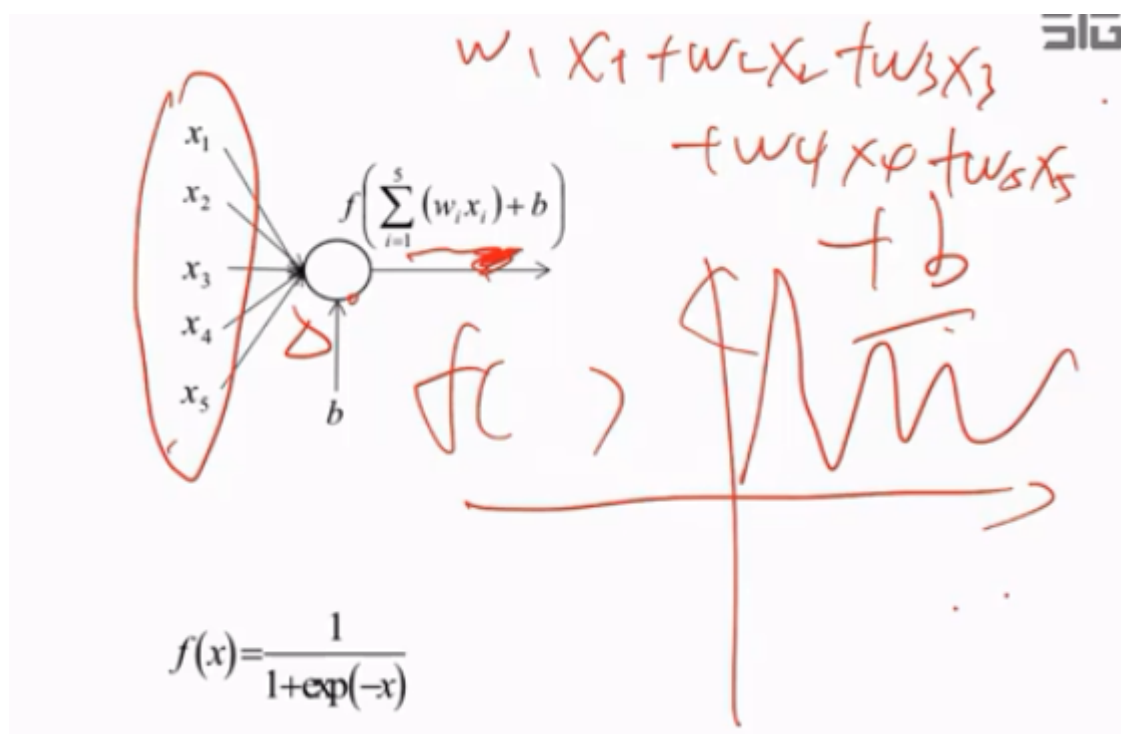


为什么需要激活函数?

向量和矩阵形式 $f(\mathbf{w}^T \mathbf{x} + b)$

$$f(x) = \frac{1}{1 + \exp(-x)}$$

看一下神经元的结构



11.5 sigmoid函数

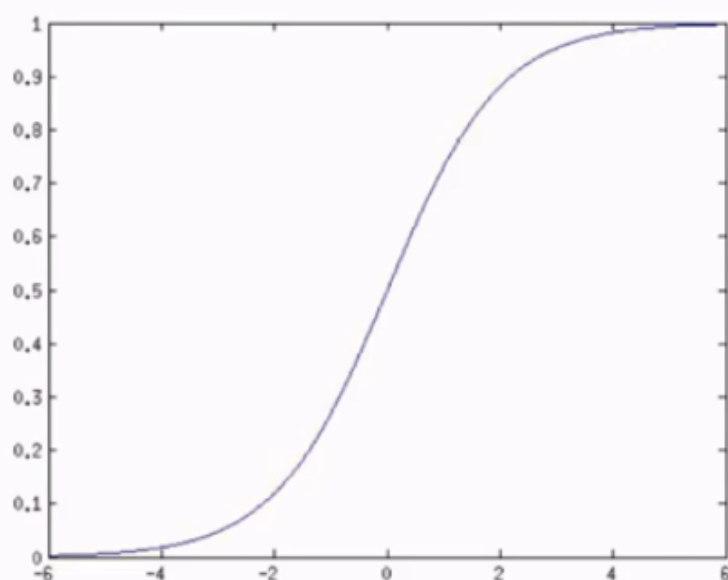
sigmoid函数

sigmoid函数

定义域

值域

单调性



函数的

- 定义域
- 值域
- 极限
- 单调性

$$f(x) = \frac{1}{1+e^{-x}}$$

$$(-\infty, +\infty)$$

$$(0, 1)$$

$$\lim_{x \rightarrow -\infty} \frac{1}{1+e^{-x}} = 0$$

\downarrow
 ∞

$$\begin{aligned}
 & \frac{1}{1+e^{-x}} \cdot e^{-x} \\
 & \frac{e^{-x}}{(1+e^{-x})^2} \\
 & \checkmark > 0 \\
 & \frac{1}{1+e^{-x}} \cdot e^{-x} \\
 & \frac{1}{1+e^{-x}}
 \end{aligned}$$

11.6 神经网络的结构

多层前馈型神经网络

神经网络结构

多层前馈型神经网络，也称为多层感知器模型MLP

是一个分层结构

输入层

隐含层

输出层

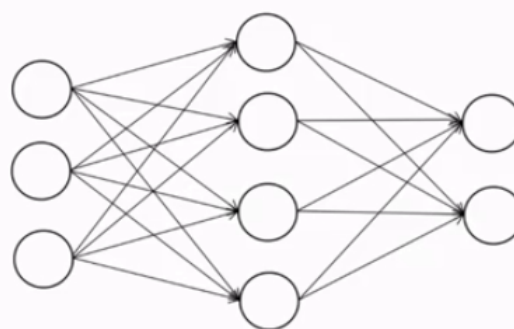
$$(x_1, x_2, x_3) \quad (y_1, y_2, y_3, y_4) \quad (z_1, z_2)$$

$$y_1 = \frac{1}{1 + \exp\left(-\left(w_{11}^{(1)}x_1 + w_{12}^{(1)}x_2 + w_{13}^{(1)}x_3 + b_1^{(1)}\right)\right)}$$

$$y_2 = \frac{1}{1 + \exp\left(-\left(w_{21}^{(1)}x_1 + w_{22}^{(1)}x_2 + w_{23}^{(1)}x_3 + b_2^{(1)}\right)\right)}$$

$$y_3 = \frac{1}{1 + \exp\left(-\left(w_{31}^{(1)}x_1 + w_{32}^{(1)}x_2 + w_{33}^{(1)}x_3 + b_3^{(1)}\right)\right)}$$

$$y_4 = \frac{1}{1 + \exp\left(-\left(w_{41}^{(1)}x_1 + w_{42}^{(1)}x_2 + w_{43}^{(1)}x_3 + b_4^{(1)}\right)\right)}$$



$$z_1 = \frac{1}{1 + \exp\left(-\left(w_{11}^{(2)}y_1 + w_{12}^{(2)}y_2 + w_{13}^{(2)}y_3 + w_{14}^{(2)}y_4 + b_1^{(2)}\right)\right)}$$

$$z_2 = \frac{1}{1 + \exp\left(-\left(w_{21}^{(2)}y_1 + w_{22}^{(2)}y_2 + w_{23}^{(2)}y_3 + w_{24}^{(2)}y_4 + b_2^{(2)}\right)\right)}$$

人工神经网络从数学上看就是一个符合函数，就是一个多层的复合函数。

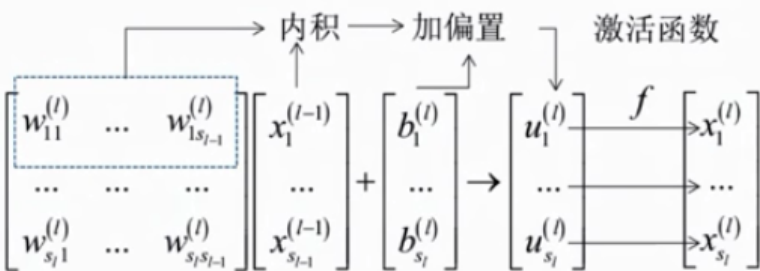
11.7 神经网络每一层完成的变换

神经网络每一层完成的变换

$$\mathbf{u}^{(l)} = \mathbf{W}^{(l)} \mathbf{x}^{(l-1)} + \mathbf{b}^{(l)}$$
$$\mathbf{x}^{(l)} = f(\mathbf{u}^{(l)})$$

← 权重矩阵的每一行为本层神经元与上一层所有神经元的连接权重

← 激活函数分别作用于每个神经元的输出值，即向量的每个分量，且使用了相同的函数



11.8 正向传播算法

完整的正向传播算法

完整的正向传播算法

设置输入值 $\mathbf{x}^{(1)} = \mathbf{x}$

循环，对于 $l = 2, \dots, m$

$$\mathbf{u}^{(l)} = \mathbf{W}^{(l)} \mathbf{x}^{(l-1)} + \mathbf{b}^{(l)}$$
$$\mathbf{x}^{(l)} = f(\mathbf{u}^{(l)})$$

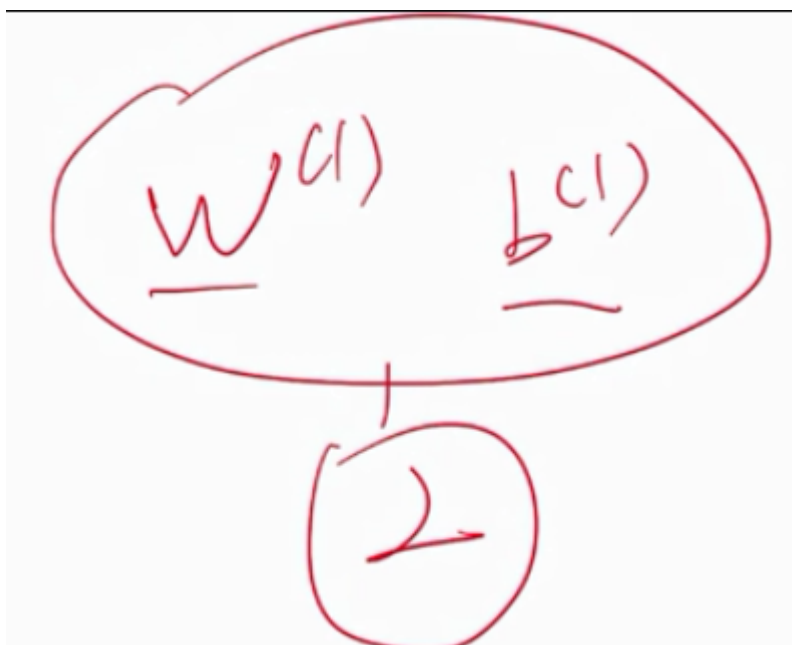
结束循环

输出 $\mathbf{x}^{(m)}$

权重 w , b 是通过训练得到的。

11.9 神经网络的本質

神经网络本质上是一个多层的复合函数
通过调整权重和偏置项实现不同的映射
权重和偏置项的值通过训练得到



怎么设置网络？后面会讲。

11.10 怎样用于实际问题

怎样用于实际问题

分类问题-输入值为特征向量或原始数据，输出值为one-hot编码

回归问题-输出值为回归函数值

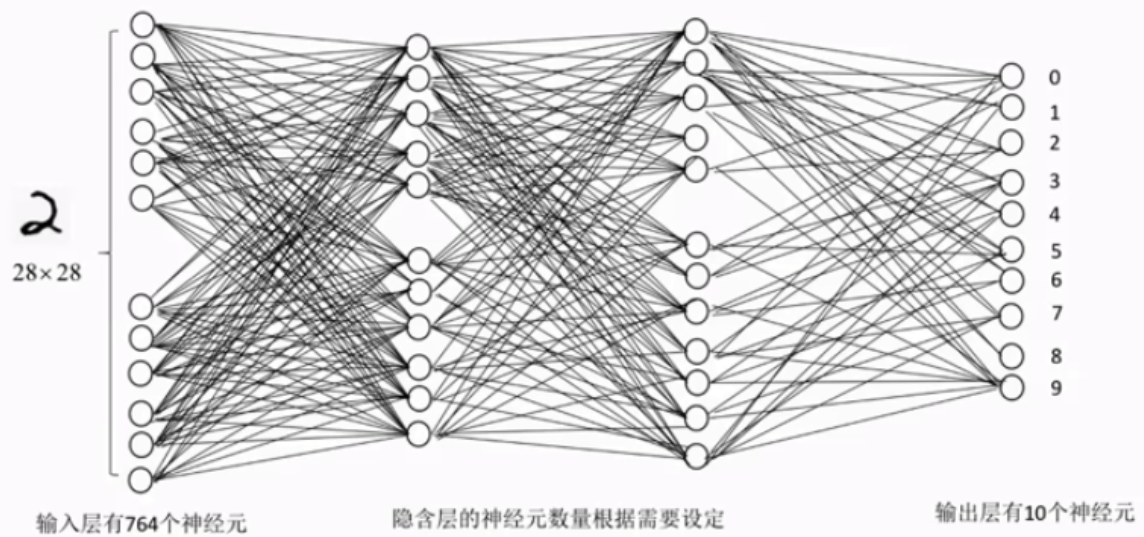
对于分类问题，分类结果为输出层神经元的最大值

对于回归问题，直接是输出层的值

- 分类
- 回归

手写数字识别的例子

分类问题-手写数字图像识别



- 输入层有764个神经元
- 隐含层的神经元数量根据需要设定
- 输出层有10个单元

这里只是预测值

对于二分类问题，对于回归问题