

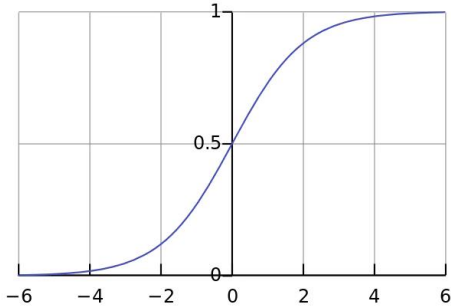
## 17.07.14 ( logistic函数和softmax函数 )

### logistic函数

其实逻辑斯谛函数也就是经常说的sigmoid函数，它的几何形状也就是一条sigmoid曲线。

logistic函数的公式形式如下：

其中， $x_0$ 表示了函数曲线的中心 ( sigmoid midpoint )， $k$ 是曲线的坡度。



逻辑斯谛回归 ( Logistic Regression，简称LR ) 作为一种对数线性模型 ( log-linear model ) 被广泛地应用于分类和回归场景中。

此外，logistic函数也是神经网络最为常用的激活函数，即sigmoid函数。

### softmax函数

softmax函数的本质就是将一个 $K$ 维的任意实数向量压缩 ( 映射 ) 成另一个 $K$ 维的实数向量，其中向量中的每个元素取值都介于 ( 0 , 1 ) 之间。

softmax函数形式如下：

其中 $j=1,2,\dots,K$

softmax函数经常用在神经网络的最后一层，作为输出层，进行多分类。

logistic具体针对的是二分类问题，而softmax解决的是多分类问题，因此从这个角度也可以理解logistic函数是softmax函数的一个特例。

softmax回归进行的多分类，类与类之间是互斥的，即一个输入只能被归为一类；

多个logistic回归进行多分类，输出的类别并不是互斥的，即“苹果”这个词语既属于“水果”类也属于“3C”类别。

---

### 为什么要对数据进行归一化处理

数据归一化后，最优解的寻优过程明显会变得平缓，更容易正确的收敛到最优解。

---

### 为什么神经网络参数不能全部初始化为全0

本来我们希望不同的结点学习到不同的参数，但是由于参数相同以及输出值都一样，不同的结点根本无法学到不同的特征！这样就失去了网络学习特征的意义了。隐藏层与其它层多个结点，其实仅仅相当于一个结点！

这样总结来看： $w$ 初始化全为0，很可能直接导致模型失效，无法收敛。

因此可以对 $w$ 初始化为随机值解决。