简单总结一下机器学习最常见的两个函数，一个是logistic函数，另一个是softmax函数，

1. logistic函数

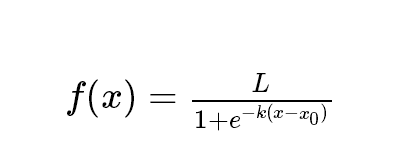
1.1 logistic函数定义

　　引用wiki百科的定义：

　　A logistic function or logistic curve is a common "S" shape (sigmoid curve).

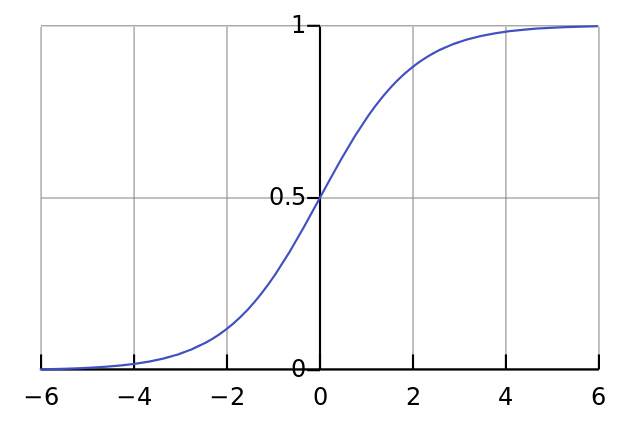
　　其实逻辑斯谛函数也就是经常说的sigmoid函数，它的几何形状也就是一条sigmoid曲线。

　　logistic函数的公式形式如下：



其中，x0表示了函数曲线的中心（sigmoid midpoint），k是曲线的坡度。

　　logistic的几何形状如下所示：



1.2 logistic函数的应用

　　logistic函数本身在众多领域中都有很多应用，我们只谈统计学和机器学习领域。

　　logistic函数在统计学和机器学习领域应用最为广泛或者最为人熟知的肯定是逻辑斯谛回归模型了。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。

**2. softmax函数**

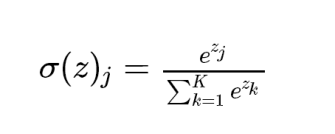
**2.1 softmax函数的定义**

　　同样，我们贴一下wiki百科对softmax函数的定义：

　　softmax is a generalization of logistic function that "squashes"(maps) a KK-dimensional vector zz of arbitrary real values to a KK-dimensional vector σ(z)σ(z) of real values in the range (0, 1) that add up to 1.

　　这句话既表明了softmax函数与logistic函数的关系，也同时阐述了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：



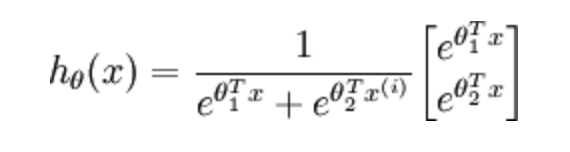
其中j=1,2,...,K

2.2 sotfmax函数的应用

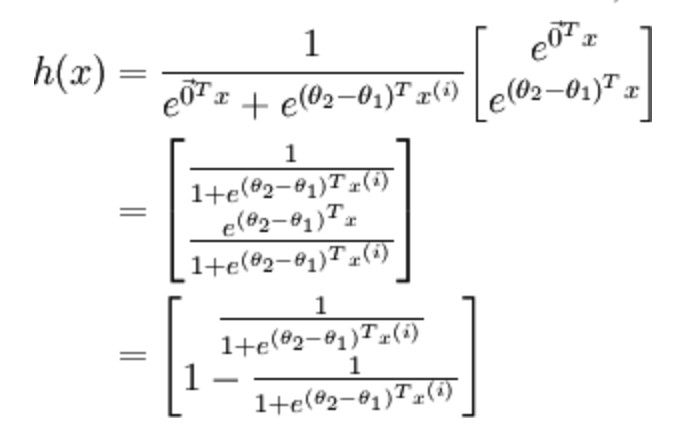
　　softmax函数经常用在神经网络的最后一层，作为输出层，进行多分类。

　　1）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。具体的推导过程如下：

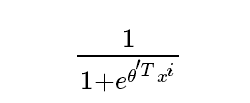
　　当分类数为2时，softmax回归的假设函数表示如下：



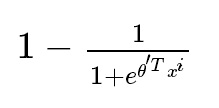
　　利用softmax回归参数冗余的特点，从两个参数向量中都减去向量θ1，得到：



最后，用θ′来表示θ2−θ1，上述公式可以表示为softmax回归器预测其中一个类别的概率为



另一个类别的概率为



　　 这与logistic回归是完全一致的。

　　2）从概率角度来看logistic和softmax函数的区别。

softmax建模使用的分布是多项式分布，而logistic则基于伯努利分布，（两点分布）

　　3）softmax回归和多个logistic回归的关系。

　　多个logistic回归通过叠加也同样可以实现多分类的效果，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。