**Tensorflow 的交叉熵损失函数和softmax回归**

假设在神经网络中，网络输出为y1,y2,...yn，由于我们需要将输出结果变成概率分布的形式提供给交叉熵损失函数计算，我们可以将softmax函数定义为：

softmax(yi)=exp(yi)/sigma(y1,….yn)

从而得到神经网络的输出的概率分布 (Q)，也就是预测的结果值。

之前已经说过交叉熵损失函数是由信息论的想法而来，假设预测值为离散结果，表示为

H(P,Q)=-EX~P(log(Q(x)))=-sigmaP(x)logQ(x)

我们通常将P(x)表示为正确答案，Q(x)为预测值。

举例：在MNIST手写数字数据集中，有one\_hot=True，表示的是将样本对应的标注信息（label）加上。如果图片信息为1，则one\_hot的表示方法为[1,0,0,0,0,0,0,0,0,0]

如果用此时用softmax得到的概率分布为[0.8,0.1,0.1,0,0,0,0,0,0,0]和[0.5,0.2,0.3,0,0,0,0,0,0,0]，我们会认为第一种分布更好，这是因为第一种的概率分布和标签的概率分布距离最近，使用H(P,Q)计算可以得到，

H1(P,Q)=-(1\*log(0.8)+0\*log(0.1)+0\*log(0.1)+7\*(0\*log(0)))~0.01

H2(P,Q) = -(1\*log(0.5)+0\*log(0.2)+0\*log(0.3)+7\*(0\*log(0)))~0.03

我们通常会选择比较小的损失函数值

在tensorflow中，提供了交叉损失函数和softmax连用的函数，tf.nn.softmax\_cross\_entropy\_with\_logits(\_sentinel=None,labels=None,logits=None,dim=-1,name=None)

\_sentinel:本质上是不用的参数，不用填

* labels:和logits具有相同的type(float)和shape的张量(tensor)
* logits:一个数据类型（type）是float32或float64
* name:操作的名字，可填可不填
* dims:类的维度，默认-1，也就是最后一维

其中labels表示正确的答案的分布，而logits表示预测概率分布。