数据结构，共有10组分类，0至9，其中每组有500个训练数据。

Matrix All\_theta =

是指代与矩阵X相乘时，使得对应的y=i的概率最大，而其余y的概率趋于0的一组参数，例如，对于，当X与之相乘时，得到的是一个向量y’，这个向量中，至这500个数是y=10的概率，而其余行的数值应该是趋于0的，实际上就是通过逻辑回归，将k=10这些个点从其余点筛选出来了。

X\*All\_thetaT= \*

max(X\*All\_thetaT，[]，2) 对矩阵逐行取最大值(maxValue)，并输出最大值的列号（index）