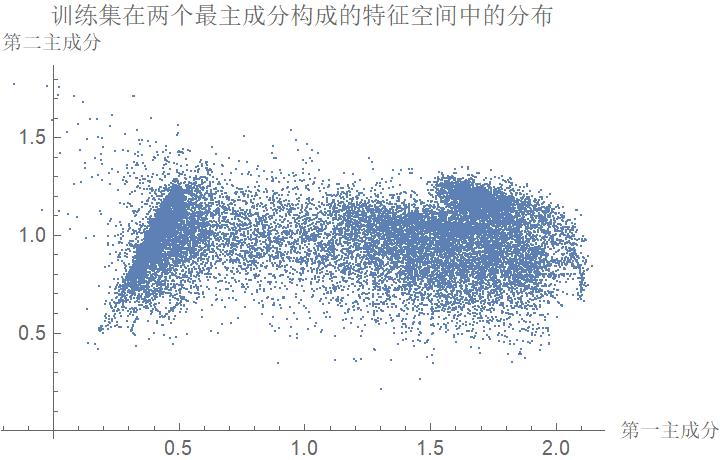
## 导入数据进行主成分分析

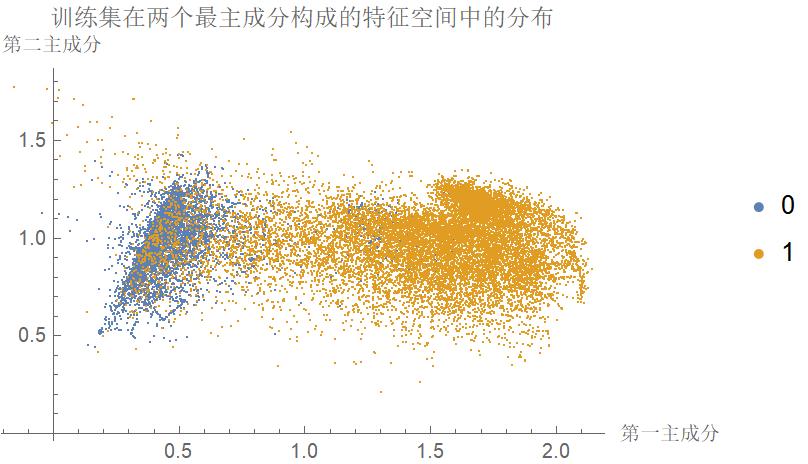
得到主向量：

并且得到主成分分析之后的数据（仍然是20000\*7的），我们对每个数据只取前两个分量（主向量上的分量）。将它们散布在特征空间上。

得到如下图像：



将它们按照标签"1"或者"0"赋予不同的颜色，得到：



我们来分析一下在这个问题中使用主成分分析的好处。

我们首先来求出协方差矩阵，

求取它的特征值，

特征值分别占总比例：

累计比例