### [主成分分析(PCA)](http://blog.csdn.net/google19890102/article/details/27969459)

#### 3.3.1.数据降维

对于现在维数比较多的数据，我们首先需要做的就是对其进行降维操作。降维，简单来说就是说在尽量保证数据本质的前提下将数据中的维数降低。降维的操作可以理解为一种映射关系，例如函数http://img.blog.csdn.net/20180105233016850，即由原来的二维转换成了一维。

#### 3.3.2.PCA的概念

PCA是一种较为常用的降维技术，PCA的思想是将维特征映射到维上，这维是全新的正交特征。这维特征称为主元，是重新构造出来的维特征。在PCA中，数据从原来的坐标系转换到新的坐标系下，新的坐标系的选择与数据本身是密切相关的。其中，第一个新坐标轴选择的是原始数据中方差最大的方向，第二个新坐标轴选取的是与第一个坐标轴正交且具有最大方差的方向，依次类推，我们可以取到这样的个坐标轴。

#### 3.3.3.PCA的操作过程

(1)去平均值，即每一位特征减去各自的平均值

(2)计算协方差矩阵

(3)计算协方差矩阵的特征值与特征向量

(4)对特征值从大到小排序

(5)保留最大的个特征向量

(6)将数据转换到个特征向量构建的新空间中