1. 通过计算训练样本的协方差矩阵的特征值和特征向量求得



1. 另一种通过SVD分解是直接对样本矩阵进行SVD变换获取特征值和特征向量求解问题。



KPCA方案针对PCA方案对线性数据无法很好的分类，采取的方案是KPCA核函数映射方案，将数据映射到高维空间中进行降维之后，再通过SVM线性分类器进行分类或者其它的操作。

PCA方案能够做的最好的是特征选取，对高维数据进行降维。现实中的SVM等分类器对于高维数据进行分类困难很大，对于高维数据进行降维之后，在使用线性分类器进行分类。而对于有些情况下，PCA进行降维之后的数据无法线性可分，那么使用线性分类器是于事无补的，因此提出的是KPCA方案，KPCA方案先经过先将数据映射到更高维空间，在高维度空间做PCA得到线性分类曲面，对数据进行先升后降维的操作，最后通过SVM线性分类器做出线性分类的操作。

而我的大论文中的方案是直接将PCA映射坐标轴的距离之值当作是数据检测的标杆，通过双重检测进行异常数据的检测。KPCA方案也是利用最大的非相似度来计算阈值做数据的异常检测。

可以做的点，根据SVM做一个基于马氏内核的SVM-Mahas异常数据检测模型。

PCA使用距离向量来做识别，我们可以当作是一类无监督的学习方案，使用已知数据的特征来做数据值的区分。一类无监督学习方案。

SVM需要找超平面的映射，PCA方案找的是最能反应数据特征的坐标轴的方向。

**使用距离PCScore来衡量数据的相似度**

RBF方案高斯方案的核心是使用高斯内核函数来计算，它考虑的是欧式距离并没有考虑节点间数据的相关性质！

**而马氏Maho距离使用的是特征距离来衡量节点之间的核函数关系，相关矩阵和核函数**