背景知识

Gram 矩阵就是协方差阵。

在做 POD 的过程中，寻找最优基向量时，发现协方差阵的特征值问题的解对应的各个特征向量刚好是各阶模态。

协方差阵是实对称阵。

实对称阵的性质：

特征值都是实数。实对称阵都是半正定阵。（什么是半正定？）

半正定的定义：

广义：设 A 是 n 阶方阵，如果对任何非零向量 X，都有 X'A*X≥0*，其中 *X'*表示 X 的转置，就称 A 为**半正定矩阵**。

狭义：设 A 为实对称矩阵，若对于每个非零实向量 *X*，都有 X*'AX*≥0，则称 A 为**半正定矩阵**，称 X'AX 为半正定二次型。（>0 就称为正定矩阵了；<0 负定；≤0 半负定；其他情况不定。）

什么样的函数是核函数？

Mercer 定理(A.D.1909)：任何半正定的函数都可作为核函数。

注解：构造协方差阵时，使用偏差的内积运算，这样得到的协方差矩阵中，每个元素是两个偏差向量的内积。类似的，也允许有其他运算 f(x)（广义的 f(x)啊），将两个偏差向量作为输入，得到的输出作为协方差矩阵的元素，如果这样得到的协方差矩阵是同样拥有实对称阵的半正定性质，那么这种运算就可以选为核函数。

核函数的确定并不困难,满足 Mercer 定理的函数都可以作为核函数。通常定义为空间中任一点 x 到某一中心 xc 之间欧氏距离的单调函数 ,可记作 k(||x-xc||)，其作用往往是局部的，即当 x 远离 xc 时函数取值很小[1]。

常用的核函数有：线性核函数，多项式核函数，径向基核函数，Sigmoid 核函数和复合核函数，傅立叶级数核，B 样条核函数和张量积核函数等[2]。其中高

斯核函数最常用，可以将数据映射到无穷维，也叫做径向基函数（Radial Basis Function 简称 RBF），是某种沿径向对称的标量函数[3]。

核函数有什么优点？

设 x,z∈X,X 属于 R（n）空间，非线性函数Φ实现输入空间 X 到特征空间 F 的映射，其中 F 属于 R（m），n<<m。根据核函数技术有：

$$K(x,z) = <\Phi(x), \Phi(z)> \tag{1}$$

其中：<, >为内积，K(x,z)为核函数。

注解 1：n 维的 x 经过Φ映射到高维空间后得到 m 维的Φ(x)；Φ(z)同理，也是 m 维。那么<Φ(x)，Φ(z) >就是两个 m 维向量的内积运算。而 K(x,z)是对 n 维的 x 和 n 维的 z 进行运算，假设 K 也是内积运算，那么它就是两个 n 维向量的内积运算，因为 n<<m，所以 K 的运算量<<m 的情况，极大地减少了获取高维空间中 1 个元素的运算次数，将成倍地加快计算。

注解 2：核函数会随着Φ的变化而变化，即式(1)中的自变量是映射关系Φ，因变量是关于 x 和 z 的函数关系 K。在 POD 中，协方差阵中的元素是偏差阵的内积，也就是说：输入映射到了偏差阵维度的空间，而输入的空间也是偏差阵的维度，即映射前后偏差阵并未变化，原来是什么列向量，映射之后还是什么列向量。POD 中的这种映射属于“线性核函数”的类别[4]。

核函数的修正

在确定了核函数之后，由于输入数据可能存在一定的误差，考虑到通用性问题，因此引入了松弛系数以及惩罚系数两个参变量来加以校正[5]。

注解：也就是让核函数在数据恶劣的条件下也能有良好的效果，即误差过滤。（松弛系数、惩罚系数不是一样的吗？不都是权系数吗？）

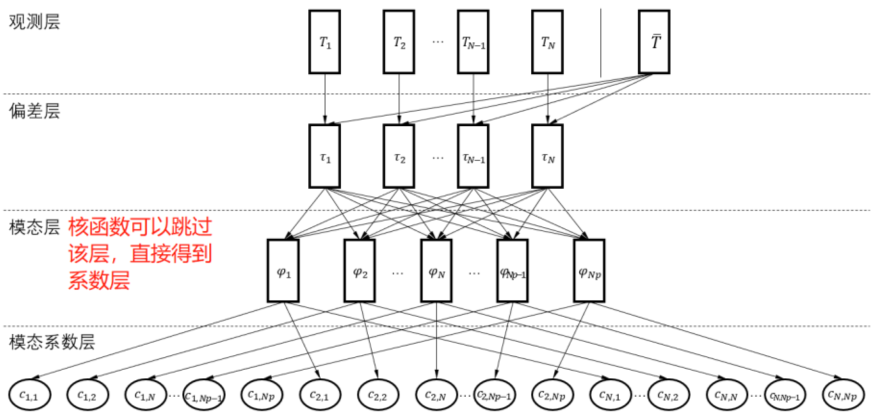
核函数有什么缺点？

核函数的引入避免了“维数灾难”，大大减小了计算量。而且输入空间的维数 n 对核函数矩阵无影响[6]，因此，核函数方法可以有效处理高维输入。但是，

避免了维数灾难的同时也会忽略掉高维空间中的信息，虽然使用核函数技巧能够减少计算量，但由于对高维信息的忽略（只是知道特征在特征方向上的投影），将导致得到的基函数不具备通用性。信息缺失将产生欠拟合，因此泛化能力不佳。

核函数的推导参考：https://blog.csdn.net/zjuPeco/article/details/77510981

如果只是为了分类，则不需要知道高维空间的模态，只需要知道输入向量映射到高维空间之后，它对应于各个模态的系数即可。只要知道了系数，即可用系数来实现聚类和分类。而 KPCA 方法通过增加核函数的方式，可以跳过对高维空间中的模态的计算，直接得到模态系数。所以 KPCA 方法更适用于分类。在这样的情况下，核函数的选择原则是使经过核函数变换之后的类协方差阵的特征值及特征向量求解尽可能简单（即矩阵尽量稀疏或三角）。



核函数公式的参考：https://www.cnblogs.com/xiaohuahua108/p/6146118.html

参考资料：

[1] https://baike.baidu.com/item/%E6%A0%B8%E5%87%BD%E6%95%B0

[2]刘明. 支持向量机中 Sigmoid 核函数的研究[D]. 西安电子科技大学, 2009.

[3] 楼俊钢, 蒋云良, 申情,等. 软件可靠性预测中不同核函数的预测能力评估[J]. 计算机学报, 2013, 36(6):1303-1311.

[4] https://blog.csdn.net/cqy_chen/article/details/77932270

[5] https://blog.csdn.net/zhazhiqiang/article/details/20146243

[6] 王华忠, 俞金寿. 核函数方法及其在过程控制中的应用[J]. 石油化工自动化, 2005(1):25-30.