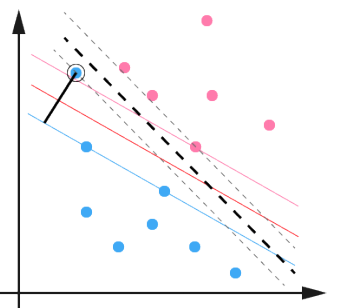
在前两篇笔记中我们介绍了两种的情况，即：本身数据是线性可分的或者本身数据线性不可分，但是通过非线性特征变换后变得可分的情况。一般而言，通过非线性特征变换后数据被映射到高维的空间，能够线性可分的概率大大增加了，但是某些情况还是很难处理。例如可能并不是因为数据本身的非线性结构，而只是因为数据有噪声，对于这种偏离正常位置很远的数据点，叫做离群点。



如上图所示，蓝色黑圈的点就是一个离群点，明显偏离蓝色群落，可能是噪声导致的，假若落在正常范围内，得到的分离面是红色线，但是由于离群点的存在，使得分离面变成了黑色虚线，这样导致间隔变小了，如果再往右上走点，可能就变得不可分了。很明显，的存在导致分类面的鲁棒性变差，抗干扰能力变差，同时，由于的存在，如果要做非线性特征变换，则需要很复杂的非线性特征变换，导致参数变多，很容易就。

然而，终究不是一个优化问题，而是一个学习问题，我们希望模型有好的效果，就需要对模型做优化处理。显然 在实际应用中由于的存在，容易导致，不太实用，因此，为了把考虑进来，我们对约束条件做些放宽，就是。

给出训练样本：



存在非线性特征变换：



 要求两类训练样本不能违反任何边界，目标函数如下：



 考虑到，对约束条件引入松弛变量，



引用上篇笔记的话：由于非线性特征变换导致维度较大，而且往往非线性变换的形式是未知的，用二次规划非常难解，但是在预备知识中我们讲过：无论原问题是什么问题，其对偶问题总是个凸优化的问题，于是我们引入预备知识中的拉格朗日乘子和乘子，得到函数：



其中：为乘子。

由预备知识可知，满足条件是解规划问题（1）的必要条件，条件如下：

1. ：

令（2）式对的偏导数为0得：



可得：



令（2）式对的偏导数为0得：



可得：



令（2）式对的偏导数为0得：



可得：



2）：



3）：



1. ：



 问题等价于带约束条件的拉格朗日函数问题：



其对偶问题为：



显然：



可以进一步证明，等号成立（强对偶）。



将（3）（4）（5）代入（7）中：

 目标函数变为：



很明显上述问题也是一个问题。

利用求出乘子之后，如何求和呢？



至于，利用（5）和（6）式：

当任何时，必有：



得到：



得到和后，可得的预测函数：



 目标函数（8）中的表示原始特征空间经过非线性特征变换得到新的特征空间，表示内积运算。所以 求解过程中主要有两个步骤要做：

1. 将原始特征空间通过非线性特征变换映射到高维空间；
2. 在高维空间中做内积运算。

那么问题来了：

首先，轻易的找到一种非线性特征变化是很困难的；

其次，这种高维空间维度可能是爆炸性的，大到内积等运算无法进行。

那是否有办法直接在低维度空间中进行计算，而不需要显示地写出映射后的计算再计算？

：

即定义一个核函数，所以我们只需要定义核函数的形式即可。

1. 线性核函数：



1. 多项式核函数：



1. 高斯核函数：



1. 核函数：



利用核函数， 目标函数变为：



预测函数为：





2017.07.30.