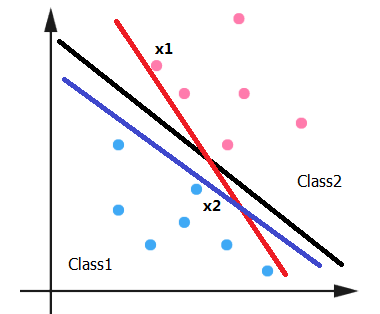
是机器学习中的一种分类算法，所谓支持向量就是支持或支撑平面上把两类或多类类别划分开的向量点，就是一个算法。

本篇笔记主要学习的和，指的就是分类面是线性的，对于任何线性分类器，我们假设其分类超平面为；表示表示所有的训练样本是可以被分开的。

给出训练样本：



先简单描述下，为了便于可视化，假设样本只有2个特征，样本分布如下图所示：



图中给出三个分类面，分别用黑、红、蓝三种颜色表示，训练样本离红色分类超平面较近，一个属于且与距离相近的样本点可能被红色分类超平面分类到中，从而造成分类错误；离相近的样本点对于蓝色分类超平面可能也有相同影响，所以红色和蓝色的分类超平面具有较差的鲁棒性。然而所有样本点到黑色分类超平面的距离都很大，所以黑色分类超平面的抗噪声能力最强，最优最强的鲁棒性。

黑色分类超平面就是要找的分类超平面，数学上描述就是：**就是要找出一个分类面使得离它最近的点到它的距离最大。**

构建和的准则：

1. **所有训练样本都能分开；**
2. **使离分类面最近的点到分类面的距离最大。**

若1）满足，则存在分类面可以把每个样本都分开，即：

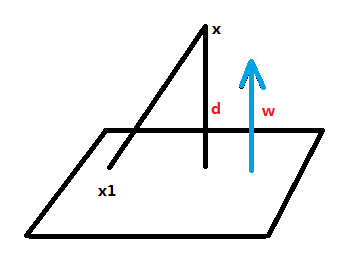








为满足2），首先回顾下如何计算点到平面的距离，记距离为，如下图所示：



如下式所示：



定义边界：离分类面最近的点到该分类面的距离，如下式所示：



结合准则，标准问题可以转化为：





令：



则：



令：



可得：



所以标准问题转换为：











得到：



至此，问题转化为了一个最优化问题，且为一个二次规划问题，目标函数为二次函数，约束条件是线性的。这个问题可以利用现成的优化包求解，求出和之后，其判别函数为：





2017.07.29.