SVN学习总结笔记

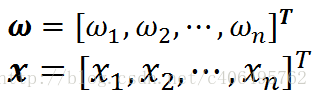
## 一、概念

支持向量机（Support Vector Machine, SVM）是一类按监督学习（supervised learning）方式对数据进行二元分类的广义线性分类器（generalized linear classifier）[1]

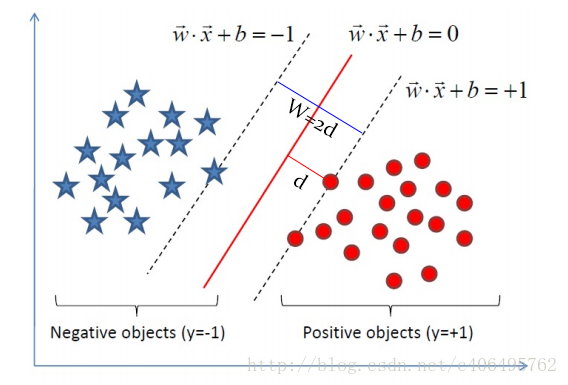
## **原理**

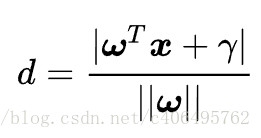
**目的**：求解能够正确划分训练数据并且几何间隔最大的分离超平面；

**支持向量：**两侧虚线所穿过的样本点，就是SVM中的支持样本点，称为"支持向量"

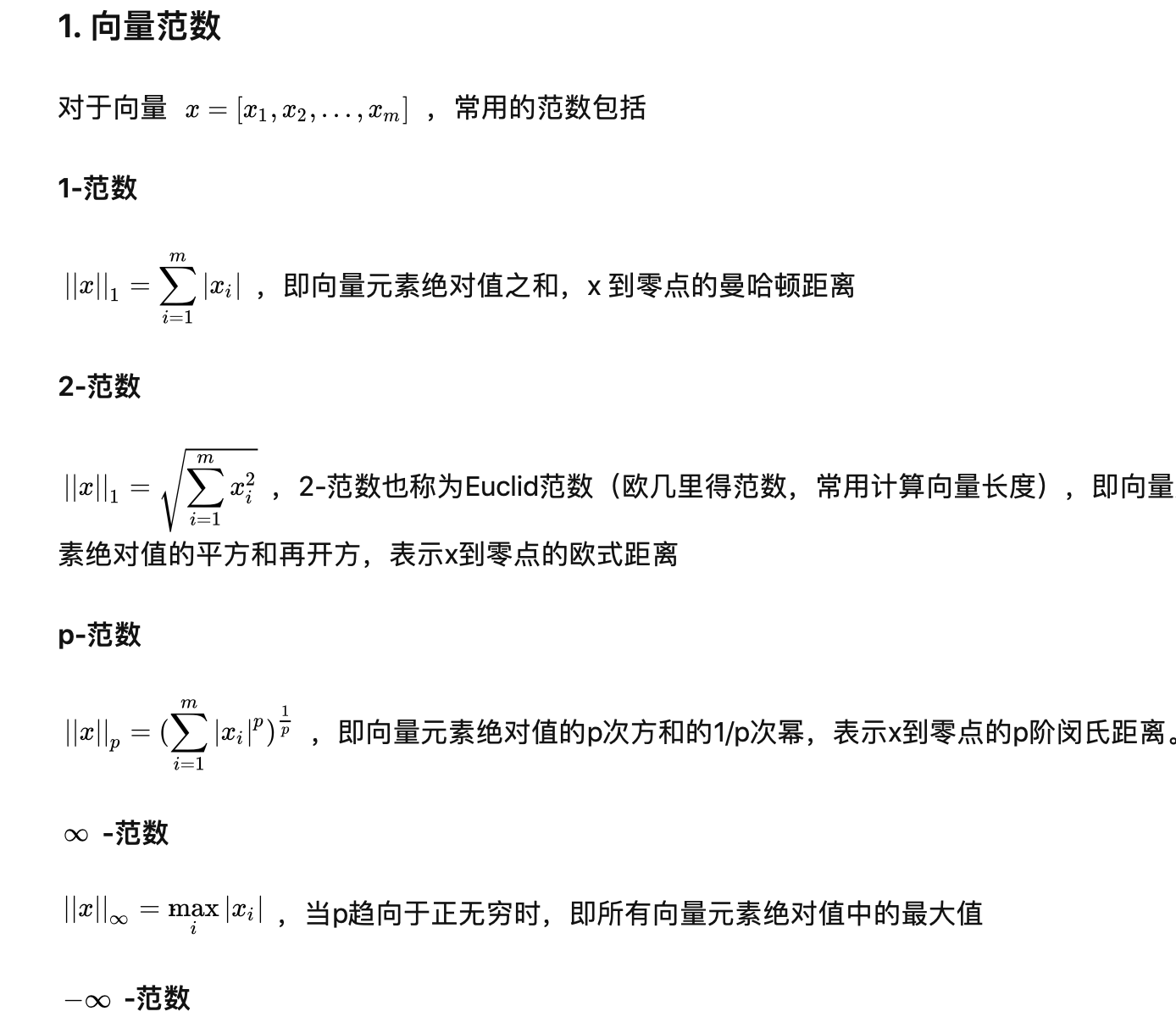
**超平面定义：**IMG_256

**二维示例图：**

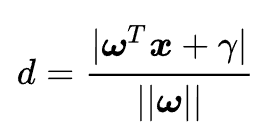
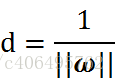


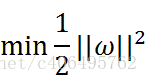
**超平面分类间隔d：**（||w|| w向量的二范数）

**范数概念：**范数，是具有“长度”概念的函数。是矢量空间内的所有矢量赋予非零的正长度或大小。

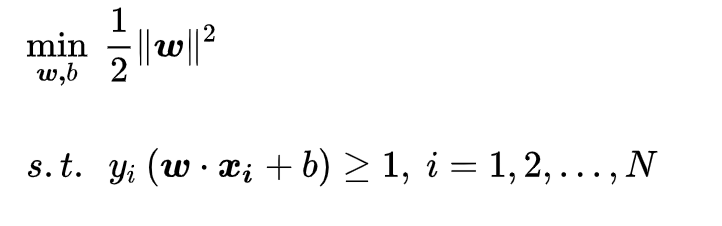


## 三、推导[2]

1. 
2. IMG_256 
3. 求超平面关于所有样本点的几何间隔的最小值，也就是求d的最大值，即求||w||的最小值，等价于



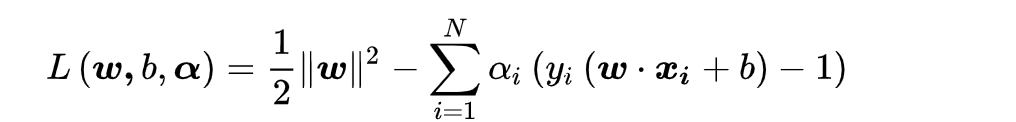
1. 最终得出：



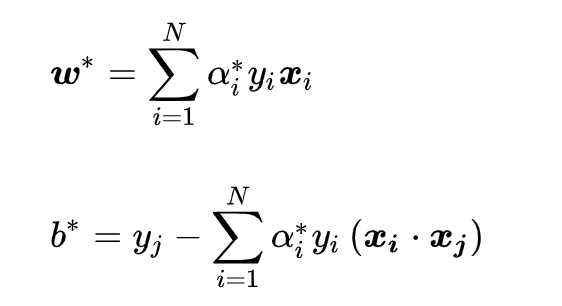
## 求解[3]：

对其使用拉格朗日乘子法得到其对偶问题（dual problem）

将有约束的原始目标函数转换为无约束的新构造的拉格朗日目标函数:



最终求得：



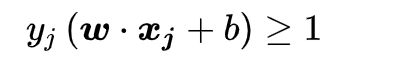
为拉格朗日乘子[4]

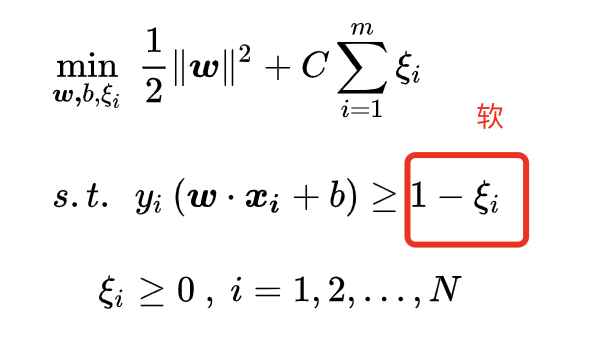
有很多教材上会把最优解记作变量的右上角添加星号，若想表示w的最优解，即为w\*

## 扩展

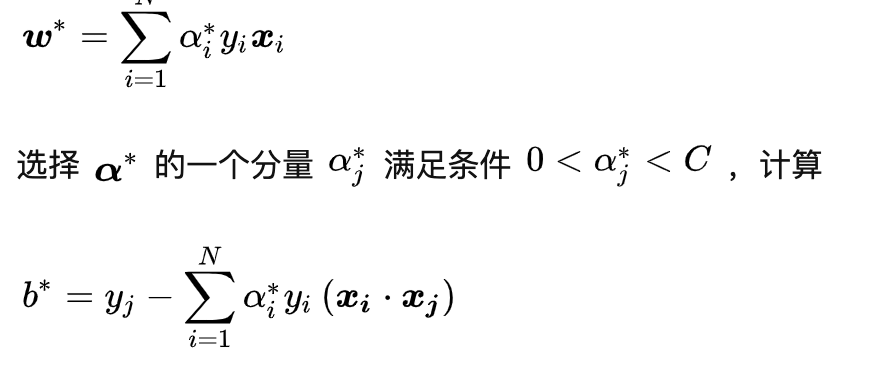
### 1、软间隔

**【四】**中是基于训练集数据线性可分的假设下进行的，但是实际情况下几乎不存在完全线性可分的数据，为了解决这个问题，引入了“软间隔”的概念，即允许某些点不满足约束。

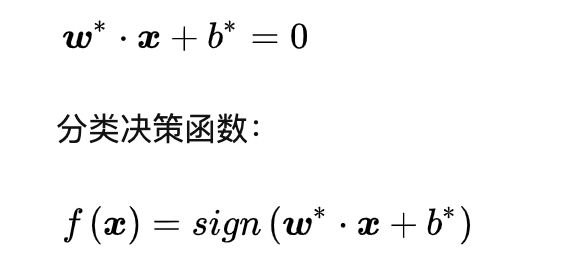
1. 
2. 采用hinge损失，将原优化问题改写为（ C称为惩罚参数）



1. 求得：

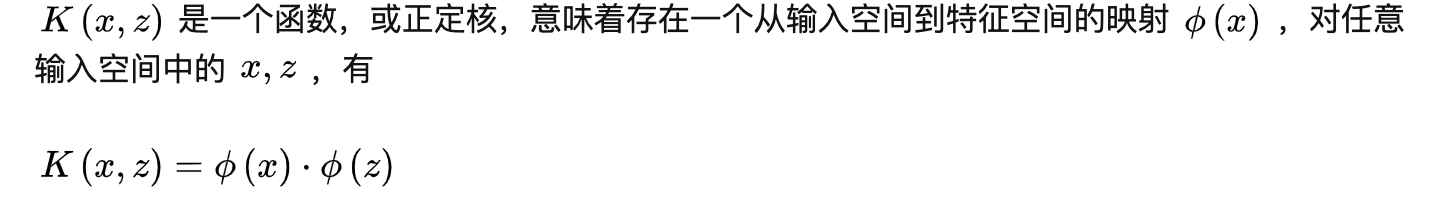


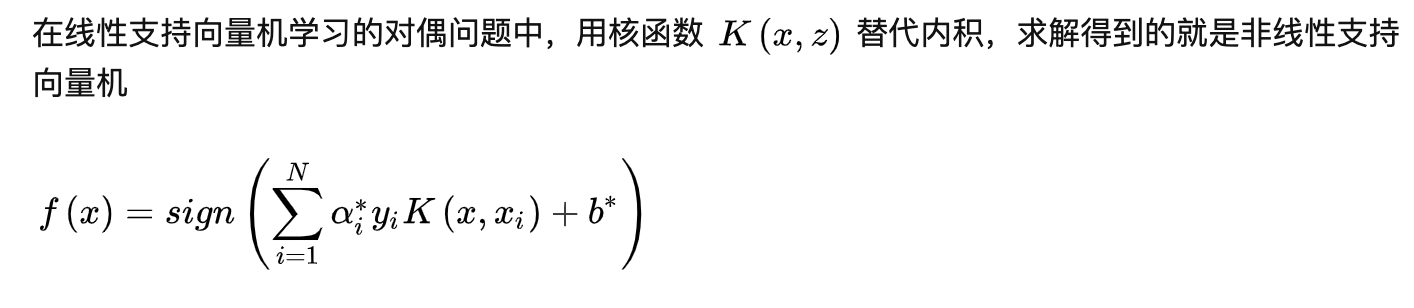
1. 最终

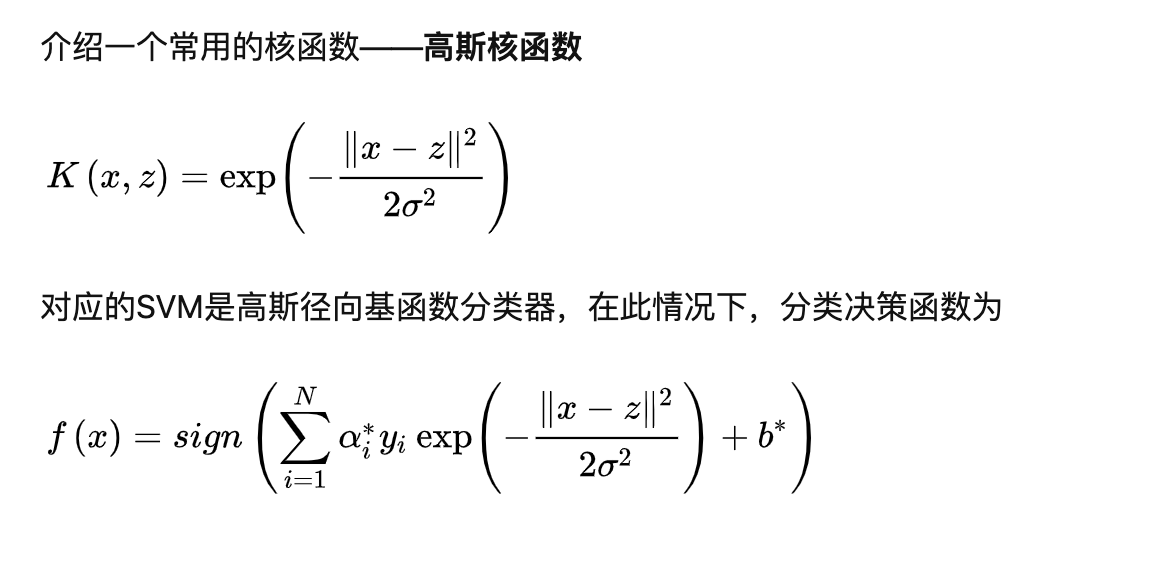


### 非线性svm[4]

对于输入空间中的**非线性分**类问题，可以通过非线性变换将它转化为**某个维特征空间中**的**线性分类**问题，在高维特征空间中学习线性支持向量机。由于在线性支持向量机学习的对偶问题里，目标函数和分类决策函数都**只涉及实例和实例之间的内积，所以不需要显式地指定非线性变换**，**而是用核函数替换当中的内积**。核函数表示，通过一个非线性转换后的两个实例间的内积。







## 代码实践

见https://github.com/zhaigang2333/learn

[1].https://baike.baidu.com/item/%E6%94%AF%E6%8C%81%E5%90%91%E9%87%8F%E6%9C%BA/9683835?fromtitle=SVM&fromid=4385807&fr=aladdin

[2].https://blog.csdn.net/c406495762/article/details/78072313#2-smo%E7%AE%97%E6%B3%95%EF%BC%89

[3].https://zhuanlan.zhihu.com/p/31886934

[4].https://www.zhihu.com/question/38586401