支持向量机主要可分为三类：线性可分支持向量机、线性支持向量机、非线性支持向量机

1.线性可分支持向量机

线性可分支持向量机也成为硬间隔（Hard margin）支持向量机，即要求所有的样本都要被分类正确。

输入：T={(x1,y1), (x2, y2) … (xn,yn)}

其中，xi表示训练数据集第i个实例，是一个向量。

yi表示第xi的类标记 yi ={-1, +1}

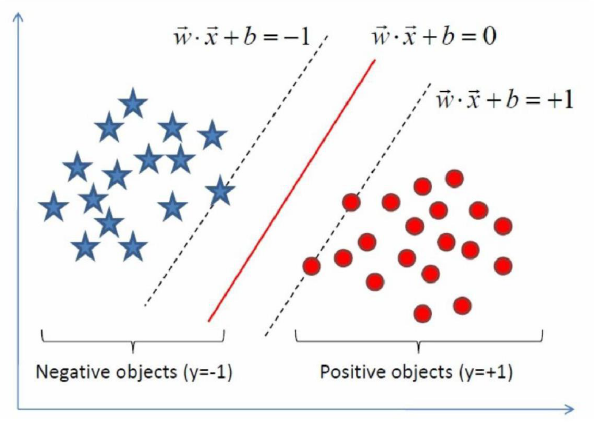
(xi, yi)为样本点

给定线性可分数据集，通过间隔最大化得到分离超平面： 相应的分类决策函数为

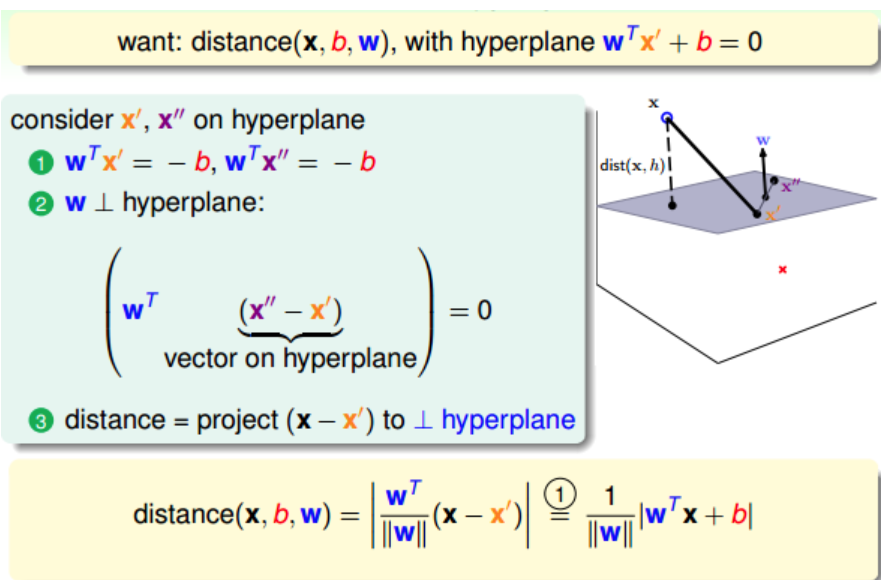
即线性可分SVM的任务是不仅要找到把样本点正确分类的超平面，而且是要找到间隔最大的超平面。

此时间隔的定义：设C和D为两不相交的凸集，则存在超平面P，P可以将C和D分离，两集合间的距离定义为两集合中元素的最短距离，C,D两集合最短距离的中垂线（多维空间即中垂面）即为超平面，任一集合到超平面的距离即为间隔。

由上可知，线性可分SVM即求距离超平面最近的点的间隔最大化。从而增强了SVM的泛化能力。



距离求解：假设x1,x2均为超平面上的点，即x1,x2两点满足 即， x1 – x2可以看成是超平面上的向量，即w为超平面法向量。点X到超平面的距离等于，X在法向量方向上的投影长度。则



则目标函数为…………………………(1)

超平面： |y|通过缩放w,b使两类的值都满足 ……………………………(2)

由1，2两式可得目标函数转换为：………………(3)

s.t.

将目标函数转换为………………………(4)

s.t.

根据新的目标函数可知，目标函数为2次式，约束条件为一次式，典型的二次规划问题。使用拉格朗日乘子法。构造拉格朗日函数

………………………(5)

s.t.

s.t. ………………………(6)

因为 且则 小于等于0，所以max()L() = , 即原问题可以表示为,其对偶问题为,SVM刚好满足这些条件。

先求 即对对(w,b)求导等于0。

= 0 即 ………………………(7)

………………………(8)

将（7）代入（5）得到 ………………………(9)

对偶问题有最优解必然满足KKT条件

s.t. ………………………(10)

由10式可以看出当时，即 即样本点在支持向量，当时，，即这个样本点是支持向量以外的点，此时9式中

可以省去很多计算，换句话说，w,b只与有关。

根据上面式子求解所有，将代入7式即可以求w, 对于任意支持向量（）有 且 ………………………(11)

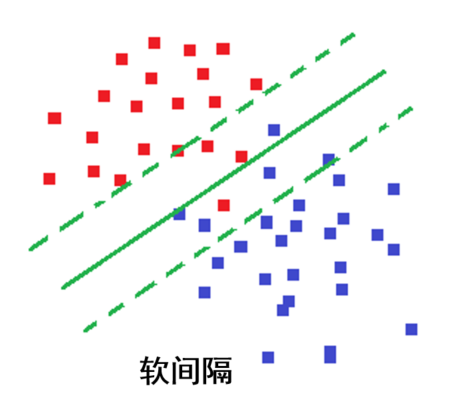
对于其他任意向量（）有 ,所以 所以

………………………(12)

支持向量点的越少，支持向量机的泛化性越好。

怎么解？对偶问题怎么求解？

线性支持向量机（soft margin）即可以允许部分分类存在错误，但是给予损失项



同上目标函数变为在线性可分支持向量机的目标函数基础上加上损失项

………………………(13)

其中c 为惩罚系数，h(z)为损失函数

常见的 0/1损失

指数损失

对率损失

SVM使用hinge loss

Hinge损失 ，当分类正确时z大于等于1，此时h(z) = 0，即无损失，不惩罚，当分类错误是，z小于1，此时，h(z) = 1-z,错误距离越大，则损失越大。

带有hinge loss的对应优化目标

………………………(14)

引入松弛变量

………………………(15)

s.t. ………………………(16)

………………………(17)

同样的方法得到对偶函数

………………………(18)

s.t.

s.t. c为惩罚系数，当c无穷大时，不允许有分类错误，即变为硬间隔。

同样怎么求解对偶问题？

SMO：通过把其他都固定住，仅仅优化某一对，这样就把N个优化变成优化两个因为对偶问题最优解必定满足KKT条件，所以当所有满足KKT条件，则优化完成，得到最优解。

**，**假设固定了 则原优化目标变为(省去常数项)：

………………………(19)

约束条件s,t.

………………………(20)

由20式可得

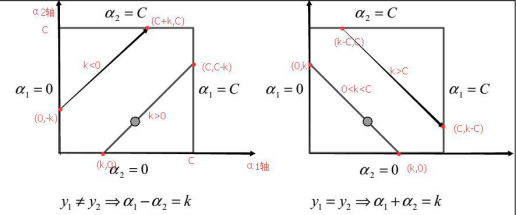
………………………(21)

将21代入到19可得关于的函数，函数求导等于0，再将代入可得

其中

由于

所以约束共通过决定可行域（同种）



由于y的取值为+1或者-1

当y1与y2同号时例如同为1时，即右边的图形

可能的几种取值

：为空值

： 此时与正方形区域焦点为(0,0)，

此时直线和正方形区域相交，即的可行区间是[0,]，即[0,]

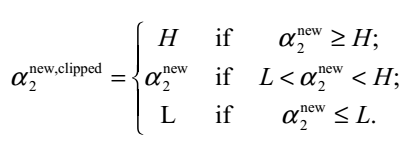
当 可行区间[]即]  
当 可行域为空集

综上的可行域为 即

同理，当y1与y2异号时

的可行域为，即

所以在更新时，要先求出的可行域，然后用之前的那个公式求出极值点**，**然后看极值点处的在不在可行域范围内，在的话就使用极值点处的，不在的话就使用边界值H或者L更新



更新，可根据公式更新

怎么选择

1. 寻找违反KKT条件的参数，因为最优解必满足KKT条件，当参数不满足KKT条件时就不是最优，所以进行优化。
2. 首先遍历数据集，对每个不满足KKT条件的参数作为第一个待修改参数，作为外层循环
3. 对整个数据集遍历完一遍之后，选择那些参数满足，开始遍历，如果发现一个不满足KKT条件的作为第一个待修改参数，然后找到第二个待修改的参数，修改完之后，重新开始遍历这个子集
4. 遍历完子集后，重新开始（1）（2），直到在（1）（2）执行时，没有任何修改就结束
5. 启发式寻找

找满足|最大

寻找一个随机位置满足的可以优化的参数进行修改

在整个数据集上寻找一个随机位置的可以优化的参数进行修改

都不行那就找下一个第一个参数