凸集：s.t. subject to 的缩写

连续型优化问题的一般形式：目标函数，约束函数，等式约束，不等式约束。

**凸集的定义：凸组合，凸包。**

若D是凸集，则D中任意m个点的凸组合仍属于D

在欧氏空间中，凸集是对于集合内的每一对点，连接该对点的直线段上的每个点也在该集合内。立方体是凸集，但是任何中空的或具有凹痕的例如月牙形都不是凸集。

凸集，实数R上（或复数C上）的向量空间中，如果集合S中任两点的连线上的点都在S内，则称集合S为凸集。

设向量https://gss1.bdstatic.com/9vo3dSag_xI4khGkpoWK1HF6hhy/baike/s%3D118/sign=b40bfb6105f41bd5de53ecf569da81a0/0b55b319ebc4b7458abcab5cc4fc1e178a82154b.jpg 如有实数https://gss0.bdstatic.com/94o3dSag_xI4khGkpoWK1HF6hhy/baike/s%3D43/sign=2b8423ab69d9f2d3241125eca8ec8498/37d12f2eb9389b501fb2fc188e35e5dde7116e2e.jpg ，且https://gss3.bdstatic.com/-Po3dSag_xI4khGkpoWK1HF6hhy/baike/s%3D66/sign=18ce25498c44ebf869716739d8f9ff5f/9358d109b3de9c82a2f6e5c56781800a18d843ec.jpg ，则称https://gss0.bdstatic.com/-4o3dSag_xI4khGkpoWK1HF6hhy/baike/s%3D52/sign=5927fe502334349b70066e87c9eaf637/b03533fa828ba61ed4d9d7654a34970a304e5940.jpg 为向量 https://gss1.bdstatic.com/-vo3dSag_xI4khGkpoWK1HF6hhy/baike/s%3D24/sign=d21b85b972310a55c024d9f0b64550db/c8ea15ce36d3d539932d10893187e950352ab063.jpg的一个凸组合(凸线性组合）。

关系：如果点https://gss2.bdstatic.com/-fo3dSag_xI4khGkpoWK1HF6hhy/baike/s%3D132/sign=9a99024e3f9b033b2888f8d927cf3620/55e736d12f2eb938d5e365b8de628535e4dd6fe9.jpg 的任意凸组合仍包含在D中，则D一定为凸集。

**两个向量的凸组合：**

https://gss2.bdstatic.com/9fo3dSag_xI4khGkpoWK1HF6hhy/baike/s%3D195/sign=89eb81668c01a18bf4eb1646ab2e0761/08f790529822720e0ecce5b070cb0a46f21fab7d.jpg

即：https://gss3.bdstatic.com/7Po3dSag_xI4khGkpoWK1HF6hhy/baike/s%3D145/sign=ac38b07851afa40f38c6cad99e65038c/29381f30e924b89976c97e7265061d950b7bf687.jpghttps://gss3.bdstatic.com/-Po3dSag_xI4khGkpoWK1HF6hhy/baike/s%3D144/sign=928a382fb819ebc4c478729db627cf79/63d9f2d3572c11df4257b869682762d0f603c2f2.jpg

表示https://gss1.bdstatic.com/-vo3dSag_xI4khGkpoWK1HF6hhy/baike/s%3D9/sign=f41498629c2bd40746c7dfcd7b10de/6c224f4a20a44623f4c7a3289322720e0cf3d76f.jpg 在连结https://gss3.bdstatic.com/-Po3dSag_xI4khGkpoWK1HF6hhy/baike/s%3D24/sign=20e8fa901ece36d3a60484343bf3f18a/314e251f95cad1c8dcdbfa6d743e6709c93d5113.jpg 两点的线段上（向量共线），https://gss2.bdstatic.com/9fo3dSag_xI4khGkpoWK1HF6hhy/baike/s%3D18/sign=405e51a3f9d3572c62e298d48b1383bc/4ec2d5628535e5ddad9ea3957dc6a7efcf1b62cd.jpg 中的情况也是一样。

三个向量https://gss2.bdstatic.com/9fo3dSag_xI4khGkpoWK1HF6hhy/baike/s%3D64/sign=c0f5575d69d0f703e2b296d809fa9d11/d4628535e5dde711da0942bcacefce1b9c1661a4.jpg 的凸组合，https://gss1.bdstatic.com/9vo3dSag_xI4khGkpoWK1HF6hhy/baike/s%3D337/sign=f6f72d232e2dd42a5b0907a8343a5b2f/adaf2edda3cc7cd9accd4c933201213fb90e9193.jpg

https://gss2.bdstatic.com/-fo3dSag_xI4khGkpoWK1HF6hhy/baike/s%3D64/sign=76e0cb4ca4c379317968852debc45138/d52a2834349b033b55a5f5901ece36d3d539bd4e.jpg 的凸组合所表示的点的全体就是以它们为顶点的三角形的边界和内部的点。

n个点https://gss1.bdstatic.com/-vo3dSag_xI4khGkpoWK1HF6hhy/baike/s%3D81/sign=592eeab070cb0a46812286386a63e705/bd315c6034a85edfd86cf66942540923dc547593.jpgZZ 的凸组合全体形成一个凸多边形

**凸函数：**https://gss2.bdstatic.com/9fo3dSag_xI4khGkpoWK1HF6hhy/baike/s%3D292/sign=af32d1ce960a304e5622a7f3e3c9a7c3/962bd40735fae6cd29a40eab0eb30f2443a70fe3.jpg

对于实数集上的凸函数，一般的判别方法是求它的二阶导数，如果其二阶导数在区间上非负，就称为凸函数。

若f为定义在凸集S上的凸函数，则对每一实数c，水平集Sc={x|x∈S，f(x)≤c}是凸集. 凸性在仿射映射下不变：也就是说，如果f（x）是凸函数，那么g(y)=f(Ay+b)也是凸函数。

凸优化中局部最优解同时是整体最优解，且最优解唯一。此凸函数意思为凸向原点有极小值。

以下问题都是凸优化问题，或可以通过改变变量而转化为凸优化问题：

最小二乘、线性规划、线性约束的二次规划 、半正定规划。

**无约束最优化方法：**

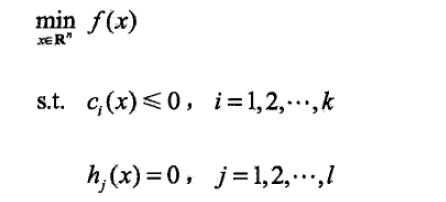
最速下降法，牛顿法

拟牛顿法（目前求解无约束优化问题最有效的方法）

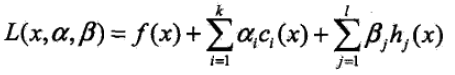
**约束最优化方法**：有约束条件，必须想办法把约束条件去掉才行， 拉格朗日函数

Lagrange对偶问题。

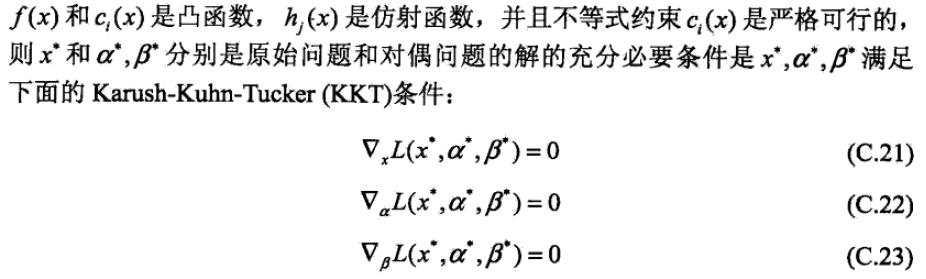
约束优化的一般形式。



引入广义Lagrange函数：

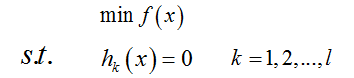


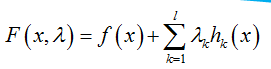
https://upload-images.jianshu.io/upload_images/8740300-1446d9a6776a49cd.png?imageMogr2/auto-orient/strip%7CimageView2/2/w/271



**拉格朗日乘子法：**

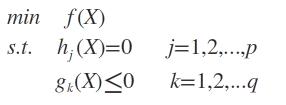
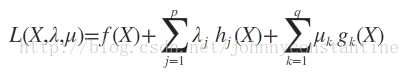
**等式约束条件：**



求解：

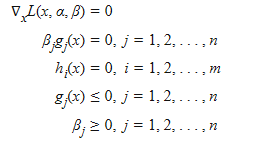


**不等式约束条件：**

引入KTT条件：

KKT条件是说最优值必须满足以下条件：



(1) ：拉格朗日取得可行解的必要条件；

(2) ：这就是以上分析的一个比较有意思的约束，称作松弛互补条件；

(3) ∼

(4) ：初始的约束条件；

(5) ：不等式约束的 Lagrange Multiplier 需满足的条件。

主要的KKT条件便是 (3) 和 (5) ，只要满足这俩个条件便可直接用拉格朗日乘子法，

**拉格朗日对偶：**

目标函数 f(x)会有多种形式：如果目标函数和约束条件都为变量

x的线性函数, 称该问题为线性规划； 如果目标函数为二次函数, 约束条件为线性函数, 称该最优化问题为二次规划; 如果目标函数或者约束条件均为非线性函数, 称该最优化问题为非线性规划。

每个线性规划问题都有一个与之对应的对偶问题，对偶问题有非常良好的性质：

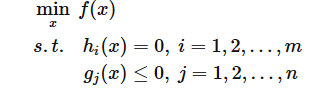
对偶问题的对偶是原问题；

无论原始问题是否是凸的，对偶问题都是凸优化问题；

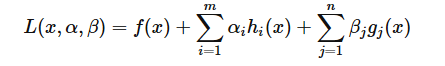
对偶问题可以给出原始问题一个下界；

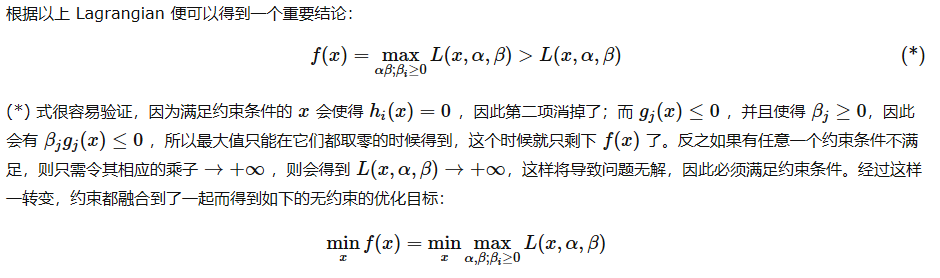
当满足一定条件时，原始问题与对偶问题的解是完全等价的；

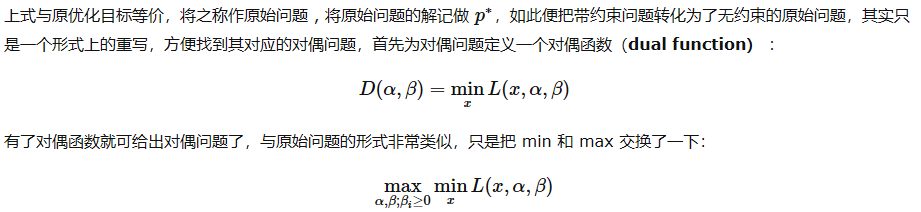
给出不等式约束优化问题：



定义拉格朗日函数：







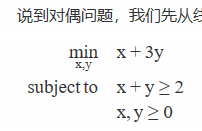
因为在强对偶成立的情况下，可以通过求解对偶问题来得到原始问题的解，在 SVM 中就是这样做的。当然并不是所有的对偶问题都满足强对偶性 ，在 SVM 中是直接假定了强对偶性的成立，其实只要满足一些条件，强对偶性是成立的，比如说 Slater 条件与KKT条件。

当弱对偶成立时，可以得到原始问题的一个下界。而如果强对偶成立，则可以直接求解对偶问题来解决原始问题。

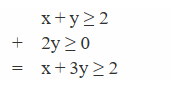
SVM 就是这样的。对偶问题由于性质良好一般比原始问题更容易求解，在 SVM 中通过引入对偶问题可以将问题表示成数据的内积形式从而使得 kernel trick 的应用更加自然）

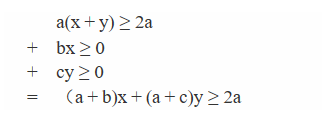
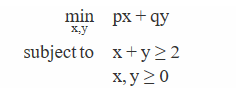
对偶问题：

线性规划：



求解：





即对偶问题：

