# 第五讲　凸规划

* 凸集
* 凸函数
* 凸优化
* 最优化条件
* 对偶原理
* 线性规划Wolfe对偶

本章主要给出凸集和凸函数的定义与性质、凸规划的基本理论，包括最优化条件和对偶定理以及Wolfe对偶。为此首先回忆凸集、凸函数的定义及其基本性质。

其中是凸函数，是凹函数，是线性函数。

## 一、凸集

本节主要介绍凸集的定义及主要性质。

**定义5.1**

设x，y为欧式空间中相异的两个点，则点集

称为通过x和y的直线。若则P为通过x和y的线段。若，P为通过x和y

的放射线**。**

**定义5.2（凸集）**

设S⊆,若对∀及∀λ∈[0,1],都有

则称S为凸集.

设，称

(其中)为的凸组合。

集合S的凸包conv(S)定义为S中所有元素的凸组合构成的集合。

换句话说，若一个集合中任意两个两点的线段仍在这个集合中，则该集合为凸集。

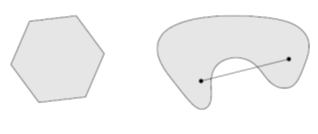


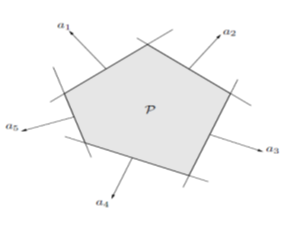
图1

图1中左边的集合为凸集，右边的集合不是凸集。

例5.1

* 平凡凸集:空集、点、线。
* Norm ball: 其中 表示任意一种范数, r为半径。
* 超平面:,其中中的任意向量。
* 半空间:
* 仿射空间:,其中 A,b。
* 多面集 :。

集合 也是多面集。

****

* 单纯形：一种特殊的多面集，表示为conv，其中仿射独立。典型的例子为概率单纯形，即Conv 。

**定义5.3（凸锥）**

设 满足对任意下式成立：

。

则称C为锥。

凸锥是指凸的锥，即满足任意的有下式成立：

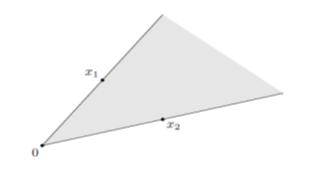


图2

**定义5.4（极点、极方向）**

S是非空凸集，,若由,其中必推出，那么是S的极点。

设S是中的闭凸集，如果对，有

则称向量为S的方向。设为S的方向，若对任意的,有，则称与是两个不同的方向。若S的方向不能表示为该集合的两个不同方向的正的线性组合，则称为S的极方向。

下面给出关于凸集的几个重要结论。

**定理5.1（凸集分离定理）**

任意两个不相交的非空凸集，必存在一个超平面可以分离这两个集合。即，若C，D是两个非空凸集且满足，则存在向量使得

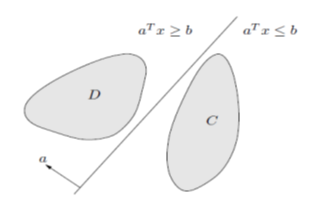


图3

**定理5.2（支撑超平面定理）**

任意一个凸集必存在一个支撑超平面与该凸集相交于凸集边界上一点。即 若C为非空凸集及，存在一个向量使得

.

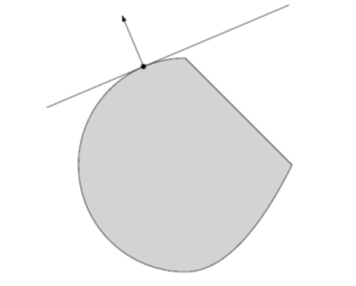


图4

**性质5.1（**凸集上的运算性质）

1、有限个凸集的交集仍为凸集；

2、凸集的简单转换仍为凸集。即若 C 为凸集,则对任意的 ，

仍为凸集。

3、凸集上的仿射变换及逆变换得到的集合仍为凸集。

即若且C 是凸集，则有

为凸集。

若D为凸集，则有

为凸集。

4、 设（）,对任意的，

.

若 是凸集，则P(C)也为凸集。若D 是凸集，则

5、线性分式方程定义为

其中。若为凸集，则为凸集。若D 为凸集，则也为凸集。

**例5.1**

设,则关于这些矩阵的线性矩阵不等式是指

。

对任意给定的，证明满足上述不等式的组成的集合C为凸集。

证明：

方法一（定义法）:

设

对任意的，根据可得

从而有

即C为凸集。

方法二（凸集上的运算性质）:

令。注意到,根据凸集的运算性质3可得C为凸集。 □

**例5.2**

对任意的阶的Fantope是指

，

其中tr(Z)=表示迹运算。 证明：是凸集。

证明：

方法一：（定义法）

设,,则有

,且 ==。则对任意的 ，满足

；

.

即，故有是凸集。

方法二：（凸集上的运算性质）

已知为三个凸集的交集，故为凸集。 □

**例5.3**

令U,V 分别为和上的随机变量， 为U,V联合组成随机变量，对任意的，定义联合分布概率为

。

令 为条件变量,即对任意 定义

。

若C是凸集，证明D是凸集。

证明：定义

则为线性分式方程, 由性质5可得D为凸集。 □

## 二、凸函数

**定义5.5（凸函数）**

设*S*是中的非空凸集， 是定义在*S*上的实函数，如果对于每一对∈*S*及每一个，0≤≤1,都有

()≤+(1－),

则称函数为S上的凸函数。上式中，若≤变为<，则称为严格凸函数。 若-为S的凸函数，则称为S上的凹函数．

*x*

严格凸 凸（凹） 凸

图5

**例 5.4**

（1）、指数函数

（2）、幂函数 对任意的 在

（3）、幂函数在

（4）、对数函数: 在上是凹函数。

（5）、仿射函数：既是凸函数又是凹函数。

（6）、二次函数 当(半正定)时为凸函数。

（7）、最小二乘损失: 是凸函数(因为为半正定矩阵)。

（8）、范数：，范数定义为

。

矩阵的谱范数定义为

矩阵的迹范数定义为

其中 是矩阵X的奇异值。

（9）、示性函数：若 C是凸集，则其示性函数

为凸函数。

（10）、支撑函数: 对任意集合C(凸集或非凸集),支撑函数

为凸函数

（11）、 为凸函数。

下面给出凸函数的重要性质。

**定理5.3**

为凸函数等价于其上图

为凸集。

该定理建立了凸函数和凸集的等价关系，从而可以通过判断一个函数的上图是否为凸集来判定该函数是否为凸集。

**定理5.4**

若为凸函数, 则对任意的，其水平集

为凸集。反之不成立。

**定理5.5（一阶性质）**

若一阶可微, 则为凸函数等价于为凸集且对任意的 ，有

。

从而，对任意可微函数，

为的极小值点。

**定理5.6（二阶性质）**

若一阶可微, 则为凸函数等价于为凸集且对任意的 ，有

。

**Jensen’s inequality:**

若为凸函数， X为上的随机变量，则有

。

**性质5.2（凸函数的性质）**

1、若 为凸函数，则对任意的，

为凸函数。

2、若对任意的, 为凸函数，则

为凸函数，此时S可以为无限集。

3、若对为凸函数且C是凸集，则

为凸函数。

4、若为凸函数，则

为凸函数。

5、设,其中 则有

* 若为凸函数且非减，为凸函数，则为凸函数。
* 若为凸函数且非增，为凹函数，则为凸函数。
* 若为凹函数且非减，为凹函数，则为凸函数。
* 若为凹函数且非增，为凸函数，则为凸函数。

理解上述结论。

6、设

其中。则有：

* 若为凸函数且对每个分量非减，为凸函数，则为凸函数。
* 若为凸函数且对每个分量非增，为凹函数，则为凸函数。
* 若为凹函数且对每个分量非减，为凹函数，则为凸函数。
* 若为凹函数且对每个分量非增，为凸函数，则为凸函数。

**例5.5**

（1）、令C为任意集合，考虑任意范数下的最大距离

。

证明该距离为凸函数。

证明：

对固定的 关于为凸函数。根据性质5.2的2可得为凸函数。

（2）令C为凸集, 考虑最小距离

。

证明该距离为凸函数。

证明：

已知对和为凸函数，由C为凸集及性质5.2的3可知为凸函数。

**例5.6**

对于,

可以近似估计，证明为凸函数。

证明：

结合性质5.2的1和4可得仅需证明

为凸函数。

则

.

该矩阵是对角占优的，因而有半正定，由定理5.6可得为凸函数。 □

## 三、凸规划基本性质

凸规划是指：

其中是凸函数，是凹函数，是线性函数。

记。

* 称为目标函数；
* 为不等式约束函数；
* 为等式约束函数；
* 若 ,且，则称为可行点；
* 设为可行点，若存在使得任意的,

有,则称为局部极小值点。

* 在所有可行点上的最小值称为最优值，记做；
* 若为可行点且,则称为最优解，也成为解或最小值点；
* 若为可行点且，则称为 ；
* 若为可行点且 , 则称在处为积极约束。
* 凸的最小化问题可以转换为凹的最大化问题

**定理5.6**

令表示凸规划（1）的最优解的集合, 则是凸集。

证明：

若，则对任意的

，

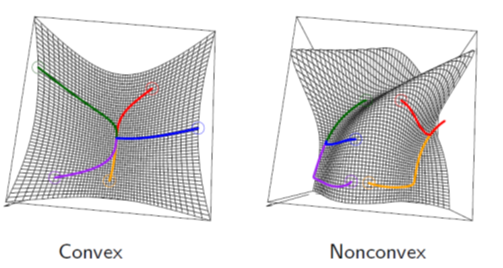
，

。

即。 □

**定理5.7**

若是严格凸函数,则最优解唯一，即仅包含一个元素。



## 四、最优性条件

**定理5.7（一阶充分条件）**

设问题

中，f 是凸函数，( i=1,2,…,m)是凹函数，S为可行域，。f和在点可微， 在点连续，且在处KKT条件成立，则为整体极小点。

优化问题

可以改写为

其中 。进而可以转换为无约束优化问题

其中是集合C上的示性函数。

**定理5.8（一阶最优性条件）**

设为凸可微函数，凸优化问题为

则可行点 为最优解的充要条件为

for all

换言之： 点处所有的可行方向与同向。若，则一阶最优性条件为

**例5.7 （等式约束极小化问题）**

考虑等式约束极小化问题:

其中可微。证明拉格朗日乘子最优性条件为

对某些成立

证明“根据一阶最优性条件，解x满足Ax=b 且对所有的y下式成立

根据，上述结论等价于对所有的

成立。

若f 关于 (x,y)为凸函数，则 关于x为凸函数且C为凸集。

**例5.8**

设,则

（2）

其中 （2）为凸规划当且仅当（1）为凸规划。

**例5.9**

将上述约束改写为。假设上述问题存在最优解。则上述优化问题等价于

其中 称为 hinge 函数。

消除等式约束

给定约束优化问题

常将约束改写为,其中 且 。因此上述优化问题等价于

注：这种方法很常见但在实际问题中不一定适用。

下面介绍新的消除等式约束的方法：引入松弛变量。

给定约束优化问题

我们可以通过引入松弛变量的方法消除等式约束

注： 该问题只有在 为仿射函数才为凸规划。

约束问题的一阶最优性条件——KKT条件

是问题（5.1）的局部最优解 则对所有的,有

即令,得

**定理5.9**

设是问题（1）的局部最优解并设处

(约束品性)

则存在Lagrange乘子向量使得

特别地，若存在处LICQ成立；或函数(i)都是线性函数，则上面的KKT条件成立。

**定理5.10**

设f是凸函数，i都是凹函数，j 都是线性函数，又设在处KKT条件成立则是问题（1）的全局最优解。

注：对凸规划问题，KKT条件也是充分条件！

练习：求约束极值问题

的KKT点。

解：

推论1： 设（NP）是线性约束的凸规划，则是整体最优解是KKT点。

推论2：问题

则是最优解存在

关于线性规划，我们将在下一节详细介绍。

五、**凸规划对偶**

**定理5.11（弱对偶定理）**

令p和d分别表示原问题和对偶问题的最优解，则。

**定理5.11（强对偶定理）**

设在可行域D上为凸函数，且 为仿射函数。若存在 使得

，, j=1,…,r

则p=d。

下面介绍Wolfe 对偶。

凸优化问题：

设f(x) 为凸连续可微函数且 g(x) 为凹连续可微函数。定义

and Equation 1等价于

对任意固定的 , L为x的凸函数且有唯一最优解。对任意固定的x, L 关于为线性函数。

定义

线性方程的最优解组成的区域为凹集，且在线性方程的导数为0时达到最优解。因此 在时有唯一最优解。 即

且函数f的全局最小值或h的全局最大值在下面条件下取得

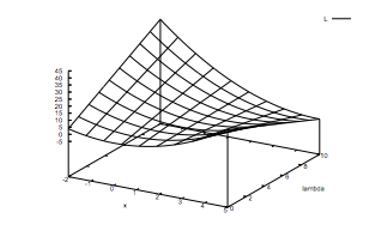
其中最优解在可行域的边界上。若最优解不在可行域的边界，则

等价于

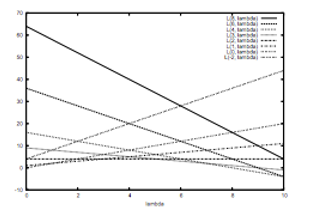
此问题为对偶问题，且最优解等于下列原问题的最优解。

**例5.10**

拉格朗日函数 形状为：



将图形投影到上，线性函数的极小值为凹的。对于这个过程的一个直观的理解就是利用**直线族的包络**，对于每个固定的x，都有一条对应的以lambda为参数的直线，那么所有的x实际上就是一个直线族，这个直线族在每个lambda截面上必定有极大**和**极小值（极值可以是无穷）。这些极值最后组合成的曲线就是包络。而这个包络所对应的极值点（如果有限）应该就是原函数的鞍点（saddle）了.



为了找到函数h, 集合

解得。将 带入 L 可得

令

则，且4为图上凹的部分的最大值点。将 带入原问题可得，且2为原问题可行域的边界点。

线性规划的Wolfe 对偶？

习题：

例：判断下列函数是否为凸函数.

参考文献

1. 《Convex Optimization Algorithm》 Dimitri P. Bertsekas， 2015
2. 《Convex Optimization Theory》Dimitri P. Bertsekas， 2009
3. 《Numerical Optimization》 Jorge Nocedal, Stephen Wright， 1999(2006)
4. 《Convex Optimization》 Stephen Boyd, Lieven Vandenberghe， 2004
5. 《最优化理论与算法》陈宝林，1989(2005)编著
6. J.P. Hiriart-Urruty and C. Lemarechal (1993), \Fundamentals of convex analysis"
7. R. T. Rockafellar (1970), \Convex analysis"