最优化方法

陆 玫

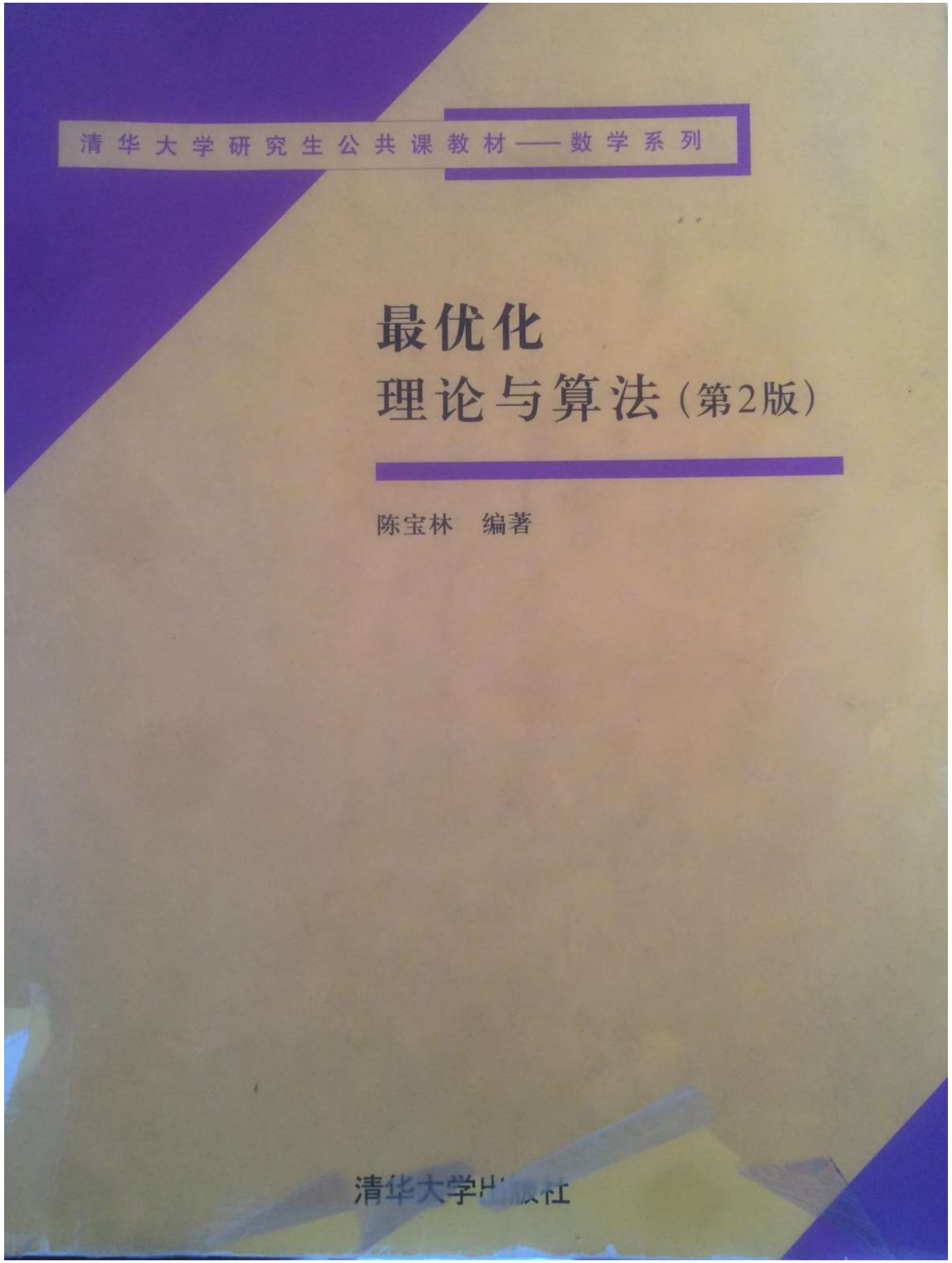
lumei@tsinghua.edu.cn

Tel：62794756

内容：

1. 线性规划
2. 非线性规划

教材



参考书

- 《数学规划》黄红选，韩继业编著
- 《运筹学》＜运筹学＞教材编写组编著
- **Linear and Nonlinear Programming, David G. Luenberger and Yinyu Ye**

作业要求与答疑安排

- 作业要求：写清名字与学号并将作业扫描（或者拍照）后在网络学堂上提交(电子版请提交pdf版本)。
- 每周四晚上12点之前提交作业。
- 答疑时间：每周三下午3：00--4：00
- 答疑地点：理科楼A414
- 通过邮箱(lumei@tsinghua.edu.cn)答疑
- 总成绩=平时成绩(20%)+期末考试成绩(75%)
- 出勤：5分

- 助教：

  田静(jingtian_xju@163.com)

  邵钰莫(syg20@mails.tsinghua.edu.cn)

  李佳明(lijiaminggood@163.com)

  吕泽群(2312231981@qq.com)

  郑瑜(zhengyu.davy@foxmail.com)

最优化问题

最优化：在一定约束条件下极大化或极小化某函数. 在所有可行方案中选取最佳方案，使效益最大或成本最低。

依照给定条件和目标，从众多方案中选择最佳方案

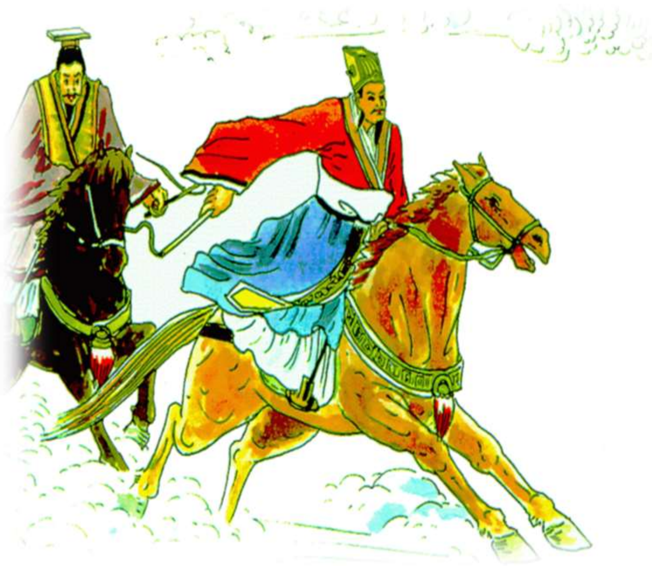
应　用

欧拉：宇宙万物无不与最小化或最大化有关

- **经济金融**：投资组合
- **交通运输**：列车运行、物流运输
- **信息科学**：数据挖掘、图像/信号处理、机器学习、强化学习、压缩感知、模式识别
- **生命医学**：医学图像、DNA序列、蛋白质折叠
- **军事运筹**：摆兵布阵、后勤保障

历史

最优化的起源：春秋战国时期的田忌赛马



	齐威王		田忌	齐威王		田忌
上	A	>	B	A	>	F
中	C	>	D	C	<	B
下	E	>	F	E	<	D
	3	:	0	1	:	2

18世纪，欧拉、拉格朗日、高斯等对与力学、天文学中极值问题的研究

欧拉

拉格朗日

高斯

第一次世界大战

- 最早进行运筹学工作的，是以英国生理学家希尔为首的英国国防部防空试验小组，他们在第一次世界大战期间进行了高射炮系统利用研究

雷达与“Blackett马戏团”

- 1935年，英国科学家沃森--瓦特(R. Watson-Wart〉发明了雷达. 但很快发现由这些雷达得到的信息常常是互相矛盾的，需要加以协调和关联.
- 1939年组建了代号为“Blackett马戏团”的研究小组，专门就改进防空系统进行研究，大大提高了英国本土的防空能力
- “Blackett马戏团”是世界上第一个使用了“Operational Research”一词；从学术思想上，他们的研究已经蕴含着整体性的概念和系统分析的思想，这也是运筹学的精髓

- 第二次世界大战

- 深水炸弹的起爆深度
- **1942**年, 麻省理工学院的物理学家莫尔斯领导的小组经过调查研究,提出两条重要建议:
(l)将反潜攻击由反潜舰艇投掷水雷改为由飞机投掷深水炸弹;且仅当潜艇浮出水面或刚下潜时,才投掷深水炸弹:炸弹的起爆深度由原来的水下**100**米左右改为水下**25**米左右.
(2)改进运送物资的船队及护航舰艇编队的方式,由小规模多批次,改进为加大规模, 减少批次,可使损失减少.

线性规划—第二次世界大战的后勤供应问题

$$
\begin{aligned}
\min\ & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} \\
\text{s.t. } & \boldsymbol{A}\boldsymbol{x} \geqslant \boldsymbol{b}
\end{aligned}
$$

线性规划单纯形方法的创始人 Dantzig

- 第二次世界大战结束时，英、美及加拿大军队中工作的运筹学者已超过**700**人，正是由于战争的需要，运筹学有了长足的发展，并且形成了一门学科。

- 用最优化方法解决实际问题，一般可经过下列步骤：
- ①提出最优化问题，收集有关数据和资料；
- ②建立最优化问题的数学模型,确定变量,列出目标函数和约束条件；
- ③分析模型，选择合适的最优化方法；
- ④求解，一般通过编制程序，用计算机求最优解；
- ⑤最优解的检验和实施。
- 上述 5个步骤中的工作相互支持和相互制约，在实践中常常是反复交叉进行。

生产计划的编制

资源 \ 产品	A	B	C	企业现有资源
钢材(吨/只)	3	4	2	600
木材（立方米/只）	2	1	2	400
人工（千小时/只）	1	3	3	300
机床(台/只)	1	4	4	200
收益（千元/只）	2	4	3	

问：企业应如何安排生产，能使总收益最大？

2、数学模型

- 决策目标：**A**、**B**、**C** 产品各生产多少台使企业总收益最大？

• 决策变量：设 x_1, x_2, x_3 为生产 A, B, C 三种产品的数量。

• 目标函数：$\max \ 2x_1 + 4x_2 + 3x_3$

• 约束条件：

$$
\begin{aligned}
3x_1 + 4x_2 + 2x_3 &\le 600 \\
2x_1 + x_2 + 2x_3 &\le 400 \\
x_1 + 3x_2 + 3x_3 &\le 300 \\
x_1 + 4x_2 + 4x_3 &\le 200
\end{aligned}
$$

•非负条件：$x_1, x_2, x_3 \ge 0$

$$
\begin{cases}
\max & 2x_1 + 4x_2 + 3x_3 \\
s.t. & 3x_1 + 4x_2 + 2x_3 \le 600 \\
 & 2x_1 + \quad x_2 + 3x_3 \le 400 \\
 & x_1 + 3x_2 + 3x_3 \le 300 \\
 & x_1 + 4x_2 + 4x_3 \le 200 \\
 & x_1, x_2, x_3 \ge 0
\end{cases}
$$

运输问题

如某建材公司有三个水泥厂 A_1，A_2，A_3 ，四个经销商 B_1，B_2，B_3，B_4 其产量、销量、运费（元/吨）见表**2．5**。

表 2.5 建材公司的数据

产地 \ 销售地	B_1	B_2	B_3	B_4	产量（吨）
A_1	8	7	3	2	2000
A_2	4	7	5	1	10000
A_3	2	4	9	6	4000
销量（吨）	3000	2000	4000	5000	

如何制定调运方案，使总的运费最小?

数学模型----线性规划问题

$$
\begin{aligned}
\min f =\ & 8x_{11} + 7x_{12} + 3x_{13} + 2x_{14} + 4x_{21} + 7x_{22} \\
& + 5x_{23} + x_{24} + 2x_{31} + 4x_{32} + 9x_{33} + 6x_{34}
\end{aligned}
$$

$$
s.t. \quad
\begin{aligned}
& x_{11} + x_{12} + x_{13} + x_{14} \le 2000 \\
& x_{21} + x_{22} + x_{23} + x_{24} \le 10000 \\
& x_{31} + x_{32} + x_{33} + x_{34} \le 4000 \\
& x_{11} + x_{21} + x_{31} \ge 3000 \\
& x_{12} + x_{22} + x_{32} \ge 2000 \\
& x_{13} + x_{23} + x_{33} \ge 4000 \\
& x_{14} + x_{24} + x_{34} \ge 5000
\end{aligned}
$$

$$x_{ij} \ge 0, i = 1,\ 2,\ 3;\ \ j = 1,\ 2,\ 3,\ 4$$

由生产基地 $A_i\ (i = 1,\ 2,\ 3)$

运到销售地 $B_j\ (j = 1,\ 2,\ 3,\ 4)$

的货运量为 x_{ij}

问题的解决方案

其解为 $x=(0,0,2000,0,1000,0,2000,5000,2000,2000,0,0)^T$， $\min f=37000$ 元。最佳运输方案见：表**2．6**

表 2.6　最佳运输方案

产地＼销售地	B_1	B_2	B_3	B_4	产量（吨）
A_1	0	0	2000	0	2000
A_2	1000	0	2000	5000	10000
A_3	2000	2000	0	0	4000
销量（吨）	3000	2000	4000	5000	

合理下料问题

现有一批长度一定的原材料钢管，由于生产的需要，要求截出不同规格的钢管若干。

试问应如何下料，既能满足生产的需要，又使得使用的原材料钢管数量最少（即废材最少）？

具体问题：料长**7.4m**，要求截成**2.9m**，**2.lm**，**1.5m**的钢管分别为**1000**根，**2000**根，**1000**根。如何截取，才使得总用料最省?

Modeling

把所有可能的下料方式、按照各种下料方式从料长7．4m的原料上得到的不同规格钢管的根数、残料长度,以及需要量列于表2．8中。

表 2.8

钢管规格（m）\ 下料方式	B_1	B_2	B_3	B_4	B_5	B_6	B_7	B_8	需要量(根)
2.9	2	1	1	1	0	0	0	0	1000
2.1	0	0	2	1	2	1	3	0	2000
1.5	1	3	0	1	2	3	0	4	1000
残料长度（m）	0.1	0	0.3	0.9	0.2	0.8	1.1	1.4	

问题转化为确定每种下料方式各用多少根7．4m的原料。

数学模型----整数线性规划问题

设 x_1，x_2，$\cdots$，x_8 分别为按照 B_1，$B_2,\ldots,B_8$ 方式下料的原料根数，

则 $\min f = x_1 + x_2 + \cdots + x_8$

$$
\begin{aligned}
s.t.\quad & 2x_1 + x_2 + x_3 + x_4 && \geq 1000 \\
& 2x_3 + x_4 + 2x_5 + x_6 + 3x_7 && \geq 2000 \\
& x_1 + 3x_2 \quad x_4 + 2x_5 + 3x_6 \quad + 4x_8 && \geq 1000 \\
& x_j \geq 0 \quad (j = 1,\ 2,\ \cdots,\ 8)\text{且为整数}
\end{aligned}
$$

其解为 $x = (0, 200, 800, 0, 200, 0, 0, 0)^T$，$\min f = 1200$ （根）

最佳下料方案为：方式 $B_2: 200$ 根，方式 $B_3: 800$ 根，方式 $B_5: 200$ 根。

人力资源安排问题

某商场是个中型的百货商场，现在需要对营业员的工作时间作出安排，营业员每周工作五天，休息两天，并要求休息的两天是连续的，问题归结为：如何安排营业员的作息时间，既能满足工作需要，又使配备的营业员人数最少？

1、有关数据

对营业员的需求进行统计分析，营业员每天的需求人数如下表所示：

时　　间	所需营业员人数
星期日	**28** 人
星期一	**15** 人
星期二	**24** 人
星期三	**25** 人
星期四	**19** 人
星期五	**31** 人
星期六	**28** 人

2、模型

决策变量：设x_j为第j天开始休息的人数$(j=1,2,\cdots,7)$

目标函数： $\min x_1+x_2+x_3+x_4+x_5+x_6+x_7$

约束条件：

$$
\begin{aligned}
& x_1+x_2+x_3+x_4+x_5 \geq 28 \\
& x_2+x_3+x_4+x_5+x_6 \geq 15 \\
& x_3+x_4+x_5+x_6+x_7 \geq 24 \\
& x_4+x_5+x_6+x_7+x_1 \geq 25 \\
& x_5+x_6+x_7+x_1+x_2 \geq 19 \\
& x_6+x_7+x_1+x_2+x_3 \geq 31 \\
& x_7+x_1+x_2+x_3+x_4 \geq 28 \\
& x_1,x_2,x_3,x_4,x_5,x_6,x_7 \geq 0, \text{整数}
\end{aligned}
$$

- 例（挑选球员问题）某篮球教练要从8名业余队员中挑选3名队员参加专业球队，使平均身高达到最高。队员的号码、身高及所擅长的位置如下。要求：中锋1人，后卫1人，前锋1人，但1号、3号与6号队员中至少保留1人给业余队。

号码	身高（米）	位置	挑选变量
1	1.92	中锋	x_1
2	1.91	中锋	x_2
3	1.90	前锋	x_3
4	1.86	前锋	x_4
5	1.85	前锋	x_5
6	1.83	后卫	x_6
7	1.80	后卫	x_7
8	1.79	后卫	x_8

$$
\begin{aligned}
\max\ & 1.92x_1 + 1.91x_2 + 1.90x_3 + 1.86x_4 \\
& + 1.85x_5 + 1.83x_6 + 1.80x_7 + 1.79x_8 \\
s.t.\quad & x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 = 3 \\
& x_1 + x_2 = 1 \\
& x_3 + x_4 + x_5 = 1 \\
& x_6 + x_7 + x_8 = 1 \\
& x_1 + x_3 + x_6 \le 2 \\
& x_i = 0 \text{ 或 } 1 \quad i = 1, 2, \cdots, 8
\end{aligned}
$$

- 例（选址问题）设有n个市场，第j个市场的位置为(a_j,b_j)，对某种货物的需要量为q_j, j=1,…,n,现计划建立m个仓库，第i个仓库的容量为c_i,i=1,…,m,试确定仓库的位置，使各仓库到各市场的运输量与路程乘积之和最小.
- 解：设第i个仓库的位置为(x_i,y_i)，运输量为w_{ij}.

$$
\begin{cases}
\min \sum_{j=1}^{n}\sum_{i=1}^{m} w_{ij}\sqrt{(x_i-a_j)^2+(y_i-b_j)^2} \\
s.t. \quad \sum_{j=1}^{n} w_{ij} \le c_i \quad i=1,2,\cdots,m \\
\qquad \sum_{i=1}^{m} w_{ij} = q_j \quad j=1,2,\cdots,n \\
\qquad w_{ij} \ge 0 \quad i=1,2,\cdots,m \ j=1,2,\cdots,n
\end{cases}
$$

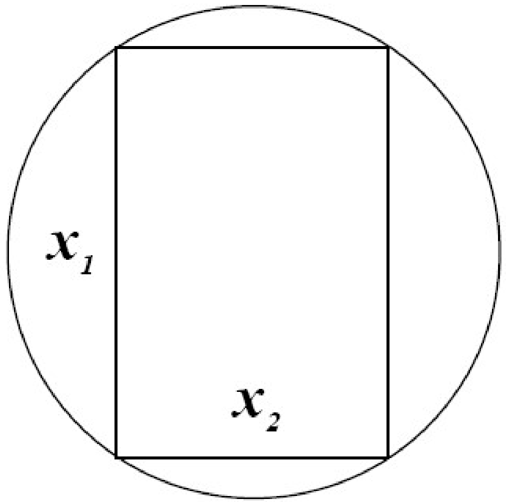
- 例（数据拟合问题）在实验数据处理或统计资料分析中常遇到如下问题：设两个变量x和y，已知存在函数关系，但其解析表达式或者是未知的、或者虽然为已知的但过于复杂。设已取得一组数据，
- $(x_i, y_i),\quad i=1,2,\ldots,m$
- 根据这组数据导出函数$y=f(x)$的一个简单而近似的解析表达式。
- 取一个简单的函数序列$g_0(x), g_1(x),\ldots,g_n(x)$

$$\min\sum_{i=1}^{m}\left[y_i-\sum_{j=1}^{n}\alpha_j g_j\left(x_i\right)\right]^2$$

例：把圆形木材加工成矩形横截面的木梁，要求木梁高度不超过H，横截面的惯性矩(高度^2 ×宽度)不小于W，而且高度介于宽度与4倍宽度之间，问如何确定木梁尺寸可使木梁成本最小。

设矩形横截面的高度为x_1，宽度为x_2

$$
\begin{aligned}
\min \quad & \pi\left(\frac{1}{4}x_1^2+\frac{1}{4}x_2^2\right) \\
s.t. \quad & x_1 \le H \\
& x_1^2 x_2 \ge W \\
& x_2 \le x_1 \le 4x_2 \\
& x_1 \ge 0, x_2 \ge 0
\end{aligned}
$$



模型与分类

数学模型

$$\min \quad f(\boldsymbol{x})$$
$$\text{s.t.} \quad \boldsymbol{x} \in \Omega$$

目标函数

可行域/决策集

subject to 缩写

决策变量

$$\Omega = \{\boldsymbol{x} \in \mathbb{R}^n \mid c_i(\boldsymbol{x}) = 0,\ i \in \mathcal{E};\quad c_i(\boldsymbol{x}) \geqslant 0,\ i \in \mathcal{I}\}$$

等式约束

等式约束指标

不等式约束

不等式约束指标集

1、根据约束划分

无约束优化 $\min\limits_{x \in R^n} \ f(x)$

约束优化
$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & x \in \Omega \end{array}$$

约束优化问题一般比无约束优化问题难解；有时可以相互转化

$$\min_{\boldsymbol{x} \in \mathbb{R}^n} \{\boldsymbol{x}^{\mathrm{T}} \boldsymbol{A} \boldsymbol{x} \mid \boldsymbol{x}^{\mathrm{T}} \boldsymbol{x} = 1\} \Longleftrightarrow \min_{\boldsymbol{x} \in \mathbb{R}^n} \frac{\boldsymbol{x}^{\mathrm{T}} \boldsymbol{A} \boldsymbol{x}}{\boldsymbol{x}^{\mathrm{T}} \boldsymbol{x}}$$

$$\min_{\boldsymbol{x} \in \mathbb{R}^n} f(\boldsymbol{x}) \Longleftrightarrow \min_{\boldsymbol{x}, t} \{t \mid t - f(\boldsymbol{x}) \geqslant 0\}$$

2、根据函数的线性度划分

线性规划：目标函数及约束函数都是线性

$$\begin{aligned} \min \quad & c^{\top} x \\ \text{s.t.} \quad & Ax \geqslant b \end{aligned}$$

非线性规划：目标函数或约束中含有非线性函数

$$\min \; -3x_1 + x_2 - x_3^2$$

$$\text{s.t.} \; x_1 + x_2 + x_3 \leqslant 0$$

$$-x_1 + 2x_2 + x_3^2 = 0$$

一类常用的非线性规划

- **二次规划：** 目标函数二次，约束函数线性

$$
\begin{aligned}
\min\ Q(\boldsymbol{x}) &= \frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{x} + \boldsymbol{g}^{\mathrm{T}}\boldsymbol{x} \\
\text{s.t.}\ \boldsymbol{a}_i^{\mathrm{T}}\boldsymbol{x} &= b_i,\ i \in \mathcal{E} \\
\boldsymbol{a}_i^{\mathrm{T}}\boldsymbol{x} &\geqslant b_i,\ i \in \mathcal{I}
\end{aligned}
$$

线性规划、二次规划问题是常见的、重要的最优化问题。

目前已有比较完善的理论和求解方法。

3、根据目标函数和可行域的凸性划分：

凸规划： 目标函数为凸函数，可行域为闭凸集。

凸函数

$$\begin{aligned} \min \quad & f(\boldsymbol{x}) \\ \text{s.t.} \quad & \boldsymbol{x} \in \Omega \end{aligned}$$

闭凸集

非凸优化： 目标函数非凸或可行域非凸

凸规划问题的局部最优解就是全局最优解，求解相对容易。

非凸优化问题难于求解，特别是全局最优解。

4、根据函数的解析性质划分

光滑优化：优化问题的所有函数都连续可微, 如多项式优化

$$\min \left(x_1 - \frac{4}{9}\right)^2 + (x_2 - 2)^2$$
$$\text{s.t.} \ -x_1^2 + x_2 \geqslant 0$$
$$x_1 + x_2 \leqslant 6$$
$$x_1, x_2 \geqslant 0$$

非光滑优化：优化问题含有不可微函数

$$\min \ \frac{1}{2}\|Ax - b\|^2 + \|x\|_1$$

非光滑优化

非光滑项

5、根据可行域中可行点的个数划分

连续优化： 可行域含有无穷多个不可数的点且可行域中的点连续变化。

离散优化： 可行域中含有有限个或可数个点, 即该优化问题在由有限个点或可数个点组成的可行域中寻求最优解。

多数情况下, 该问题的可行域是通过某些元素的排列组合产生的. 因此又称**组合优化**。

典型案例： 装箱问题、 旅行商问题(销售商问题)、中国邮路问题。

有多个结构相同、大小相等的箱子及多个不同大小的物品。寻找一种方法，使得用最少的箱子将全部物品装入箱内，或用一个箱子装入尽可能多的物品.

给定多个城市和每两城市间的距离，寻求一条最短回路，使得每座城市访问一次并最终回到起始城市

给定多个街道/村庄及每两街道间的距离。寻求一条最短回路，使走遍所有街道，并最终回到邮局。
(中国数学家管梅谷1950s)

根据变量的特殊要求，离散优化又分离出

整数规划：　所有变量都取整数（离散优化）

0-1规划：　所有变量取0或1（离散优化）

混合整数规划:部分变量为整数变量,其余变量为连续变量

混合0-1规划：部分变量取0或1,其余变量为连续变量

稀疏优化：要求最优解稀疏，即非零个数尽可能少

华罗庚先生的统筹法：一种安排工序的组合方法

泡茶喝　水壶要洗，茶壶，茶杯要洗；火已生了，需要烧水，茶叶有了。先干什么，后干什么？才能在最短的时间喝上茶？

办法1：洗水壶，灌凉水，放在火上；在等待水开的时间里，洗茶壶、洗茶杯、拿茶叶；等水开了，泡茶喝。

办法2：先做好一些准备工作，洗水壶，洗茶壶茶杯，拿茶叶；一切就绪，灌水烧水；坐待水开了泡茶喝。

办法3：洗净水壶，灌上水，放在火上，坐待水开；水开之后，找茶叶，洗茶壶茶杯，泡茶喝。

第一种办法最省时间，后两种办法都窝了工。这就是组合优化中的排序问题。

基本概念

- 可行点（可行解）：在线性规划和非线性规划中，满足约束条件的点.
- 可行集或可行域**S**：全体可行点组成的集合.
- 无约束问题：如果一个问题的可行集是整个空间.

- 对于一个规划问题，下面三种情况必占其一：
- **(1)** $S = \Phi$，则称该问题无解或不可行；
- (2) $S \neq \Phi$，但目标函数在S上无界，则称该问题无界；
- (3) $S \neq \Phi$且目标函数有有限的最优值，则称该问题有最优解.

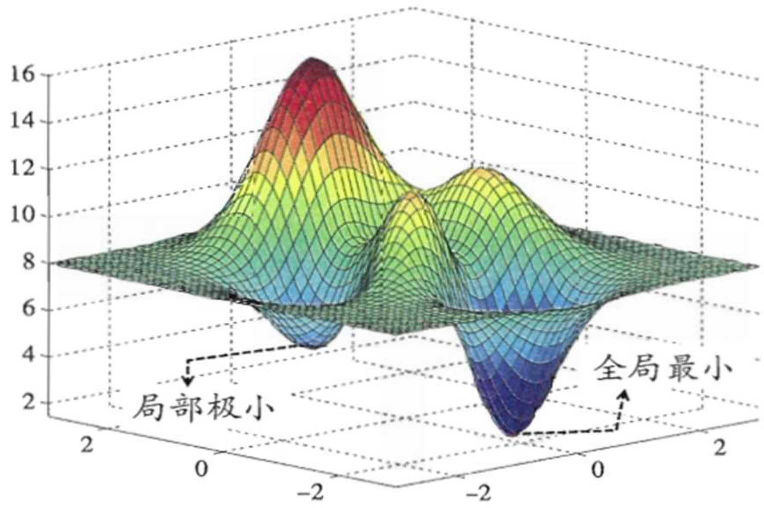
$$\begin{aligned} \min \quad & f(\boldsymbol{x}) \\ \text{s.t.} \quad & \boldsymbol{x} \in \Omega \end{aligned}$$

可行解:　　$\boldsymbol{x} \in \Omega$

最优解:$\boldsymbol{x}^* \in \Omega$.对任意的 $\boldsymbol{x} \in \Omega$, 都有 $f(\boldsymbol{x}^*) \leqslant f(\boldsymbol{x})$

局部最优解： $\boldsymbol{x}^* \in \Omega$. 存在该点的邻域 $N(\boldsymbol{x}^*, \delta)$, 使对任意的 $\boldsymbol{x} \in N(\boldsymbol{x}^*, \delta) \cap \Omega$, 都有 $f(\boldsymbol{x}^*) \leqslant f(\boldsymbol{x})$。

严格局部最优解： 若 $\boldsymbol{x}^*$ 还满足对任意的 $\boldsymbol{x} \in N(\boldsymbol{x}^*, \delta) \cap \Omega$, $\boldsymbol{x} \neq \boldsymbol{x}^*$, 都有 $f(\boldsymbol{x}^*) < f(\boldsymbol{x})$



最优值：优化问题的目标函数在最优解处的值

若优化问题在可行域上有下界，但没有最优解，此时称目标函数在可行域上的**下确界**为优化问题的最优值。

二元函数 $f(x) = x_1^2 + (1 - x_1 x_2)^2$ 在二维空间中的最优值为0，但只在 $x_1 = 1/x_2$ 且 $x_2 \to \infty$ 时才能达到。

考虑到最优值可达和不可达的情况，有时将最优值记为

$$f^* = \inf_{x \in \Omega} f(x)$$

预备知识

线性相关与线性无关：

设V为向量空间，$v^1, v^2, \cdots, v^k \in V$,若存在不全为零的数$\lambda_1, \lambda_2, \cdots, \lambda_k$使得

$$\lambda_1 v^1 + \lambda_2 v^2 + \cdots + \lambda_k v^k = 0$$

则称$v^1, v^2, \cdots, v^k$为线性相关的向量组，否则称为线性无关的向量组。

范数

若函数$\|\cdot\|$：$\Re^n \to \Re$满足下面条件：

（1）正定型：$\forall x \in \Re^n, \|x\| \geq 0$, 并且$\|x\| = 0 \Leftrightarrow x = 0$;

（2）三角不等式：$\forall x, y \in \Re^n, \|x+y\| \leq \|x\| + \|y\|$;

(3) 齐次性：$\forall \alpha \in \Re, \forall x \in \Re^n, \|\alpha x\| = |\alpha| \|x\|$.

则称$\|\cdot\|$为$\Re^n$上的范数.

常用范数：

$$\|x\|_2 = \sqrt{\sum_{i=1}^{n} x_i^2} \qquad \|x\|_1 = \sum_{i=1}^{n} |x_i| \qquad \|x\|_\infty = \max_{1 \leq i \leq n} |x_i|$$

序列的极限

- 定义：设$\{x^{(k)}\}$是R^n中的一个向量序列，$\bar{x}\in R^n$,如果对任给的$\varepsilon>0$存在正整数K_ε，使得当$k>K_\varepsilon$时就有$\|x^{(k)}-\bar{x}\|<\varepsilon$,则称序列收敛到$\bar{x}$或称序列以$\bar{x}$为极限，记作$\lim\limits_{k\to\infty} x^{(k)}=\bar{x}$.
- 结论：序列若存在极限，则任何子序列有相同的极限，即序列极限是唯一的。
- 定义：设$\{x^{(k)}\}$是R^n中的一个向量序列，如果存在一个子序列$\left\{x^{(k_j)}\right\}$，使得$\lim\limits_{k\to\infty} x^{(k_j)}=\hat{x}$，则称$\hat{x}$是序列$\{x^{(k)}\}$的一个聚点。

- 定义：设$\{x^{(k)}\}$是R^n中的一个向量序列， 如果对任给定的$\varepsilon>0$，总存在正整数K_ε，使得当m, $l>K_\varepsilon$时就有$\left\|x^{(m)}-x^{(k)}\right\|<\varepsilon$，则称序列$\{x^{(k)}\}$为**Cauchy**序列。

- 定理：设 $\left\{x^{(j)}\right\}\subset R^n$ 为**Cauchy**序列，则 $\left\{x^{(j)}\right\}$ 的聚点必为极限点。

集合

x^0的$\varepsilon-$邻域：$N_\varepsilon(x^0)=\{x \mid \| x-x^0 \|<\varepsilon, \varepsilon>0\}$

内点： 设$x^0 \in S \subset \Re^n$，若存在$\varepsilon>0$，使得$N_\varepsilon(x^0)\subset S$，

则称x^0为S的一个内点。

补集： 集合S的补集定义为$S^C=\{x \mid x \notin S, x \in \Re^n\}$

开集： 若对$\forall x \in S, x$为内点，则称S为开集。

闭集： 若集合S的补集S^C为开集，则称S为闭集。

有界集 ：若存在正数$M>0$，使得$\forall x \in S, \| x \| \leq M$成立，

则称S为有界集。

紧集： 有界闭集称为紧集.

性质：

（1）集合$S \subset \Re^n$是闭集，当且仅当对任意的无穷序列$\{x^k\} \subset S$，若$x^k \to x^*$，则$x^* \in S$。

（2）集合$S \subset \Re^n$是紧集当且仅当对任意的无穷序列$\{x^k\} \subset S$，必存在收敛于S中点的子序列$\{x^{k_i}\}$.

函数的展开

梯度：$$\nabla f(x) = \left(\frac{\partial f(x)}{\partial x_1}, \cdots, \frac{\partial f(x)}{\partial x_n} \right)^T$$

Hesse矩阵：

$$\nabla^2 f(x) = \begin{bmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \cdots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(x)}{\partial x_2 \partial x_1} & \cdots & \frac{\partial^2 f(x)}{\partial x_2 \partial x_n} \\ \cdots\ \cdots\ \cdots\ \cdots & & \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f(x)}{\partial x_n^2} \end{bmatrix}$$

Taylor展开

定理：

设函数$f:\Re^n \to \Re$连续可微，向量$p \in \Re^n$，则

$$f(x+p) = f(x) + \nabla f(x)^T p + o(\| p \|)。$$

若函数f是二阶连续可微，则

$$f(x+p) = f(x) + \nabla f(x)^T p + \frac{1}{2} p^T \nabla^2 f(x) p + o(\| p \|^2).$$

二次型的正定性

定义：对实二次型$f(X)=X^TAX$，若$\forall X\neq 0$，都有$f(X)=X^TAX>0$成立，则称$f(X)$为正定二次型，A为正定矩阵。

定理：对于n阶实对称矩阵A，下列命题等价：

（1）X^TAX是正定二次型（或A是正定矩阵）；

（2）A的n个顺序主子式都大于零；

（3）A的n个特征值都大于零；

（4）存在可逆矩阵P，使得$A=P^TP$.

二次型的半正定性

定义：对实二次型$f(X)=X^TAX$，若$\forall X\neq 0$，都有$f(X)=X^TAX\geq 0$成立，且存在$X\neq 0$使得$f(X)=X^TAX=0$，则称$f(X)$为半正定二次型，A为半正定矩阵。

定理：对于n阶实对称矩阵A，下列命题等价：

（1）X^TAX是半正定二次型（或A是半正定矩阵）；

（2）A的所有主子式（行号与列号取成相同的子式）都大于等于零，而且至少有一个等于零；

（3）A的n个特征值都大于等于零，而且至少有一个等于零。

凸集(convex set)

- 定义：设x,y为欧式空间E^n中相异的两个点，则点集
- $$P=\{\lambda x+(1-\lambda)y|\lambda\in R\}$$
- 称为通过x和y的直线。
- 定义：设$S\subseteq E^n$, **若对**$\forall x^{(1)}, x^{(2)}\in S$**及**$\forall\lambda\in[0,1]$, **都有**
- $$\lambda x^{(1)}+(1-\lambda)x^{(2)}\in S$$
- **则称S为凸集。**
- **设**$x^{(1)}, x^{(2)}, \cdots, x^{(k)}\in S$, **称**
- $$\lambda_1 x^{(1)}+\lambda_2 x^{(2)}+\ldots+\lambda_k x^{(k)}$$
- **(其中**$\lambda_i\geq 0$, $\lambda_1+\lambda_2+\cdots+\lambda_k=1$) **为**$x^{(1)}, x^{(2)}, \cdots, x^{(k)}$**的凸组合**.

- **H={$x|p^{\mathrm{T}}x = a$}** ——————超平面

- **H⁻={$x|p^{\mathrm{T}}x \leqslant a$}** ——————（闭）半空间

- L={$\boldsymbol{x} \mid \boldsymbol{x}=\boldsymbol{x}^{(0)}+\lambda\,\boldsymbol{d},\ \lambda \geqslant 0$}————射线

凸集的性质

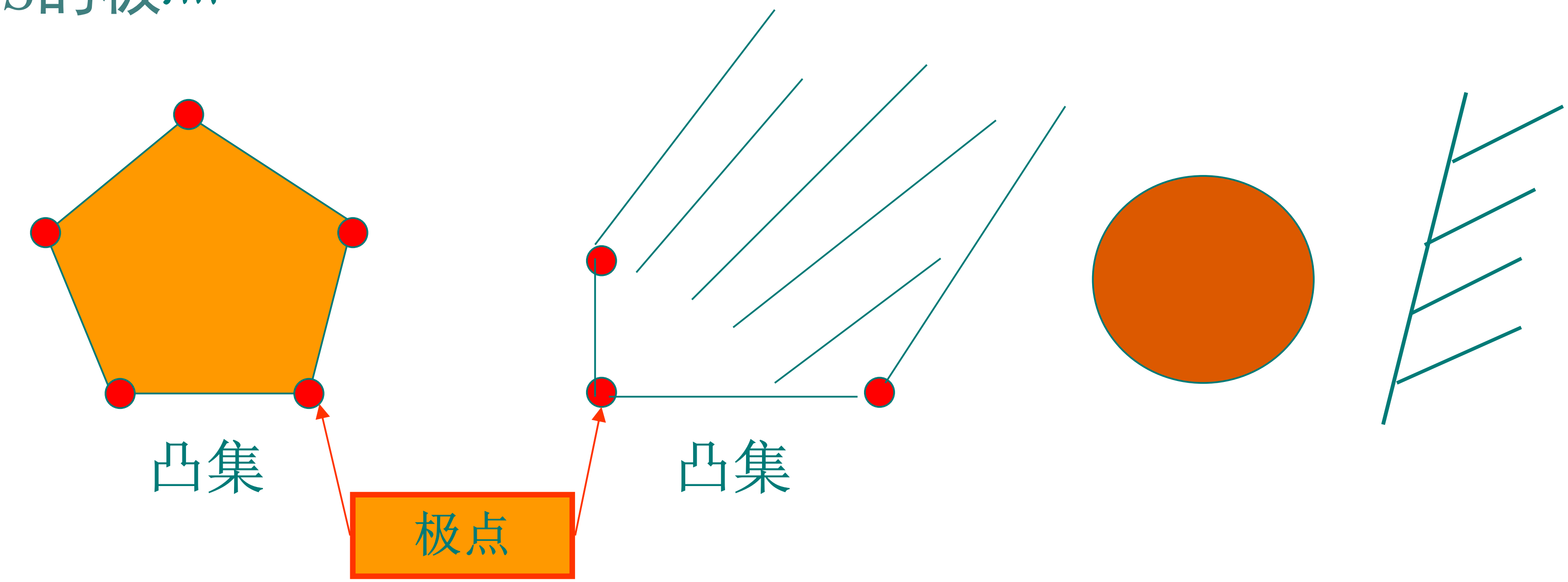
- 设S_1和S_2为E^n中的两个凸集，β是实数，则
- (1) $\beta S_1=\{\beta x|x\in S_1\}$为凸集。
- (2) $S_1\cap S_2$为凸集。
- (3) $S_1+S_2=\{x^{(1)}+x^{(2)}|x^{(1)}\in S_1,\ x^{(2)}\in S_2\}$为凸集。
- (4) $S_1-S_2=\{x^{(1)}-x^{(2)}|x^{(1)}\in S_1,\ x^{(2)}\in S_2\}$为凸集。

定义：有限个闭半空间的交$\{x\mid Ax\le b\}$称为多面体。

结论：多面体是凸集

极点(extreme point)

定义： 设S是非空凸集，$x \in S$，若由$x = \lambda x^{(1)} + (1-\lambda)x^{(2)}$，其中$\lambda \in (0,1)$, $x^{(1)}, x^{(2)} \in S$，必推出$x = x^{(1)} = x^{(2)}$，则称 x是S的极点.



极方向(extreme direction)

定义:设S为E^n中的闭凸集，$d \in E^n, d \neq 0$，如果对$\forall x \in S$，有

$$\{x+\lambda d \mid \lambda \geq 0\} \subset S$$

则称向量d为S的方向。设$d^{(1)}, d^{(2)}$为S的方向，若对任意的$\lambda > 0$，有$d^{(1)} \neq \lambda d^{(2)}$,则称$d^{(1)}$与$d^{(2)}$是两个不同的方向。若$S$的方向$d$不能表示为该集合的两个不同方向的正的线性组合，则称d为S的极方向。

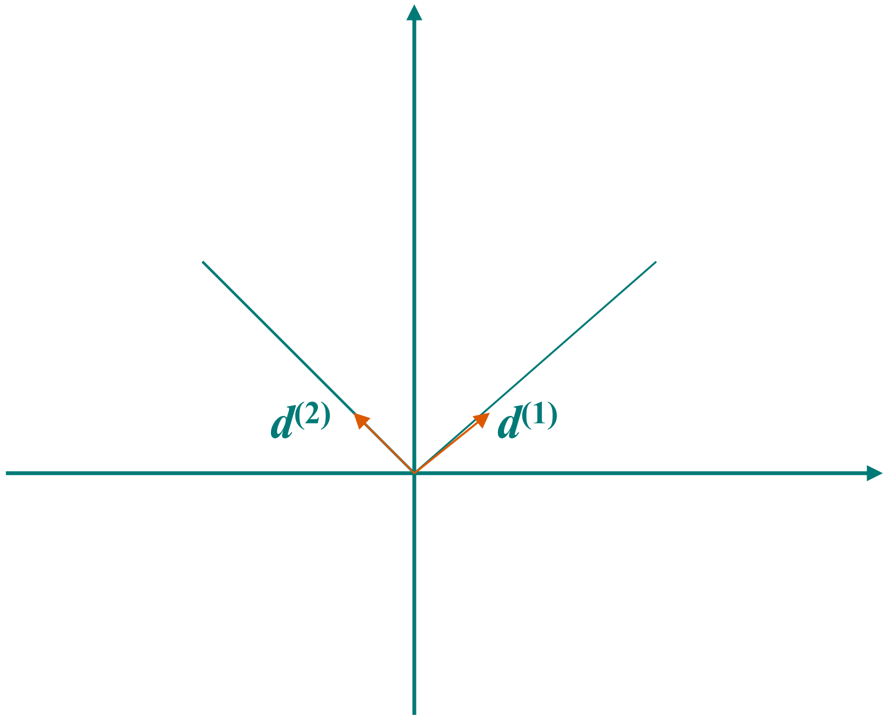
$$S=\{(x_1,x_2)^T \mid x_1 \geq 0, x_2 \geq 0\}$$

向量$d \geq 0, d \neq 0$是S的方向

例：

设$S=\{(x_1,x_2)^T \mid x_2 \ge |x_1|\}, d^{(1)}=(1,1)^T, d^{(2)}=(-1,1)^T$,

则$d^{(1)}, d^{(2)}$是S的极方向。



例：设$S=\{(x_1,x_2)^T \mid x_2 \geq |x_1|\}, d^{(1)}=(1,1)^T, d^{(2)}=(-1,1)^T$，

则$d^{(1)}, d^{(2)}$是S的极方向。

对$\forall x \in S, \forall \lambda \geq 0$, 有

$$x+\lambda d^{(1)}=\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}+\lambda\begin{pmatrix} 1 \\ 1 \end{pmatrix}=\begin{pmatrix} x_1+\lambda \\ x_2+\lambda \end{pmatrix},$$

$\because x \in S, \quad \therefore x_2 \geq |x_1|,$

而 $x_2+\lambda \geq |x_1|+\lambda \geq |x_1+\lambda|$

$\therefore \quad \left\{x+\lambda d^{(1)} \mid \lambda \geq 0\right\} \subset S$

$\Rightarrow \quad d^{(1)}$是$S$的方向。

设$d^{(1)}=\lambda_1\begin{pmatrix}x_1\\x_2\end{pmatrix}+\lambda_2\begin{pmatrix}y_1\\y_2\end{pmatrix}$, 其中$\lambda_1$，$\lambda_2>0$, $\begin{pmatrix}x_1\\x_2\end{pmatrix},\begin{pmatrix}y_1\\y_2\end{pmatrix}$是$S$的方向,

则有 $\begin{cases}1=\lambda_1x_1+\lambda_2y_1\\1=\lambda_1x_2+\lambda_2y_2\end{cases}\Rightarrow\lambda_1x_1+\lambda_2y_1=\lambda_1x_2+\lambda_2y_2$

$$\Rightarrow\quad x_1=\frac{\lambda_2}{\lambda_1}(y_2-y_1)+x_2$$

$$\because\begin{pmatrix}x_1\\x_2\end{pmatrix},\begin{pmatrix}y_1\\y_2\end{pmatrix}\text{是}S\text{的方向},\therefore x_2\geq|x_1|,\ y_2\geq|y_1|,\begin{pmatrix}x_1\\x_2\end{pmatrix}\neq0,\begin{pmatrix}y_1\\y_2\end{pmatrix}\neq0$$

$$\Rightarrow x_2\geq|x_1|=\left|\frac{\lambda_2}{\lambda_1}(y_2-y_1)+x_2\right|\Rightarrow y_2\leq y_1$$

$$\because y_2\geq|y_1|,\therefore y_2=y_1\Rightarrow x_2=x_1\Rightarrow\begin{pmatrix}x_1\\x_2\end{pmatrix}=\frac{x_1}{y_1}\begin{pmatrix}y_1\\y_2\end{pmatrix}$$

所以，$d^{(1)}$是S的极方向。

定理： 设 $S=\{x\mid Ax=b, x\geq 0\}\neq\varnothing$, d 是非零向量，则 d 是 S 的方向 $\Leftrightarrow d\geq 0$，且 $Ad=0$.

证明："$\Leftarrow$" $\forall x\in S, \lambda\geq 0$，有 $x+\lambda d\geq 0$，且

$$A(x+\lambda d)=Ax+\lambda Ad=b$$

$\therefore x+\lambda d\in S$，即 d 是 S 的方向。

"$\Rightarrow$" 设 d 是 S 的方向，则由 $x\in S$，得

$x+\lambda d\in S$，其中 $\lambda>0$

$\therefore A(x+\lambda d)=b$

由 $Ax=b, \lambda>0$，得 $Ad=0$.

又对 $\forall\lambda>0, x+\lambda d\geq 0, \therefore d\geq 0$.

多面集的表示定理

- 定理：设$\boldsymbol{S}=\{\boldsymbol{x}|\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}, \boldsymbol{x}\geq 0\}$为非空多面集，则有
- （1）极点集非空，且存在有限个极点$x^{(1)},\cdots,x^{(k)}$.
- （2）极方向集合为空集的充要条件是S有界；若S无界，则存在有限个极方向 $d^{(1)},d^{(2)},\cdots,d^{(l)}$.
- （3） $x\in S$的充要条件是

$$x=\sum_{j=1}^{k}\lambda_j x^{(j)}+\sum_{j=1}^{l}\mu_j d^{(j)}$$

其中$\lambda_j\geq 0,\ j=1,2,\cdots,k,\ \sum_{j=1}^{k}\lambda_j=1$

$$\mu_j\geq 0,\ j=1,2,\cdots,l.$$