最优化方法

哈尔滨工业大学
数学学院

杨　畅

第 2 章

基本概念
和
理论基础

第2章 基本概念和理论基础

2.1数学规划模型的一般形式

$$(fS)\begin{cases}\text{Min} \quad f(\boldsymbol{x}) & \text{——目标函数} \\ \text{s.t.} \quad \boldsymbol{x}\in S & \text{——约束集合，可行集}\end{cases}$$

其中，$S\subseteq \mathrm{R}^n$，$f:S\rightarrow \mathrm{R}$，$\boldsymbol{x}\in S$称（$f\ S$）的可行解

最优解：$\boldsymbol{x}^*\in S$，满足$f(\boldsymbol{x}^*)\leq f(\boldsymbol{x})$, $\forall\ \boldsymbol{x}\in S$，则称

$\boldsymbol{x}^*$为（$f\ S$）的全局最优解（最优解），记为

g.opt.（global optimum），简记为opt.。

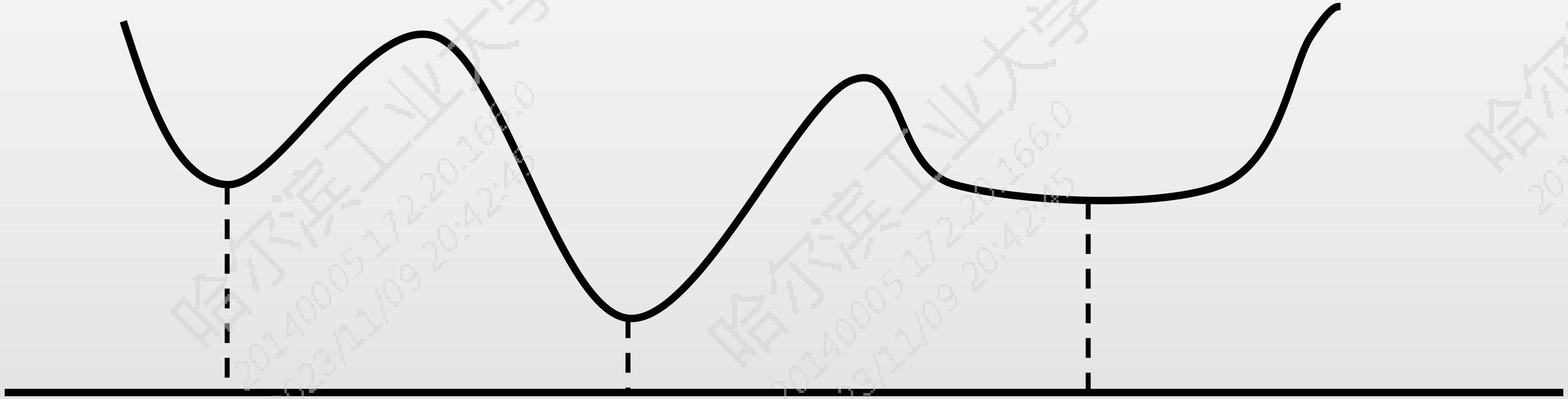
最优值：$\boldsymbol{x}^*$为（$f\ S$）的最优解， 则称 $f^*=f(\boldsymbol{x}^*)$ 为

（$f\ S$）的最优值（最优目标函数值）。

局部最优解：$\boldsymbol{x}^*\in S$，$\exists$ $\boldsymbol{x}^*$ 的邻域 $N(\boldsymbol{x}^*)$ ，使满足

$f(\boldsymbol{x}^*)\le f(\boldsymbol{x})$, $\forall$ $\boldsymbol{x} \in S \cap N(\boldsymbol{x}^*)$,则称 $\boldsymbol{x}^*$为（f S）的局部最优解，记为l .opt.（local optimum） 。

在上述定义中，当$\boldsymbol{x} \neq \boldsymbol{x}^*$ 时有严格不等式成立，则分别称 $\boldsymbol{x}^*$ 为（f S）的严格全局最优解和严格局部最优解。



函数形式： $f(\boldsymbol{x}),\ g_i(\boldsymbol{x}),\ h_j(\boldsymbol{x}):\ \mathbf{R}^n \to \mathbf{R}$

$$(fgh)\quad \begin{array}{ll} \text{Min} & f(\boldsymbol{x}) \\ \text{s.t.} & g_i(\boldsymbol{x}) \le 0,\quad \boldsymbol{i} = 1,2,\ldots,\boldsymbol{m} \\ & h_j(\boldsymbol{x}) = 0,\quad \boldsymbol{j} = 1,2,\ldots,\boldsymbol{l} \end{array}$$

矩阵形式：

$$(fgh)\quad \begin{array}{ll} \text{Min} & f(\boldsymbol{x}),\quad f(\boldsymbol{x}):\ \mathbf{R}^n \to \mathbf{R} \\ \text{s.t.} & \boldsymbol{g}(\boldsymbol{x}) \le \mathbf{0},\quad \boldsymbol{g}(\boldsymbol{x}):\ \mathbf{R}^n \to \mathbf{R}^m \\ & \boldsymbol{h}(\boldsymbol{x}) = \mathbf{0},\quad \boldsymbol{h}(\boldsymbol{x}):\ \mathbf{R}^n \to \mathbf{R}^l \end{array}$$

当 $f(\boldsymbol{x}),\ g_i(\boldsymbol{x}),\ h_j(\boldsymbol{x})$均为线性函数时，称其为线性规划；若其中有非线性函数时，称其为非线性规划。

2.2凸集、凸函数和凸规划

1.凸集

（1）凸集的概念

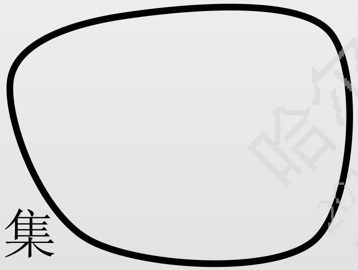
定义： 设集合 $S\subseteq \mathbf{R}^n$，若 $\forall \boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}\in S$，$\lambda\in[0,1]$， 必有

$$\lambda \boldsymbol{x}^{(1)}+(1-\lambda)\boldsymbol{x}^{(2)}\in S$$

，则称 S 为凸集。

规定： 单点集 $\{x\}$ 为凸集，空集$\varnothing$为凸集。

注： $\lambda \boldsymbol{x}^{(1)}+(1-\lambda)\boldsymbol{x}^{(2)}=\boldsymbol{x}^{(2)}+\lambda(\boldsymbol{x}^{(1)}-\boldsymbol{x}^{(2)})$ 是连接 $\boldsymbol{x}^{(1)}$与$\boldsymbol{x}^{(2)}$的线段 。


凸集


非凸集


非凸集

例： 证明集合 $S=\{\boldsymbol{x}\mid \boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}\}$ 是凸集。其中，$\boldsymbol{A}$为 $m\times n$矩阵，$\boldsymbol{b}$为m维向量。

例： 证明集合 $S=\{\boldsymbol{x}\mid \boldsymbol{A}\boldsymbol{x}\leq \boldsymbol{b}\}$ 是凸集。其中，$\boldsymbol{A}$为 $m\times n$矩阵，$\boldsymbol{b}$为m维向量。

凸组合：设 $\boldsymbol{x}^{(1)},\boldsymbol{x}^{(2)},\ldots,\boldsymbol{x}^{(m)}\in\mathrm{R}^n$，$\lambda_j\geqslant 0$，

$\sum_{j=1}^{m}\lambda_j=1$，那么称 $\sum_{j=1}^{m}\lambda_j\boldsymbol{x}^{(j)}$ 为$\boldsymbol{x}^{(1)},\boldsymbol{x}^{(2)},\ldots,\boldsymbol{x}^{(m)}$的凸组合。

注：若$\sum_{j=1}^{m}\lambda_j>0$，那么称 $\sum_{j=1}^{m}\lambda_j\boldsymbol{x}^{(j)}$ 为$\boldsymbol{x}^{(1)},\boldsymbol{x}^{(2)},\ldots,\boldsymbol{x}^{(m)}$的半正组合。

凸包：设$S\subseteq R^n$为非空集合（不一定是凸集），由S中所有的有限点的凸组合所构成的集合，被称为S的凸包，记为cov(S)。显然有，如果S是凸集，那么S=cov(S)。

- **比较**：$\boldsymbol{z}=\sum_{j=1}^{m}\alpha_j\boldsymbol{x}^{(j)}$

$\alpha_j\in\mathbf{R}$ —— 构成线性组合 —— 线性子空间

$\alpha_j\geq 0$，$\sum\alpha_j>0$ —— 构成半正组合 —— 凸锥

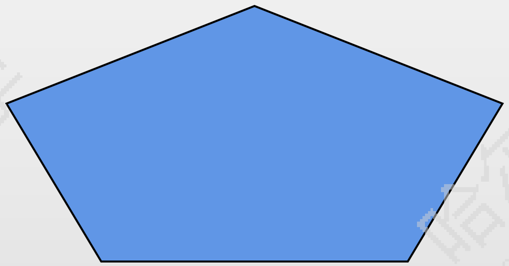
$\alpha_j\geq 0$，$\sum\alpha_j=1$ —— 构成凸组合 —— 凸集

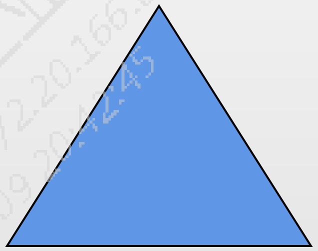
定理 S是凸集$\Leftrightarrow S$中任意有限点的凸组合属于S。

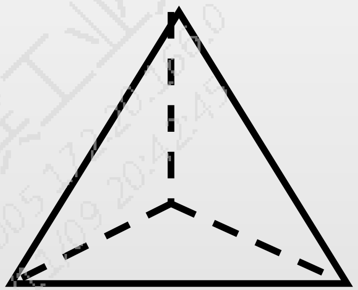
多胞形 $H(\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \dots, \boldsymbol{x}^{(m)})$：由 $\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \dots, \boldsymbol{x}^{(m)}$ 的所有凸组合构成。

单纯形：若多胞形 $H(\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \dots, \boldsymbol{x}^{(m)})$满足，

$$\boldsymbol{x}^{(2)} - \boldsymbol{x}^{(1)}, \boldsymbol{x}^{(3)} - \boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(m)} - \boldsymbol{x}^{(1)}$$ 线性无关。


多胞形


单纯形


单纯形

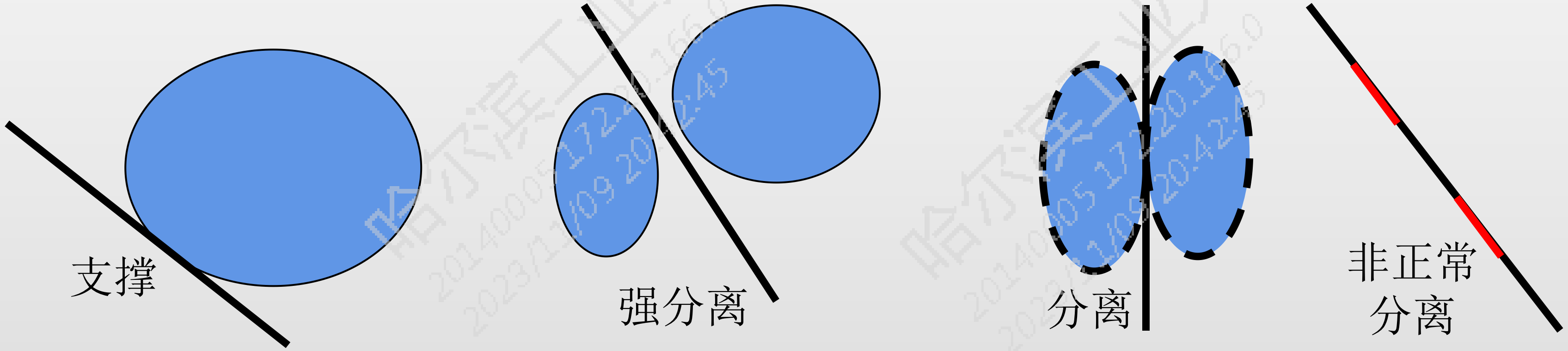
（2）凸集的性质

1) 凸集的交集是凸集；（并？）

2) 凸集的内点集是凸集；

3) 凸集的闭包是凸集。

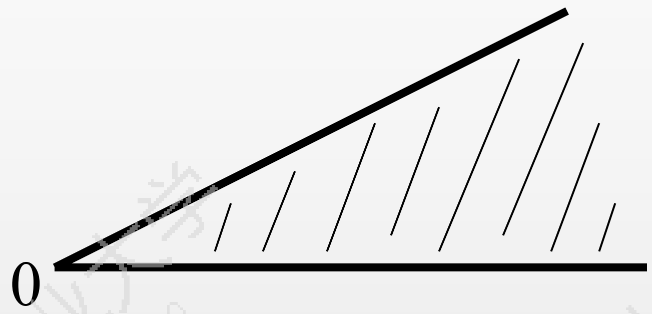
4) 分离与支撑：凸集边界上任意点存在支撑超平面；两个互相不交的凸集之间存在分离超平面。



（3）凸锥

定义 $C \subseteq \mathbf{R}^n$，若 $\boldsymbol{x} \in C$，$\lambda > 0$，有 $\lambda \boldsymbol{x} \in C$，则称 C 是以 $\mathbf{0}$ 为顶点的锥。如果 C 还是凸集，则称为凸锥。

集合 $\{\mathbf{0}\}$、$\mathbf{R}^n$ 是凸锥。



命题：C是凸锥$\Leftrightarrow$$C$中任意有限点的半正组合属于$C$ 。

2.凸函数

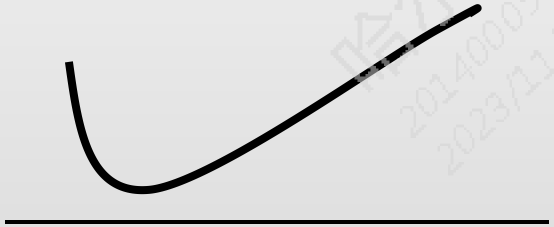
（1）凸函数及水平集

定义 设集合 $S \subseteq \mathbf{R}^n$ 为凸集，函数 $f: S \to \mathbf{R}$。若 $\forall\ \boldsymbol{x}^{(1)}$，$\boldsymbol{x}^{(2)} \in S$，$\lambda \in (0, 1)$，均有

$$f(\lambda \boldsymbol{x}^{(1)} + (1-\lambda)\boldsymbol{x}^{(2)}) \leq \lambda f(\boldsymbol{x}^{(1)}) + (1-\lambda) f(\boldsymbol{x}^{(2)}),$$

则称 $f(\boldsymbol{x})$ 为凸集 S 上的凸函数。

若进一步有上面不等式以严格不等式成立，则称 $f(\boldsymbol{x})$ 为凸集 S 上的严格凸函数。当 $-f(\boldsymbol{x})$ 为凸函数（严格凸函数）时，则称 $f(\boldsymbol{x})$ 为凹函数（严格凹函数）。



严格凸函数

凸函数

严格凹函数

定理 $f(x)$ 为凸集 S 上的凸函数$\Leftrightarrow$ S 上任意有限点的凸组合的函数值不大于各点函数值的凸组合。

思考：设f_1, f_2是凸函数。

1) 设λ_1, λ_2 > 0，$\lambda_1 f_1+\lambda_2 f_2$, $\lambda_1 f_1 - \lambda_2 f_2$是否为凸函数？

2）$f(x)= \max\{ f_1(x) , f_2(x) \}$, $g(x)= \min\{ f_1(x) , f_2(x) \}$是否为凸函数？

定义 设集合 $S \subseteq \mathbf{R}^n$，函数 $f:S\to\mathbf{R}$，$\alpha \in \mathbf{R}$，称 $S_\alpha = \{ \boldsymbol{x} \in S \mid f(\boldsymbol{x}) \leq \alpha \}$ 为 $f(x)$ 在 S 上的 α 水平集。

定理 设集合 $S \subseteq \mathbf{R}^n$是凸集，函数 $f:S\to\mathbf{R}$是凸函数，则对 $\forall \alpha \in \mathbf{R}$，$S_\alpha$ 是凸集。

注：

1) 水平集的概念相当于在地形图中，海拔高度不高于某一数值的区域。
2) 上述定理的逆不真。

考虑分段函数$f(x)$=1(x≥0)或0(x<0)，函数非凸，但任意水平集是凸集。

方向导数：设 $S \subseteq \mathbf{R}^n$ 为非空凸集，函数 $f: S \to \mathbf{R}$ ，再设 $\boldsymbol{x}^* \in S$, $\boldsymbol{d}$ 为方向，使当 $\lambda > 0$ 充分小时有 $\boldsymbol{x}^*+\lambda\boldsymbol{d} \in S$, 如果

$$\lim_{\lambda \to +0} \left[f(\boldsymbol{x}^*+\lambda\boldsymbol{d}) - f(\boldsymbol{x}^*) \right] / \lambda$$ 存在（包括 $\pm\infty$）

则称 $f(\boldsymbol{x})$ 为在点 $\boldsymbol{x}^*$ 沿方向 $\boldsymbol{d}$ 的方向导数存在，记

$$f'(\boldsymbol{x}^*;\boldsymbol{d}) = \lim_{\lambda \to +0} \left[f(\boldsymbol{x}^*+\lambda\boldsymbol{d}) - f(\boldsymbol{x}^*) \right] / \lambda$$

- 若 $f(\boldsymbol{x})$ 在 $\boldsymbol{x}^*$ 可导，则 $f'(\boldsymbol{x}^*;\boldsymbol{d}) = [\nabla f(\boldsymbol{x}^*)]^T \boldsymbol{d}$ 。

（2）凸函数的性质

以下设 $S \subseteq \mathbf{R}^n$ 为非空凸集，函数 $f: S \to \mathbf{R}$

1）若f 凸，则 f 在 S 的内点集上连续。

注： f 在 S 上不一定连续。

例如：$f(\boldsymbol{x})=2$(当$|\boldsymbol{x}|=1$)； $f(\boldsymbol{x})=\boldsymbol{x}^2$ (当$|\boldsymbol{x}|<1$) 。

2）设f 凸，则对任意方向的方向导数存在(前提首先要是“可行方向”)。

3）设 S 是开集，f 在 S 上可微，则f凸$\Leftrightarrow$ $\boldsymbol{x}^* \in S$，有

$$f(\boldsymbol{x}) \geq f(\boldsymbol{x}^*)+\nabla f^{T}(\boldsymbol{x}^*)(\boldsymbol{x}-\boldsymbol{x}^*), \forall \boldsymbol{x} \in S。$$

4） 设 S 是开集，f 在 S 上二次可微，则

①f 凸$\Leftrightarrow$ $\forall \boldsymbol{x} \in S$，$\nabla^2 f(\boldsymbol{x})$ 半正定；

②若 $\forall \boldsymbol{x} \in S$，$\nabla^2 f(\boldsymbol{x})$ 正定，则f严格凸。

例：

1) $f(x) = x_1^2 + 2x_1x_2 + 2x_2^2 + 10x_1 - 4$

2) $f(x) = -3x_1^2 + x_1x_2 - x_2^2 - 2x_3^2 - 2x_2x_3 + 26$

3) $f(x) = 3x_1^2 + ax_1x_2 + 2x_2^2 - 4x_1 + 6$（$a$=5, 4.5）

3.凸规划

当（f S）中，S为凸集，f是S上的凸函数（求min）时，称（f S）为凸规划。

对于（fgh），当f，g_i为凸函数，h_j为线性函数时，（fgh）为凸规划。

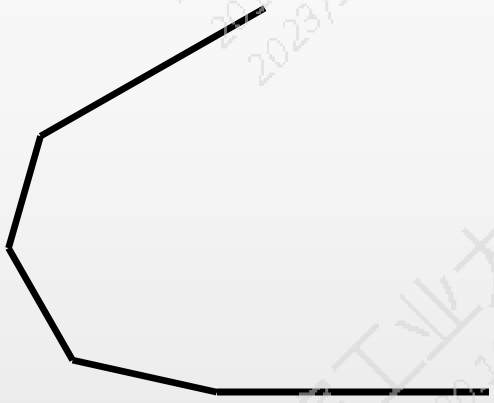
定理 设集合 $S \subseteq \mathbf{R}^n$ 为凸集，函数 $f: S \to \mathbf{R}$ 。

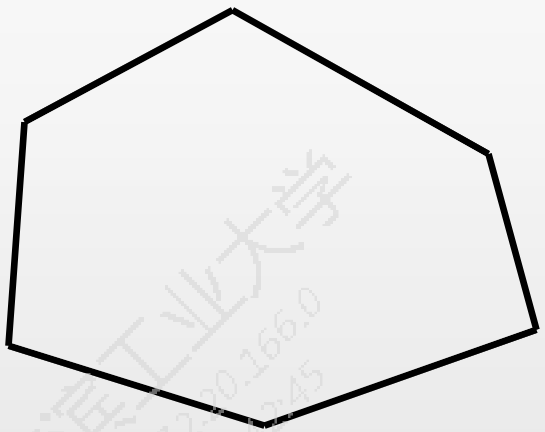
$f(x)$ 为凸集 S 上的凸函数。$\boldsymbol{x}^*$为问题（fs）的l.opt.，则$\boldsymbol{x}^*$为g.opt.；

又如果f是严格凸函数,那么$\boldsymbol{x}^*$是（fs）的唯一g.opt.。

2.3 多面体、极点、极方向

（1）**多面体**：有限个半闭空间的交





例：$S=\{\,x\in R^n \mid Ax=b\,,\ x\geq 0\,\}$

（2）**多面体的极点（顶点）**：$x\in S$，不存在 S 中的另外两个点$x^{(1)}$和$x^{(2)}$，及 $\lambda\in(0,1)$，使

$$x = \lambda x^{(1)}+(1-\lambda)x^{(2)}$$ 。

（3）**方向**：$x\in S$，$d\in R^n$，$d\neq 0$ 及 $\lambda>0$，总有 $x+\lambda d\in S$(可行方向)。其中，当$d^{(1)}=\lambda d^{(2)}$（$\lambda>0$）时，称 $d^{(1)}$和$d^{(2)}$同方向。

（4）**极方向**：方向 d 不能表示为两个不同方向的非负组合。

定理（极点特征）设 A 满秩，x 是 S 极点的充分必要条件是:

存在分解 $A=(B, N)$，其中B为m阶非奇异矩阵，使 $x^{T}=(x_B^{T}, x_N^{T})$，这里

$x_B=B^{-1}b\geq 0$，$x_N=0$。

注：S中必存在有限多个极点（$\leq C_n^m$）。

例 多面体：$S=\{x \mid Ax=b, x\geq 0\}$

其中$A=\begin{pmatrix}3 & 1 & 2 & 0\\ 0 & 4 & 5 & 1\end{pmatrix}$, $b=\begin{pmatrix}3\\ 6\end{pmatrix}$

解：取$B_1=\begin{pmatrix}3 & 1\\ 0 & 4\end{pmatrix}$,则$B_1^{-1}b=\begin{pmatrix}1/3 & -1/12\\ 0 & 1/4\end{pmatrix}\begin{pmatrix}3\\ 6\end{pmatrix}=\begin{pmatrix}1/2\\ 3/2\end{pmatrix}\geq 0$

所以$x^{(1)}=(1/2,3/2,0,0)^T$是一个极点

取$B_2=\begin{pmatrix}2 & 0\\ 5 & 1\end{pmatrix}$,则$B_2^{-1}b=\begin{pmatrix}1/2 & 0\\ -5/2 & 1\end{pmatrix}\begin{pmatrix}3\\ 6\end{pmatrix}=\begin{pmatrix}3/2\\ -3/2\end{pmatrix}\ngeq 0$

所以这个分解不对应极点

取$B_3=\begin{pmatrix}3 & 0\\ 0 & 1\end{pmatrix}$,则$B_3^{-1}b=\begin{pmatrix}1/3 & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}3\\ 6\end{pmatrix}=\begin{pmatrix}1\\ 6\end{pmatrix}\geq 0$

所以$x^{(3)}=(1,0,0,6)^T$是一个极点

例 考虑多面体 $S=\{\, x\in \mathbf{R}^n \mid Ax=b,\ x\geq 0 \,\}$，其中

$$A=\begin{pmatrix} 3 & 2 & 1 & 0 & 0 \\ 2 & 1 & 0 & 1 & 0 \\ 0 & 3 & 0 & 0 & 1 \end{pmatrix} \qquad b=\begin{pmatrix} 65 \\ 40 \\ 75 \end{pmatrix}$$

即

$$\begin{aligned} 3x_1+2x_2+x_3 &= 65 \\ 2x_1+x_2+x_4 &= 40 \\ 3x_2+x_5 &= 75 \\ x_1,\ x_2,\ x_3,\ x_4,\ x_5 &\geq 0 \end{aligned}$$

$$\boldsymbol{A} = (\boldsymbol{p}_1, \boldsymbol{p}_2, \boldsymbol{p}_3, \boldsymbol{p}_4, \boldsymbol{p}_5) = \begin{pmatrix} 3 & 2 & 1 & 0 & 0 \\ 2 & 1 & 0 & 1 & 0 \\ 0 & 3 & 0 & 0 & 1 \end{pmatrix}$$

A矩阵包含以下10个3×3的子矩阵：

$B_1=(p_1, p_2, p_3)$ $B_2=(p_1, p_2, p_4)$

$B_3=(p_1, p_2, p_5)$ $B_4=(p_1, p_3, p_4)$

$B_5=(p_1, p_3, p_5)$ $B_6=(p_1, p_4, p_5)$

$B_7=(p_2, p_3, p_4)$ $B_8=(p_2, p_3, p_5)$

$B_9=(p_2, p_4, p_5)$ $B_{10}=(p_3, p_4, p_5)$

其中$|B_4|=0$，因而B_4不能构成极点和极方向。其余均为非奇异方阵，因此该问题共有9个可构成极点、极方向的子矩阵，我们称之为基。

对于基$B_3=(p_1, p_2, p_5)$，令$x_3=0$，$x_4=0$，在等式约束中令$x_3=0$，$x_4=0$，解线性方程组

$$\begin{aligned}3x_1+2x_2+0x_5&=65\\2x_1+x_2+0x_5&=40\\0x_1+3x_2+x_5&=75\end{aligned}$$

得到$x_1=15$，$x_2=10$，$x_5=45$，对应的极点

$$\boldsymbol{x}=(x_1, x_2, x_3, x_4, x_5)^{\mathrm{T}}$$
$$=(15, 10, 0, 0, 45)^{\mathrm{T}}$$

类似可得到极点

$x^{(2)} = (5, 25, 0, 5, 0)^T$ （对应B_2）

$x^{(7)} = (20, 0, 5, 0, 75)^T$ （对应B_5）

$x^{(8)} = (0, 25, 15, 15, 0)^T$ （对应B_7）

$x^{(9)} = (0, 0, 65, 40, 75)^T$ （对应B_{10}）

而 $x^{(3)} = (0, 32.5, 0, 7.5, -22.5)^T$ （对应B_9）

$x^{(4)} = (65/3, 0, 0, -10/3, 75)^T$ （对应B_6）

$x^{(5)} = (7.5, 25, -7.5, 0, 0)^T$ （对应B_1）

$x^{(6)} = (0, 40, -15, 0, -45)^T$ （对应B_8）

不是极点。

定理（极方向特征）设 $A=(a_1,a_2,\dots,a_n)$ 秩为m

(1) d 是 S 方向的充分必要条件是$Ad=0$,且$d\geq 0$

(2) d 是 S 极方向的充分必要条件是:

存在分解 $A=(B,N)$，其中B为m阶非奇异矩阵，对于N中的列向量 a_j 使 $B^{-1}a_j\leq 0$,

$$d^T=(d_B^T,d_N^T),$$

这里$d_B=-B^{-1}a_j$，$d_N=(0,\dots,\underset{j}{1},\dots,0)^T$

注：S中必存在有限多个极方向 $(\leq (n-m)C_n^m)$。

例 多面体：$S=\{x \mid Ax=b, x\geq 0\}$

其中$A=\begin{pmatrix}3 & 1 & 2 & -6 & 0\\ 0 & 4 & 5 & -5 & 1\end{pmatrix}, b=\begin{pmatrix}3\\ 6\end{pmatrix}$

解：考虑分解$B_1=\begin{pmatrix}3 & 1\\ 0 & 4\end{pmatrix}, N=\begin{pmatrix}2 & -6 & 0\\ 5 & -5 & 1\end{pmatrix}$

取A的第4列$a_4=\begin{pmatrix}-6\\ -5\end{pmatrix}$

则$B_1^{-1}a_4=\begin{pmatrix}1/3 & -1/12\\ 0 & 1/4\end{pmatrix}\begin{pmatrix}-6\\ -5\end{pmatrix}=\begin{pmatrix}-19/12\\ -5/4\end{pmatrix}\leq 0$

所以一个极方向为d= $(19/12, 5/4, 0, 1, 0)^T$

总结

极点

$$\begin{cases} x_B = B^{-1}b \geq 0 \\ x_N = 0 \end{cases}$$

极方向

$$\begin{cases} d_B = -B^{-1}a_j \geq 0 \\ d_N = e_j \end{cases}$$

多面体 $S = \{ x \in R^n \mid Ax = b, x \geq 0 \}$的极点和极方向

定理（表示定理） 考虑上述多面体 S，设 A 满秩，$x^{(1)}, x^{(2)}, \ldots, x^{(k)}$ 为所有极点，$d^{(1)}, d^{(2)}, \ldots, d^{(l)}$ 为所有极方向。那么，对于 $\forall x \in S$，$\exists \lambda_i \geq 0$，$i=1, 2, \ldots, k$，且 $\lambda_1 + \lambda_2 + \ldots + \lambda_k = 1$，$\mu_j \geq 0$，$j = 1, 2, \ldots, l$，有

$$x = \lambda_1 x^{(1)} + \lambda_2 x^{(2)} + \ldots + \lambda_k x^{(k)} + \mu_1 d^{(1)} + \mu_2 d^{(2)} + \ldots + \mu_l d^{(l)}$$