高级算法设计与分析

（最优化课件研制组）

线性规划 Part 1

主讲：刘国义

退出

开始

线性规划 Part 1

在最优化中，目标函数和约束函数皆为线性函数的优化问题称为线性规划（LP），它是相对简单的最优化问题。本章是有关线性规划的理论与求解方法的内容。

2.1 线性规划的各种形式

1. 标准形式和典范形式

如下形式的线性规划

$$\left.\begin{aligned} &\min \sum_{j=1}^{n} c_j x_j \\ &s.t.\ \sum_{j=1}^{n} a_{ij} x_j = b_i, \quad i = 1,2,\cdots,m \\ &\qquad x_j \ge 0, \quad j = 1,2,\cdots,n. \end{aligned}\right\} \qquad (2.1)$$

称为线性规划的**标准形式**。其中各 c_j 称为价格系数，各 b_i 称为右端项。

采用向量一矩阵表示法，标准形式可以简写为。

$$\left.\begin{aligned} &\min \vec{c}^T \vec{x}; \\ &s.t.\ A\vec{x} = \vec{b} \\ &\qquad \vec{x} \ge 0. \end{aligned}\right\} \qquad (2.2)$$

在进行理论分析时，有时需要把 A 表示成

$$A=[\vec{a}_1,\vec{a}_2,\cdots,\vec{a}_n],$$

这样，（2.2）中的 $A\vec{x}=\vec{b}$ 又可写成

$$\sum_{j=1}^{n} x_j\vec{a}_j=\vec{b}$$

若 A 中有 m 个列向量可以合并成为单位矩阵，且 $\vec{b}\geq\vec{0}$，此时（2.2）则称为线性规划的**典范形式**。

例如，如下线性规划

$$\begin{aligned}
&\min 3x_1-4x_2+2x_3-5x_4;\\
&s.t.\ \ 2x_1 \qquad\qquad -6x_3+x_4=5,\\
&\qquad 4x_1+\ \ x_2-3x_3 \qquad\quad =8,\\
&\qquad\quad x_1,\quad x_2,\quad x_3,\quad x_4\geq 0,
\end{aligned}$$

就呈现为典范形式，因为 $\vec{a}_4=[1,0]^T$ 和 $\vec{a}_2=[0,1]^T$ 可合并成单位矩阵。

不失一般性，假定单位矩阵位于前 m 列，即典范形式呈现为

$$\left.\begin{array}{l}
\min c_1x_1+c_2x_2+\cdots+c_nx_n \\
s.t.\quad x_1 \qquad\qquad +a_{1,m+1}x_{m+1}+\cdots+a_{1,n}x_n=b_1 \\
\qquad\qquad x_2 \qquad +a_{2,m+1}x_{m+1}+\cdots+a_{2,n}x_n=b_2 \\
\qquad\qquad\qquad\qquad \vdots \qquad\qquad\qquad \vdots \\
\qquad\qquad\qquad x_m+a_{m,m+1}x_{m+1}+\cdots+a_{m,n}x_n=b_m \\
\quad x_j\ge 0,\quad j=1,2,\cdots,n
\end{array}\right\} \quad (2.3)$$

其中 $b_i\ge 0(i=1,2,\cdots,m)$ 。

用向量一矩阵表示法，那么（2.3）可简写成

$$\left.\begin{array}{l}
\min \vec{c}_I^{\,T}\vec{x}_I+\vec{c}_N^{\,T}\vec{x}_N \\
s.t.\quad I\vec{x}_I+N\vec{x}_N=\vec{b} \\
\qquad\quad \vec{x}_I,\vec{x}_N\ge 0.
\end{array}\right\} \quad (2.4)$$

2. 一般形式

线性规划

$$
\left.
\begin{aligned}
&\min \sum_{j=1}^{t} c_j x_j \\
&s.t.\ \sum_{j=1}^{t} a_{pj} x_j \le b_p, \quad p = 1,2,\cdots,u \\
&\qquad \sum_{j=1}^{t} a_{qj} x_j \ge b_q, \quad q = u+1,\cdots,u+v \\
&\qquad \sum_{j=1}^{t} a_{rj} x_j = b_r, \quad r = u+v+1,\cdots,m \\
&\qquad\qquad x_j \ge 0, \quad j = 1,2,\cdots,t.
\end{aligned}
\right\} \qquad (2.5)
$$

3. 一般形式与标准形式的关系

（1）松弛变量

考虑“≤”约束中的第 p 个不等式

$$\sum_{j=1}^{t} a_{pj}x_j \le b_p, \qquad （2.6）$$

引入非负变量 x_{t+p}，迫使

$$\sum_{j=1}^{t} a_{pj}x_j + x_{t+p} = b_p. \qquad （2.7）$$

使不等式约束（2.6）变为等式约束（2.7）的非负变量 x_{t+p} 称为**松弛变量**。

（2）剩余变量

考虑“≥”约束中的第 q 个不等式

$$\sum_{j=1}^{t} a_{qj}x_j \ge b_q, \qquad （2.8）$$

引入非负变量 x_{t+q}，迫使

引入非负变量 x_{t+q} ，迫使

$$\sum_{j=1}^{t} a_{qj}x_j - x_{t+q} = b_q. \qquad (2.9)$$

使不等式约束（2.8）变为等约束（2.9）的非负变量x_{t+q}称为**剩余变量**。

在引入 u 个松弛变量、v 个剩余变量后，线性规划（2.5）可化成标准形式：

$$\left.\begin{aligned}
&\min \sum_{j=1}^{t} c_j x_j \\
&s.t.\ \sum_{j=1}^{t} a_{pj}x_j + x_{t+p} = b_p, \quad p = 1,2,\cdots,u \\
&\qquad \sum_{j=1}^{t} a_{qj}x_j - x_{t+q} = b_q, \quad q = u+1,\cdots,u+v \\
&\qquad \sum_{j=1}^{t} a_{rj}x_j \qquad\quad = b_r, \quad r = u+v+1,\cdots,m \\
&\qquad\qquad x_j \ge 0, \quad j = 1,2,\cdots,t+u+v.
\end{aligned}\right\} \qquad (2.10)$$

它含有 $t+u+v=n$ 个变量、m 个等式约束。

注意，新引入变量的价格系数全部设为零。因此，在（2.10）的目标函数中没有出现新变量。

下面说明线性规划（2.5）与其标准形式（2.10）是等价的。

首先，它们的容许点是一一对应的，且对应的容许点的函数值相等。

因为若 $[x_1, x_2, \cdots, x_t]^T$ 是（2.5）的一个容许点，那么按公式（2.7）和（2.9）引入非负的松弛变量和剩余变量 $x_{t+1}, \cdots, x_n$ 后，显然 $[x_1, x_2, \cdots, x_t, x_{t+1}, \cdots, x_n]^T$ 将是（2.10）的容许点。反之亦然。故（2.5）的容许点 $[x_1, x_2, \cdots, x_t]^T$ 与（2.10）的容许点 $[x_1, x_2, \cdots, x_t, x_{t+1}, \cdots, x_n]^T$ 一一对应。

又（2.5）与（2.10）的目标函数相同，且都只是 $x_1, \cdots, x_t$ 的函数，所以（2.5）与（2.10）所对应的容许点的函数值相等。

其次，若 $[x_1^*, x_2^*, \cdots, x_t^*, x_{t+1}^*, \cdots, x_n^*]^T$ 是（2.10）的最优点，它所对应的最优值为 $z^* = \sum_{j=1}^{t} c_j x_j^*$ ，那么，由前面的证明可知，其前 t 个分量组成的向量 $[x_1^*, x_2^*, \cdots, x_t^*]^T$ 也一定（2.5）的最优点。反之亦然。

因此，线性规划（2.5）与其标准形式（2.10）是等价的。

该结论表明，可以只讨论标准线性规划。

例2.1 将线性规划

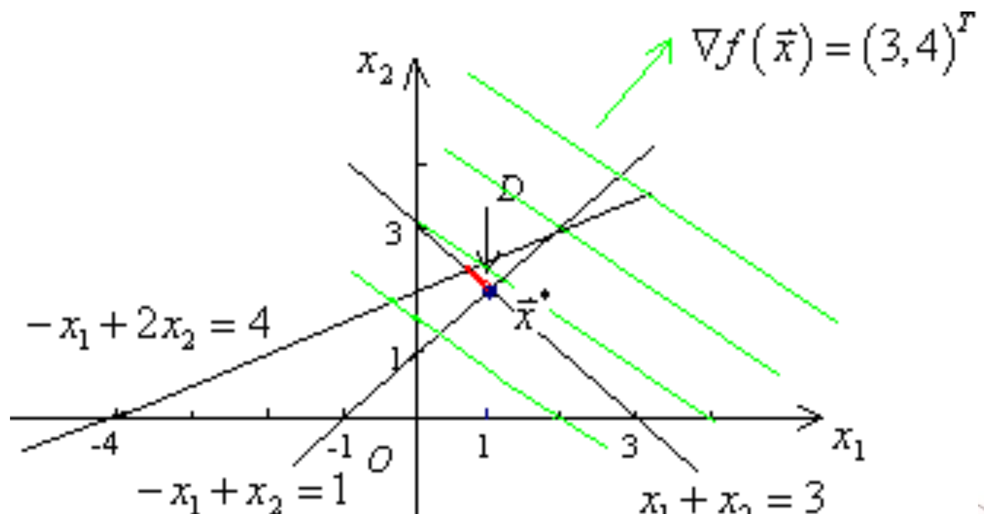
$$
\begin{aligned}
& \min 3x_1 + 4x_2 \\
& s.t. -x_1 + 2x_2 \le 4 \\
& \quad -x_1 + \ \ x_2 \ge 1 \\
& \quad\ \ x_1 + \ \ x_2 = 3 \\
& \quad\ \ x_1, \ \ x_2 \ge 0.
\end{aligned}
$$

化为标准形式，并用图解法求解原问题，给出标准形式的解。

解 对第1个约束引入松弛变量 x_3，对第2个约束引入剩余变量 x_4 。于是，该线性规划的标准形式为

$$
\begin{aligned}
&\min 3x_1 + 4x_2 \\
&s.t. -x_1 + 2x_2 + x_3 \qquad\quad = 4 \\
&\qquad -x_1 + \ x_2 \qquad\quad - x_4 = 1 \\
&\qquad\ \ x_1 + \ x_2 \qquad\qquad\quad\ = 3 \\
&\qquad\ \ x_1,\quad x_2,\quad x_3,\ x_4 \geq 0.
\end{aligned}
$$

图解法求解原线性规划如下：



最优解 $\vec{x}^*$ 在直线 $-x_1+x_2=1$ 与 $x_1+x_2=3$ 的交汇处，即 $\vec{x}^*=[1,2]^T$ 。相应的标准形式的最优解为 $\vec{x}'^*=[1,2,1,0]^T$ 。

（3）自由变量

以上讨论都考虑变量的取值是非负的（当变量的取值非正，那么它的负变量的取值即是非负的）。实际中，如果某些变量没有这种约束，也就是说，某些变量可以任意取值，那么这些变量称为**自由变量**。自由变量可以通过以下两种方法把它消除。

例如，假若 x_1 是自由变量。

第一种方法：引入两个非负变量 x_1^+ 和 x_1^- ，令

$$x_1=x_1^+-x_1^- \qquad （2.11）$$

将其代入到线性规划的目标函数和约束函数中，自由变量 x_1 就消除了。注意，求出新线性规划的最优点后，再利用（2.11）便可以定出 x_1 。

第二种方法：取一个包含 x_1 的等式约束，例如

$$a_{i1}x_1+a_{i2}x_2+\cdots+a_{in}x_n=b_i$$

由此解出

$$x_1=\frac{b_i}{a_{i1}}-\frac{a_{i2}}{a_{i1}}x_2-\cdots-\frac{a_{in}}{a_{i1}}x_n \qquad (2.12)$$

将此式代入到线性规划的目标函数和其他约束函数中，自由变量 x_1 也消除了。求出新线性规划的最优点后，利用（2.12）再定出x_1。

第一种方法将增加变量的数目，导致问题的维数增大。第二种方法正好相反。

2.2 基本定理

考虑标准线性规划（2.2），即

$$\begin{aligned}&\min \vec{c}^T\vec{x};\\&s.t.\ A\vec{x}=\vec{b}\\&\qquad \vec{x}\ge 0.\end{aligned}$$

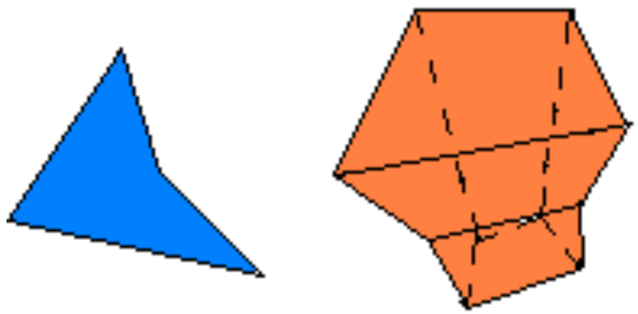
记容许集 $D=\{\vec{x} \mid A\vec{x}=\vec{b}, \vec{x}\geq\vec{0}\}$。不妨假定 $R(A)=m<n$，$\vec{b}\geq\vec{0}$。

1. 极点与基本容许解

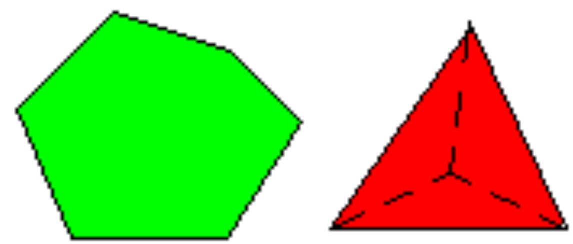
定义2.2 有限个半空间的交称为**凸多面体**。

半空间是凸集，故凸多面体是凸集。

边界为直线或平面是多面体的几何特征。

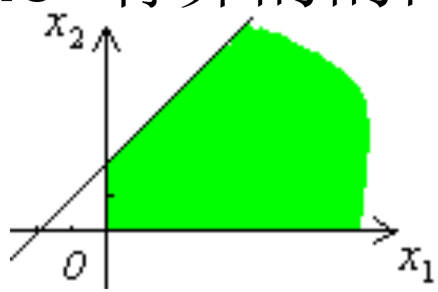


多面体

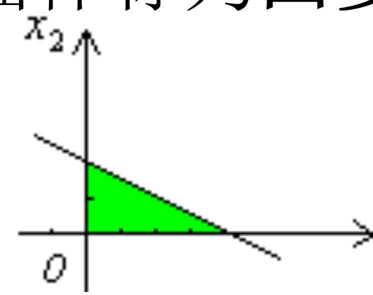


凸多面体

定义2.3 有界的的凸多面体称为**凸多胞形**。



凸多面体



凸多胞形

定理2.1 线性规划（2.2）的容许集 D 是凸多面体。

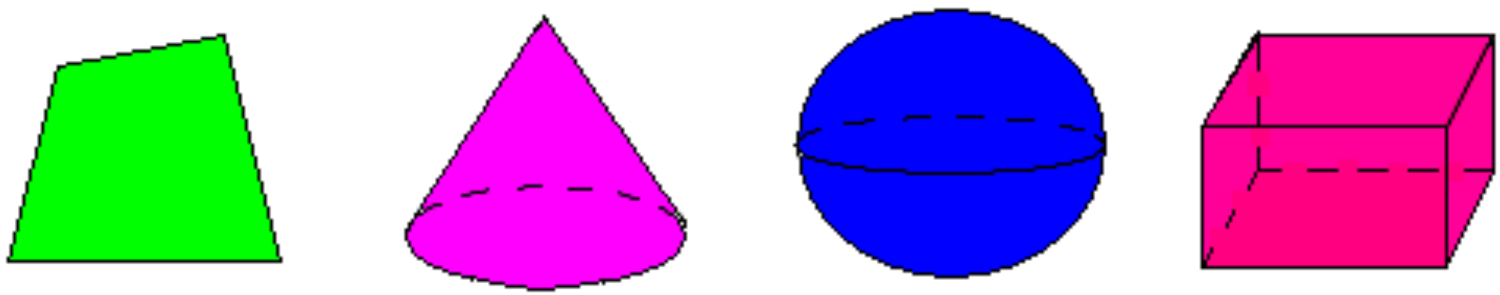
(2)凸集的极点

与凸集相关的另一个重要概念是凸集的极点。

定义2.5 若凸集 D 中的某个点 $\vec{x}$ 不能表示为这个集合中另外两个不同点的严格凸组合，即

$$\vec{x} \neq \lambda\vec{x}_1 + (1-\lambda)\vec{x}_2, \quad \forall\vec{x}_1, \vec{x}_2 \in D, \vec{x}_1 \neq \vec{x}_2; \forall\lambda \in (0,1)$$

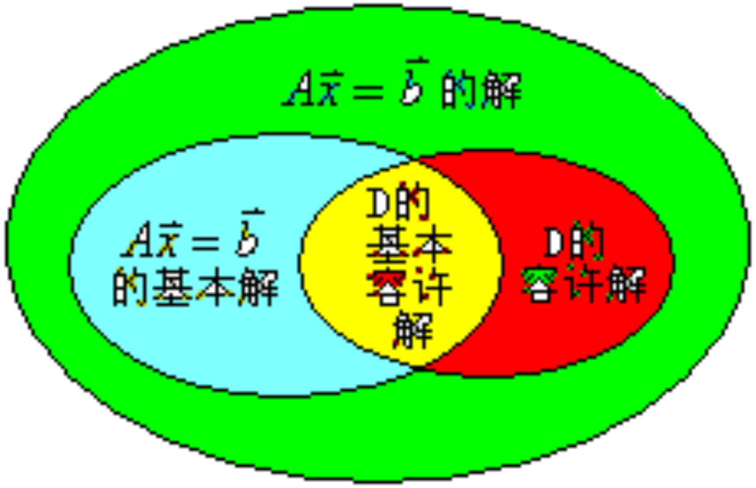
则 $\vec{x}$ 称为**凸集 D 的极点**。



(3)基本容许解

当（2.2）的容许解又是“$A\vec{x}=\vec{b}$ 的基本解”时，就称其为（2.2）的基本容许解。

方程组 $A\vec{x}=\vec{b}$ 满足“基本”条件的的解称为 $A\vec{x}=\vec{b}$ 的基本解。

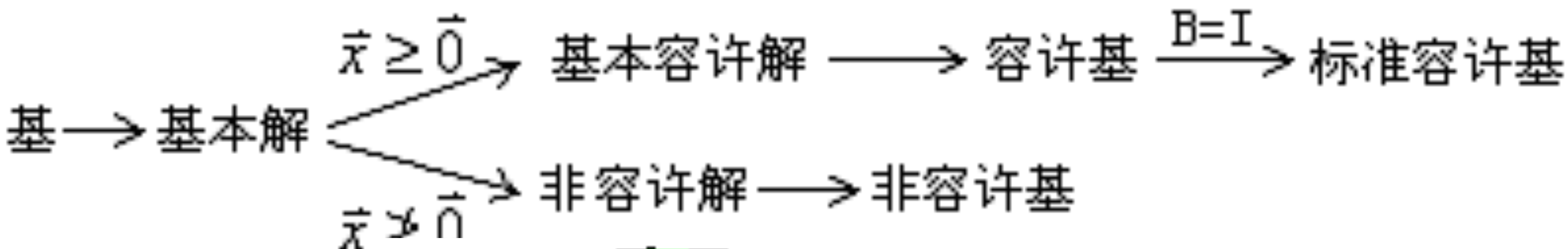


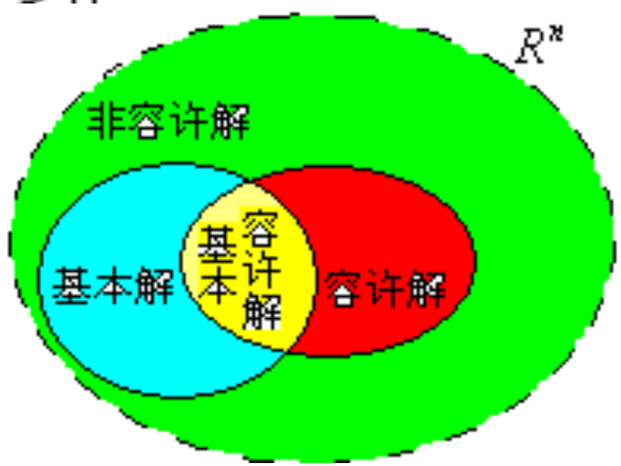
$A\vec{x}=\vec{b}$ 的基本解是如何定义的呢？

将 $A\vec{x}=\vec{b}$ 表示成向量形式

$$x_1\vec{a}_1+x_2\vec{a}_2+\cdots+x_n\vec{a}_n=\vec{b}$$

因为 $R(A)=m$ ，故 A 中必含有 m 阶的可逆矩阵 B，称为**基**。**基** B 的每个列向量称为**基向量**，A 的其余列向量称为**非基向量**。与基向量对应的变量称为**基变量**，与非基向量对应的变量称为**非基变量**。若基是单位矩阵，则称为**标准基**。非基变量取值皆为0的 $A\vec{x}=\vec{b}$ 解称为**基本**解。满足 $\vec{x}\geq\vec{0}$ 的基本解称为线性规划（2.2）关于基 B 的**基本容许解**。





不失一般性，设$A=[B,N],\vec{x}_B=[x_1,\cdots,x_m]^T,\vec{x}_N=[x_{m+1},\cdots,x_n]^T$，则

$$A\vec{x}=\vec{b}\Rightarrow(B,N)\begin{pmatrix}\vec{x}_B\\ \vec{x}_N\end{pmatrix}=\vec{b}\Rightarrow B\vec{x}_B+N\vec{x}_N=\vec{b}\Rightarrow\vec{x}_B=B^{-1}\vec{b}-B^{-1}N\vec{x}_N$$

令 $\vec{x}_N=\vec{0}$，得基本解 $\vec{x}_B=B^{-1}\vec{b}$。若 $B^{-1}\vec{b}\geq\vec{0}$，那么该基本解是关于基 B 的基本容许解。

例2.3 P48

定义2.10 基变量取值皆不为0的基本解称为**非退化的**，否则称为**退化的**；若（2.2）的所有基本容许解都是非退化的，则线性规划（2.2）称为**非退化的**，否则称为**退化的**。

例2.4 P49

(3)基本容许解与极点的关系

$$D=\{\vec{x}\,|\,A\vec{x}=\vec{b},\vec{x}\geq\vec{0}\}$$

引理2.2 设 $A=[\vec{a}_1,\vec{a}_2,\cdots,\vec{a}_n]_{\mathrm{m\times n}}$ ，$R(A)=m$ ，$\vec{x}\in D$。则 $\vec{x}$是基本容许解 $\Leftrightarrow$ $\vec{x}$的正分量所对应的 A的列向量线性无关。

定理2.3 设 $\vec{x}\in D$,则 $\vec{x}$ 是（2.2）的基本容许解 $\Leftrightarrow$ $\vec{x}$ 是D的极点。

从几何角度讲，例2.3中的约束条件

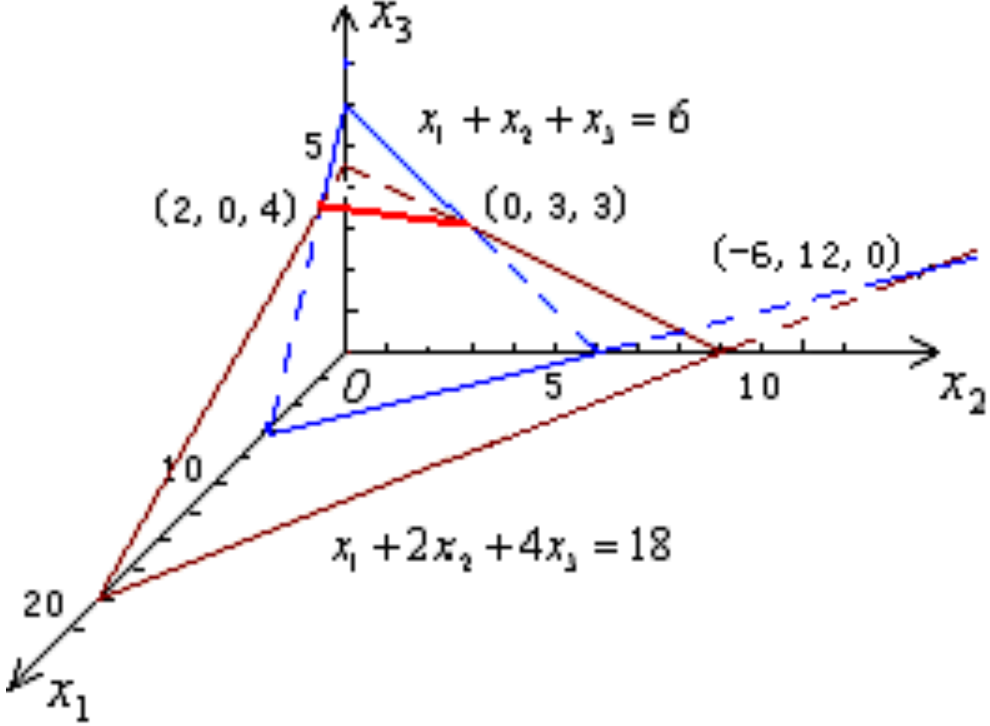
$$x_1+x_2+x_3=6$$

$$x_1+2x_2+4x_3=18$$

$$x_1,x_2,x_3\geq 0$$

表示空间一条直线在第一象限中的直线段部分。

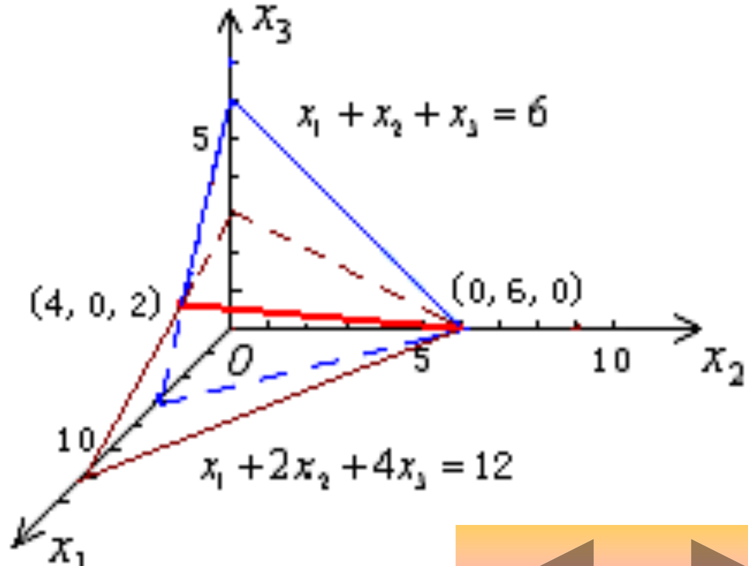
如图所示：



红线部分为容许集 D。D 有两个极点 $(0,3,3)^T$ 和 $(2,0,4)^T$，它们恰恰是两个基本容许解。

例2.4如图所示：

D 有两个极点 $(0,6,0)^T$ 和 $(4,0,2)^T$



推论2.4 容许集 D 的极点个数有限。

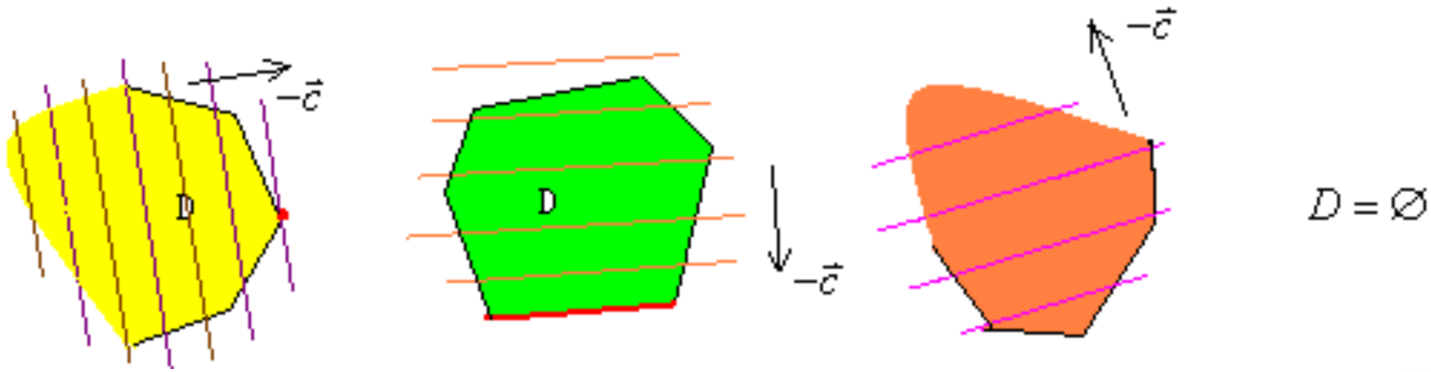
2. 线性规划基本定理

引理2.5 设 $\vec{x}\in D$,不妨设$\vec{x}=[x_1,\cdots,x_l,0,\cdots,0]^T\ (x_i>0,1\le i\le l)$,且 $\vec{a}_1,\vec{a}_2,\cdots,\vec{a}_l$ 线性相关。那么一定存在最多有 $l-1$ 个正分量的容许解。

定理2.7 线性规划（2.2）若有容许解，则必有基本容许解。

定理2.8 线性规划（2.2）若有最优解，则必有最优基本容许解。

从几何上看，解线性规划存在以下几种可能情形：

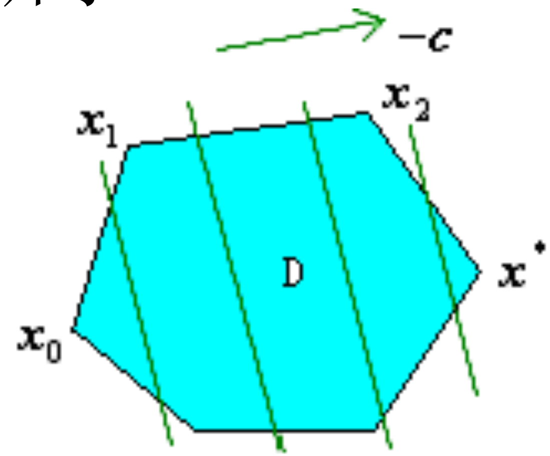


唯一解　　无穷多解　　解无界　　无解

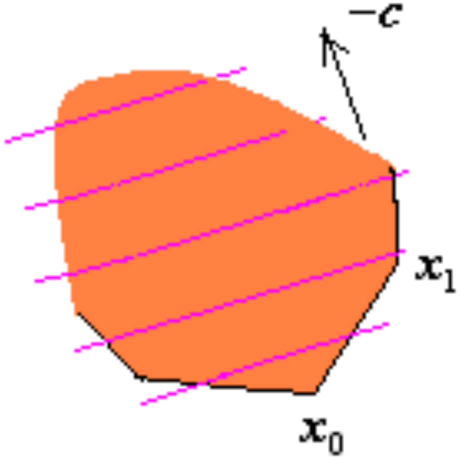
2.3 单纯形法

由前述知，（2.2）的容许集 D 是凸多面体，D 中的极点个数有限。（2.2）若有容许解，则必有基本容许解；若有最优解，则必有最优基本容许解。据此，对于线性规划，我们只关心基本容许解即可。因此，一个显而易见的求解方法是求出全部基本容许解（极点），比较目标函数值就能确定最优解。可是，当 n,m 数值较大时，这种法的计算量相当大，逆矩阵也不易求。Dantzig给出的单纯形法基本上解决了这个问题。<u>单纯形法的基本思想</u>是：首先找到（求出）一个极点（基本容许解）$\vec{x}_0$，并依据最优性准则判断其最优性，如非最优，则沿凸多面体 D 的一条棱找到（迭代到）使目标值降低（不找比 $\vec{x}_0$ 的目标值高的）的下一个极点（基本容许解）$\vec{x}_1$，……。因为极点个数是有限的，所以在一定的条件下，总可以找到（迭代到）最优极点（最优基本容许解）或者说明解无界。

如图所示。



有最优点



解无界

Dantzig单纯形法的思想涉及以下三个具体问题：

一、初始基本容许解的产生；

二、最优性准则；

三、基本容许解的改进。

1. 最优性准则

考虑标准线性规划

$$\left.\begin{aligned} \min \vec{c}^T\vec{x} \\ s.t.\quad A\vec{x}=\vec{b} \\ \vec{x}\geq\vec{0}. \end{aligned}\right\} \qquad (2.21)$$

设 B 为容许基。并不妨设

$$A=[B,N]\ ,\ \vec{x}=\begin{bmatrix}\vec{x}_B\\ \vec{x}_N\end{bmatrix},\ \vec{c}=\begin{bmatrix}\vec{c}_B\\ \vec{c}_N\end{bmatrix}\ 。$$

将（2.21）改写为

$$\left.\begin{aligned} \min \vec{c}_B^T\vec{x}_B+\vec{c}_N^T\vec{x}_N \\ s.t.\quad B\vec{x}_B+N\vec{x}_N=\vec{b} \\ \vec{x}_B,\vec{x}_N\geq\vec{0}. \end{aligned}\right\} \qquad (2.22)$$

令 $\vec{x}_N=\vec{0}$，得关于 B 的基本容许解 $\bar{\vec{x}}=\begin{bmatrix} B^{-1}\vec{b} \\ \vec{0} \end{bmatrix}$，

目标函数值 $\bar{z}=\vec{c}_B^T B^{-1}\vec{b}$

设 $\vec{x}=[\vec{x}_B^T,\vec{x}_N^T]^T$ 为任一容许解，则由（2.22）解出

$$\vec{x}_B=B^{-1}\vec{b}-B^{-1}N\vec{x}_N$$

代入目标函数中得

$$z=\vec{c}_B^T(B^{-1}\vec{b}-B^{-1}N\vec{x}_N)+\vec{c}_N^T\vec{x}_N$$

$$=\bar{z}-(\vec{c}_B^T B^{-1}N-\vec{c}_N^T)\vec{x}_N.$$

令 $\vec{\sigma}_N^T=\vec{c}_B^T B^{-1}N-\vec{c}_N^T$，于是

$$z=\bar{z}-\vec{\sigma}_N^T\vec{x}_N \qquad （2.26）$$

故

$$z-\bar{z}=-\vec{\sigma}_N^T\vec{x}_N$$

<u>因为</u> $\vec{x}_N\ge\vec{0}$ <u>，因此，只要</u> $\vec{\sigma}_N\le\vec{0}$ <u>，必有</u> $z\ge\bar{z}$.

由此得到判断基本容许解是最优解的一个充分条件。

定义2.11 在标准线性规划（2.21）中，设B是一个基，则

$$\sigma_j = \vec{c}_B^T B^{-1} \vec{a}_j - c_j, \quad j = 1, 2, \cdots, n$$

称为变量 x_j 关于基 B 的**判别数**。

判别数的向量形式为

$$\vec{\sigma} = \vec{c}_B^T B^{-1} A - \vec{c}$$

易知 $\vec{\sigma}_B = \vec{0}$ 。

定理2.9（线性规划最优性准则） 在标准线性规划（2.21）中，设 B是容许基。若所有变量关于基 B的判别数皆非正，则关于基 B 的基本容许解是最优解。

定义2.12 在标准线性规划（2.21）中，若关于基B基本容许解是最优解，则称 B 是**最优基**。

例2.5 考虑标准线性规划

$$\min x_1 + x_2 + 2x_3 + 5x_4$$

$$\begin{aligned} s.t.\ x_1 \quad\quad - 2x_3 + \ x_4 &= 1 \\ x_1 + x_2 + \ x_3 \quad\quad &= 4 \\ x_1,\ x_2,\ x_3,\ x_4 \geq 0. \end{aligned}$$

试分别求关于基 $B_0 = [\vec{a}_4, \vec{a}_2]$和 $B_1 = [\vec{a}_1, \vec{a}_2]$的基本解。若是容许解，则判别其最优性。

解 关于B_0 的基变量取值 $\vec{x}_{B_0} = \begin{bmatrix} x_4 \\ x_2 \end{bmatrix} = B_0^{-1}\vec{b} = \vec{b} = \begin{bmatrix} 1 \\ 4 \end{bmatrix} \geq \vec{0},$

故关于 B_0 的基本解 $\vec{x}_0 = [0, 4, 0, 1]^T$ 是基本容许解。计算

$$\vec{u}_0 = \vec{c}_{B_0}^T,$$

$$B_0^{-1} = [5,1]\sigma_2 = \sigma_4 = 0,$$

$$\sigma_1 = \vec{u}_0\vec{a}_1 - c_1 = [5,1]\begin{bmatrix} 1 \\ 1 \end{bmatrix} - 1 = 5,$$

$$\sigma_4 = \vec{u}_0\vec{a}_3 - c_3 = [5,1]\begin{bmatrix} -2 \\ 1 \end{bmatrix} - 2 = -11,$$

最优性准则未满足，因此不能断定 $\vec{x}_0$ 是否是最优解。

$$B_1 = [\vec{a}_1, \vec{a}_2] = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \quad B_1^{-1} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix},$$

因此关于B_1的基变量取值 $\vec{x}_{B_1} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = B_1^{-1}\vec{b} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 4 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix} \geq \vec{0},$

故关于 B_1 的基本解 $\vec{x}_1 = [1,3,0,0]^T$

也是基本容许解。计算

$$\vec{c}_{B_1} = [c_1, c_2]^T = [1,1]^T,$$

$$\vec{u}_1 = \vec{c}_{B_1}^T B_1^{-1} = [0,1],$$

$$\sigma_1 = \sigma_2 = 0,$$

$$\sigma_3 = \vec{u}_1 \vec{a}_3 - c_3 = [0,1]\begin{bmatrix} -2 \\ 1 \end{bmatrix} - 2 = -1,$$

$$\sigma_4 = \vec{u}_0 \vec{a}_4 - c_4 = [0,1]\begin{bmatrix} 1 \\ 0 \end{bmatrix} - 5 = -5,$$

最优性准则满足，因此断定 $\vec{x}_1$是最优解。