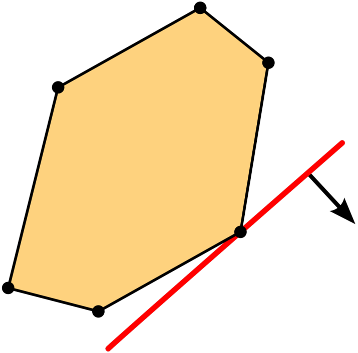
第2章 线性规划

线性规划：

目标函数与约束函数都是变量的线性函数的最优化问题称为线性规划 (Linear Programming, LP)。工程与管理科学中大量的问题都是变量数目为成百上千，乃至上万或数十万的线性规划问题。



线性规划发展重要事件

- 法国数学家傅里叶和瓦莱－普森分别于1832和1911年独立地提出线性规划的想法。
- 1939年苏联数学家康托罗维奇在《生产组织与计划中的数学方法》一书中提出线性规划问题。
- 1947年美国数学家G. B.丹齐克提出线性规划的通用方法——单纯形法，为这门学科奠定了基础。
- 1950～1956年，线性规划的对偶理论出现。1954年，C.莱姆基提出对偶单纯形法。
- 1978年，苏联数学家哈奇扬提出求解线性规划问题的多项式时间算法（内点算法），具有重要理论意义。
- 1984年，在美国贝尔实验室工作的印度裔数学家卡玛卡提出有效求解实际线性规划问题的多项式时间算法－Karmarkar算法。

本章主要学习内容

2-1　LP问题基本概念与最优性

数学模型

线性规划(Linear Programming)数学模型：

$$\begin{array}{ll} \text{min} & \boldsymbol{c}^T\boldsymbol{x}, \\ \text{s.t.} & A_1\boldsymbol{x} = \boldsymbol{b}_1, A_2\boldsymbol{x} \geq \boldsymbol{b}_2, \end{array}$$

其中：

- 价值向量： $\boldsymbol{c} = (c_1, c_2, \cdots, c_n)^T$
- 右端向量： $\boldsymbol{b} = (b_1, b_2, \cdots, b_m)^T$, $\boldsymbol{b}_2 = (b_{m+1}, \cdots, b_p)^T$,
- 约束矩阵：

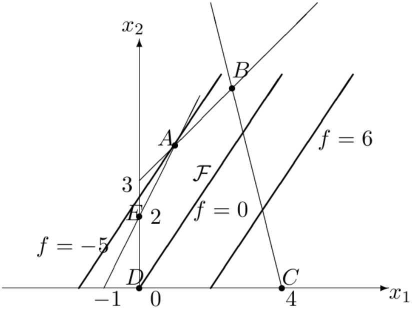
$$A_1 = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ldots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, A_2 = \begin{bmatrix} a_{m+1,1} & a_{m+1,2} & \cdots & a_{m+1,n} \\ \vdots & \vdots & \ldots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pn} \end{bmatrix}.$$

图解法

图解法：通过在可行域内移动等值线来确定最优解的方法。

例2.1.1 用图解法求解以下线性规划问题：

$$\begin{cases} \min & 3x_1 - 2x_2, \\ \text{s.t.} & -x_1 + x_2 \le 3, \\ & -2x_1 + x_2 \le 2, \\ & 4x_1 + x_2 \le 16, \\ & x_1 \ge 0, \quad x_2 \ge 0. \end{cases}$$



观察分析：

- 可行域由线段AB,BC,CD,DE,EA围成的凸多边形。可行域的顶点A为最优解，最优函数值为-5。
- 最优解A可以由方程组$-x_1 + x_2 = 3$，$-2x_1 + x_2 = 2$的解直接确定，即有$x_1^* = 1$，$x_2^* = 4$。

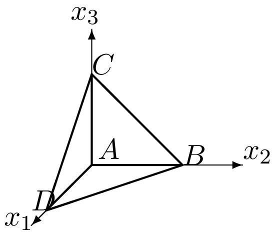
图解法

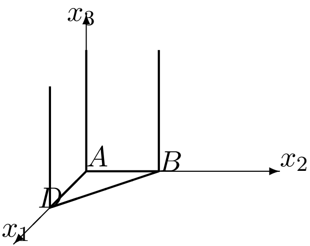
若干结论：

1. 若LP问题可行域非空，则可行域为凸多面体；
2. 若LP问题有最优解，则最优解为可行域的某个顶点；
3. 若可行域有界且顶点个数为有限，则问题的最优解必在某个顶点处达到；
4. 最优解可由最优顶点处的有效约束形成的方程组确定。

LP问题可行域

两个变量情形，可行域是凸多边形; 三个变量情形:可行域为凸多面体。维数大于3时，可行域为一些超平面所围成的超凸多面体。凸多面体可能有界，也可能无界。





有界可行域： 由约束条件$x_1+x_2+x_3 \leq 1, x_1 \geq 0, x_2 \geq 0, x_3 \geq 0$确定.

无界可行域： 由约束条件$x_1 + x_2 \leq 1, x_1 \geq 0, x_2 \geq 0, x_3 \geq 0$确定.

LP问题标准形式

标准形式: 只有线性等式约束和变量的非负约束

$$\begin{cases} \min & f(\boldsymbol{x}) = c_1x_1 + c_2x_2 + \cdots + c_nx_n, \\ \text{s.t.} & a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = b_i, \quad i = 1, 2, \cdots, m, \\ & x_j \geq 0, j = 1, 2, \cdots, n, \end{cases}$$

矩阵向量表达式:

$$\begin{cases} \min & f(\boldsymbol{x}) = \boldsymbol{c}^T\boldsymbol{x}, \\ \text{s.t.} & A\boldsymbol{x} = \boldsymbol{b}, \\ & x \geq 0. \end{cases} \qquad A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ldots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

约束系数矩阵

其中:

- 价值系数向量-$\boldsymbol{c} = (c_1, c_2, \cdots, c_n)^T$;
- 约束右端向量-$\boldsymbol{b} = (b_1, b_2, \cdots, b_m)^T \in \mathbb{R}^m$ (要求$\boldsymbol{b} \geq 0$).

化标准形

三个环节:

- 目标函数的转换:
  - 函数$f(\boldsymbol{x})$极大化问题转换成求$-f(\boldsymbol{x})$极小化问题
  - 删除目标函数中常数项
- 约束条件的转换:
  - 不等式两边乘以-1保证右端向量分量非负($\boldsymbol{b} \geq 0$)
  - 引入松弛变量和剩余变量化不等式约束为等式约束
- 变量的非负约束:
  - 自由变量写成两个非负变量的差, $x_j = x_j^{'} - x_j^{''}$.

化标准形例题

例2.1.2 将以下LP问题转化为标准型：

$$\begin{cases} \max & 3x_1 - 2x_2 + x_3, \\ \text{s.t.} & -x_1 + x_2 \geq 3, \\ & -2x_1 + x_2 \leq 2, \\ & 4x_1 + x_2 - x_3 = -16, \\ & x_1 \geq 0, \quad x_2 \geq 0. \end{cases}$$

解： 引入剩余变量x_4,松弛变量x_5, 两非负变量x_3', x_3'', 改写为标准型:

$$\begin{cases} \min & -3x_1 + 2x_2 - x_3' + x_3'', \\ \text{s.t.} & -x_1 + x_2 - x_4 = 3, \\ & -2x_1 + x_2 + x_5 = 2, \\ & -4x_1 - x_2 + x_3' - x_3'' = 16, \\ & x_1 \geq 0, x_2 \geq 0, x_3' \geq 0, x_3'' \geq 0, x_4 \geq 0, x_5 \geq 0. \end{cases}$$

约束方程组

约束相容性与满秩假设：

- **情形1：** ($m > n$) 约束方程组或者不相容，可行域为空集，或者删除相关约束得$m \leq n$。
- **情形2：** ($m = n$且满秩) 则约束方程组有唯一解，且还满足变量非负条件，则此唯一解为最优解；否则，无最优解。
- **情形3：** ($m < n$且A行满秩) 若A不是行满秩，删除相关约束不影响问题的可行域,剩下的约束形成满秩系数矩阵。

线性规划搜索空间：

当$m < n$时，从无限多可行解中确定目标函数值最优的解。

标准形LP问题的解

考虑标准形LP划问题等式约束$A\boldsymbol{x}=\boldsymbol{b}$, 将$A$(行满秩)进行分块:

$$A=[B_{m\times m}\ N_{m\times(n-m)}],\quad \boldsymbol{x}=\begin{pmatrix}\boldsymbol{x}_B\\ \boldsymbol{x}_N\end{pmatrix}\text{(对应分块)}$$

其中:

$$B=\begin{bmatrix}a_{11} & a_{12} & \cdots & a_{1m}\\ a_{21} & a_{22} & \cdots & a_{2m}\\ \vdots & \vdots & \vdots & \vdots\\ a_{m1} & a_{m2} & \cdots & a_{mm}\end{bmatrix}\quad N=\begin{bmatrix}a_{1,m+1} & a_{1,m+2} & \cdots & a_{1n}\\ a_{2,m+1} & a_{2,m+2} & \cdots & a_{2n}\\ \vdots & \vdots & \vdots & \vdots\\ a_{m,m+1} & a_{m,m+2} & \cdots & a_{mn}\end{bmatrix}$$

定义2.1.1 B与N分别称为**基矩阵**(非奇异)与**非基矩阵**; 将相应于B的m阶子向量$\boldsymbol{x}_B$称为**基向量**; 将相应于N的$(n-m)$阶子向量$\boldsymbol{x}_N$称为**非基向量**。$\boldsymbol{x}_B$和$\boldsymbol{x}_N$的分量分别称为**基变量**与**非基变量**。

标准形LP问题的解

方程组$A\boldsymbol{x}=\boldsymbol{b}$可以改写成:

$$A\boldsymbol{x}=B\boldsymbol{x}_B+N\boldsymbol{x}_N=\boldsymbol{b}$$

由于B非奇异，则有:

$$\boldsymbol{x}_B=B^{-1}(\boldsymbol{b}-N\boldsymbol{x}_N)=B^{-1}\boldsymbol{b}-B^{-1}N\boldsymbol{x}_N.$$

由此得到:

$$\boldsymbol{x}=\begin{pmatrix}\boldsymbol{x}_B\\ \boldsymbol{x}_N\end{pmatrix}=\begin{pmatrix}B^{-1}\boldsymbol{b}-B^{-1}N\boldsymbol{x}_N\\ \boldsymbol{x}_N\end{pmatrix}.$$

对任意$\boldsymbol{x}_N\in\mathbb{R}^{(n-m)}$, 上述$\boldsymbol{x}$都是方程组$A\boldsymbol{x}=\boldsymbol{b}$的解。

基本可行解

定义2.1.2： 当$\boldsymbol{x}_N = 0$时，LP问题对应的解为$\boldsymbol{x} = \begin{pmatrix} B^{-1}\boldsymbol{b} \\ 0 \end{pmatrix}$，称$\boldsymbol{x}$为LP问题的**基本解**。

定义2.1.3： 对于LP问题的基本解$\boldsymbol{x}$,

- 若$\boldsymbol{x} = \begin{pmatrix} B^{-1}\boldsymbol{b} \\ 0 \end{pmatrix} \geq 0$，则称$\boldsymbol{x}$为LP问题的**基本可行解**。
- 当$B^{-1}\boldsymbol{b}$存在小于零的分量时，称$\boldsymbol{x}$为LP问题的**不可行基本解**。

基矩阵B至多有:

$$C_n^m = \begin{pmatrix} n \\ m \end{pmatrix} = \frac{n!}{m!(n-m)!}$$

个选择，故基本解个数的上界为C_n^m，基本可行解(即顶点)的个数要少于基本解的个数。

退化线性规划

退化情形：

- 若基本可行解中基变量取零值，则称其为**退化的**基本可行解。
- 存在退化顶点的线性规划称为**退化线性规划**。

退化现象是由于约束条件中存在可删除的约束导致的。通常状况下，对线性规划作**非退化假定**。

基本解例题分析

例2.1.3 考察如下LP问题：

$$\begin{cases} \min & 3x_1 - 2x_2, \\ \text{s.t.} & -x_1 + x_2 \le 3, \\ & -2x_1 + x_2 \le 2, \\ & 4x_1 + x_2 \le 16, \\ & x_1 \ge 0, \quad x_2 \ge 0. \end{cases}$$

的约束构成可行域，验证基本可行解同可行域顶点的对应关系。

分析： 引入松弛变量x_3，x_4，x_5，将约束转换为标准形

$$\begin{aligned} -x_1 + x_2 + x_3 \qquad\qquad\qquad &= 3 \\ -2x_1 + x_2 \quad + x_4 \qquad\quad &= 2 \\ 4x_1 + x_2 \qquad\qquad + x_5 &= 16 \\ x_1 \ge 0, x_2 \ge 0, x_3 \ge 0, x_4 \ge 0, x_5 \ge 0. & \end{aligned}$$

由于$n=5$, $m=3$, 对于B共有：

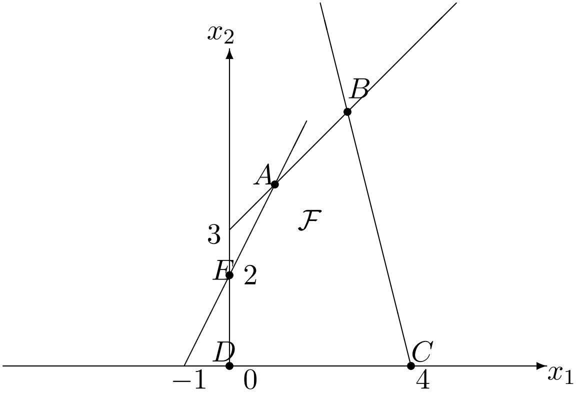
$$C_5^3 = 10$$

种可能的选择, 即存在如下10个基本解

$$\begin{aligned}
\boldsymbol{x}^{(1)} &= (0,0,3,2,16)^T & \boldsymbol{x}^{(2)} &= (0,2,1,0,14)^T \\
\boldsymbol{x}^{(3)} &= (4,0,7,10,0)^T & \boldsymbol{x}^{(4)} &= (\tfrac{13}{5}, \tfrac{28}{5}, 0, \tfrac{18}{5}, 0)^T \\
\boldsymbol{x}^{(5)} &= (1,4,0,0,8)^T & \boldsymbol{x}^{(6)} &= (\tfrac{14}{6}, \tfrac{40}{6}, -\tfrac{8}{6}, 0, 0)^T \\
\boldsymbol{x}^{(7)} &= (-3,0,0,-4,28)^T & \boldsymbol{x}^{(8)} &= (-1,0,2,0,20)^T \\
\boldsymbol{x}^{(9)} &= (0,3,0,-1,13)^T & \boldsymbol{x}^{(10)} &= (0,16,-13,-14,0)^T.
\end{aligned}$$

注意: 前5个是基本可行解(无负分量): 后5个是不可行基本解。

可行域的5个顶点：

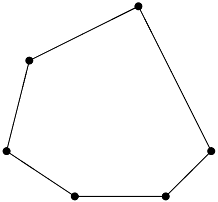


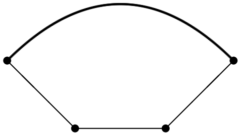
$\boldsymbol{x}^{(1)}$对应D,$\boldsymbol{x}^{(2)}$对应E, $\boldsymbol{x}^{(3)}$对应C, $\boldsymbol{x}^{(4)}$对应B, $\boldsymbol{x}^{(5)}$对应A。

可行域顶点

定义2.1.4

设$D \subset \mathbb{R}^n$为凸集,且$\boldsymbol{x} \in D$, 若$\boldsymbol{x}$不能表示成D中任何不同于$\boldsymbol{x}$的两个点的凸组合，即$\nexists \boldsymbol{y}, \boldsymbol{z} \in D$, $\boldsymbol{y} \neq \boldsymbol{x}$, $\boldsymbol{z} \neq \boldsymbol{x}$, 使得$\boldsymbol{x} = \lambda \boldsymbol{y} + (1-\lambda)\boldsymbol{z}, \lambda \in (0,1)$成立, 则称$\boldsymbol{x}$为凸集$D$的顶点.





LP问题的可行域为凸集, 其顶点$\boldsymbol{x}$也不可能表示成其任意不同于$\boldsymbol{x}$的两个可行点的凸组合。

定理2.1.1 基本可行解与顶点一一对应充要条件.

LP问题的可行点$\boldsymbol{x} \in \mathcal{F} = \{\boldsymbol{x} \mid A\boldsymbol{x} = \boldsymbol{b},\ \boldsymbol{x} \geq 0\}$是$\mathcal{F}$的顶点的充要条件为$\boldsymbol{x}$是LP问题的基本可行解。

证明： **充分性** 设$\boldsymbol{x}$为基本可行解, 即存在分块$A = [B, N]$，使得:

$$\boldsymbol{x} = \begin{pmatrix} \boldsymbol{x}_B \\ \boldsymbol{x}_N \end{pmatrix} = \begin{pmatrix} B^{-1}\boldsymbol{b} \\ 0 \end{pmatrix} \geq 0$$

为方便，假定B由A的前m列构成, N由后面的$n-m$列构成。**反证法**, 假定$\boldsymbol{x}$不是顶点, 则$\exists \boldsymbol{y}, \boldsymbol{z} \in \mathcal{F}, \boldsymbol{y} \neq \boldsymbol{x}, \boldsymbol{z} \neq \boldsymbol{x}$,以及$\lambda \in (0,1)$, 有:$\boldsymbol{x} = \lambda\boldsymbol{y} + (1-\lambda)\boldsymbol{z}$。对$\boldsymbol{y}, \boldsymbol{z}$作类似的分块:

$$\boldsymbol{y} = \begin{pmatrix} \boldsymbol{y}_B \\ \boldsymbol{y}_N \end{pmatrix} \quad \boldsymbol{z} = \begin{pmatrix} \boldsymbol{z}_B \\ \boldsymbol{z}_N \end{pmatrix}$$

证明续： 根据$\boldsymbol{y},\boldsymbol{z}$的可行性,

$$\boldsymbol{y}_N \geq 0, \boldsymbol{z}_N \geq 0, 0 = \boldsymbol{x}_N = \lambda \boldsymbol{y}_N + (1-\lambda)\boldsymbol{z}_N$$

易知: $\boldsymbol{y}_N = \boldsymbol{z}_N = 0$。又由$\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}$可行性得:

$$B\boldsymbol{x}_B = B\boldsymbol{y}_B = B\boldsymbol{z}_B = \boldsymbol{b}$$

由于B为非奇异矩阵, 方程组$B\boldsymbol{u} = \boldsymbol{b}$有唯一解, 故$\boldsymbol{x}_B = \boldsymbol{y}_B = \boldsymbol{z}_B$, 与$\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}$互不相同矛盾, 所以$\boldsymbol{x}$为可行域顶点。

必要性 假设$\boldsymbol{x}$为顶点，下证其为基本可行解。由$\boldsymbol{x}$的可行性, 根据$\boldsymbol{x}$的非零分量与零分量将其进行分块:

$$\boldsymbol{x} = \begin{pmatrix} \boldsymbol{x}_B \\ \boldsymbol{x}_N \end{pmatrix}, \quad \boldsymbol{x}_B > 0, \quad \boldsymbol{x}_N = 0$$

定理2.1.1证明

证明续： 假定dim($\boldsymbol{x}_B$)=m', 对应地将矩阵A进行分块：

$$A = [B_{m\times m'} \quad N_{m\times(n-m')}].$$

下证B的列向量线性无关。

反证法 假定B的列向量线性相关, 即存在非零向量$\boldsymbol{w} \in R^{m'}$满足$B\boldsymbol{w} = 0$.

结合$\boldsymbol{x}$的可行性$B\boldsymbol{x}_B = \boldsymbol{b}$，对所有$\lambda \in \mathbb{R}$, 有如下成立：

$$B(\boldsymbol{x}_B \pm \lambda\boldsymbol{w}) = B\boldsymbol{x}_B \pm \lambda B\boldsymbol{w} = \boldsymbol{b}$$

选取合适的λ值, 可保证$\boldsymbol{x}_B \pm \lambda\boldsymbol{w} \geq 0$, 令:

$$\boldsymbol{y} = \begin{pmatrix} \boldsymbol{x}_B + \lambda\boldsymbol{w} \\ 0 \end{pmatrix}, \boldsymbol{z} = \begin{pmatrix} \boldsymbol{x}_B - \lambda\boldsymbol{w} \\ 0 \end{pmatrix}$$

易证明$\boldsymbol{y}, \boldsymbol{z}$均可行($A\boldsymbol{y} = A\boldsymbol{z} = \boldsymbol{b}$), 且$\boldsymbol{x} = \frac{\boldsymbol{y}+\boldsymbol{z}}{2}$, 与$\boldsymbol{x}$为顶点矛盾。

定理2.1.1证明

证明续： 由于B列线性无关，则必有$m' \leq m$.

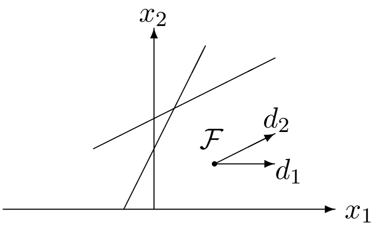
- 若$m' = m$，此时B即非奇异基矩阵，由$\boldsymbol{x}_N = 0$和$B\boldsymbol{x}_B + N\boldsymbol{x}_N = \boldsymbol{b}$推得$\boldsymbol{x}_B = B^{-1}\boldsymbol{b} > 0$，故$\boldsymbol{x}$为基本可行解。
- 若$m' < m$，由A行满秩，故可从N选取$m - m'$列并入B构成新的非奇异矩阵$\bar{B}$，N中其余列向量构成$\bar{N}$，$\boldsymbol{x}_B$补充相应的零分量，以及$\boldsymbol{x}_N$中去掉相应零分量，可得到:$\bar{\boldsymbol{x}}_{\bar{B}} = \bar{B}^{-1}\boldsymbol{b} \geq 0, \bar{\boldsymbol{x}}_{\bar{N}} = 0$，所以

$$\boldsymbol{x} = \begin{pmatrix} \boldsymbol{x}_B \\ \boldsymbol{x}_N \end{pmatrix} = \begin{pmatrix} \bar{\boldsymbol{x}}_{\bar{B}} \\ \bar{\boldsymbol{x}}_{\bar{N}} \end{pmatrix}$$

为基本可行解。 ■

无界方向

定义2.1.5: 设$D \subseteq \mathbb{R}^n$为凸集，向量$\boldsymbol{d} \in \mathbb{R}^n$，若$\forall \boldsymbol{x} \in D$和$\forall \alpha \geq 0$有：$\boldsymbol{x} + \alpha \boldsymbol{d} \in D$成立，则称$\boldsymbol{d}$为集合$D$的无界方向.



对LP问题的可行域$\mathcal{F} = \{\boldsymbol{x} | A\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geq 0\}$, 由**凸多面体表示定理**知：

- 若$\mathcal{F}$有界，则任何$\boldsymbol{x} \in \mathcal{F}$都可表示为$\mathcal{F}$顶点的凸组合；
- 若$\mathcal{F}$无界，则任何$\boldsymbol{x} \in \mathcal{F}$都可表示为$\mathcal{F}$顶点的凸组合与一个无界方向的向量和。

无界可行域与最优性

引理2.1.2.

考虑标准LP问题，设可行域$\mathcal{F}$无界且最优函数值有限，则对$\mathcal{F}$的任意无界方向$\boldsymbol{d}$均有：

$$\boldsymbol{c}^T\boldsymbol{d} \geq 0.$$

成立.

证明： 设$\boldsymbol{x}^*$为标准形LP问题的最优解，由无界方向的定义有：对$\forall\lambda \geq 0$,$\boldsymbol{x}^* + \lambda\boldsymbol{d} \in \mathcal{F}$始终成立。由点$\boldsymbol{x}^*$的最优性有:

$$\boldsymbol{c}^T(\boldsymbol{x}^* + \lambda\boldsymbol{d}) = \boldsymbol{c}^T\boldsymbol{x}^* + \lambda\boldsymbol{c}^T\boldsymbol{d} \geq \boldsymbol{c}^T\boldsymbol{x}^*.$$

由$\lambda > 0$得出$\boldsymbol{c}^T\boldsymbol{d} \geq 0$的结论。 ■

定理2.1.3 顶点最优性.

若标准形LP问题存在有限最优解，则必可在其可行域的某个顶点取得。

证明： 设可行域顶点集合为$V=\{\boldsymbol{v}_1,\boldsymbol{v}_2,\cdots,\boldsymbol{v}_k\}$. 假定$\boldsymbol{v}_j$是最优顶点，即有$\boldsymbol{c}^T\boldsymbol{v}_j=\min\{\boldsymbol{c}^T\boldsymbol{v}_i|\boldsymbol{v}_i\in V\}$.

下证顶点$\boldsymbol{v}_j$为最优解。对$\forall\boldsymbol{x}\in\mathcal{F}$, 根据可行点表示形式$\boldsymbol{x}=\sum_{i=1}^k\alpha_i\boldsymbol{v}_i+\boldsymbol{d}$, 其中$\boldsymbol{d}$或为零，或为$\mathcal{F}$的一个无界方向。根据前面**引理2.1.2**, 顶点$\boldsymbol{v}_j$在为最优顶点,以及$\sum_{i=1}^k\alpha_i=1,\ \alpha_i\geq 0,\ \ i=1,2,\cdots,k$得:

$$\boldsymbol{c}^T\boldsymbol{x}=\sum_{i=1}^k\alpha_i\boldsymbol{c}^T\boldsymbol{v}_i+\boldsymbol{c}^T\boldsymbol{d}\geq\sum_{i=1}^k\alpha_i\boldsymbol{c}^T\boldsymbol{v}_j=\boldsymbol{c}^T\boldsymbol{v}_j.$$

由$\boldsymbol{x}$的任意性，证明了顶点$\boldsymbol{v}_j$为最优解。 ■